Atti

1. *Il controllo terminologico delle risorse elettroniche in rete: tavola rotonda, Firenze 27 gennaio 2000*, a cura di Paola Capitani, 2001

2. *Commemorazione di Michele Della Corte,* a cura di Laura Della Corte, 2001

3. *Disturbi del comportamento alimentare: dagli stili di vita alla patologia*, a cura di Corrado D Agostini, 2002

4. *Proceedings of the third International Workshop of the IFIP WG5.7 Special interest group on Advanced techniques in production planning & control : 24-25 February 2000, Florence, Italy*, edited by Marco Garetti, MarioTucci, 2002

5. *DC-2002: Metadata for E-Communities: Supporting Diversity And Convergence 2002: Proceedings of the International Conference on Dublin Core and Metadata for e-Communities, 2002, October 13-17, 2002, Florence, Italy*, organized by Associazione Italiana Biblioteche [et al.], 2002

6. *Scholarly Communication and Academic Presses: Proceedings of the International Conference, 22 March 2001, University of Florenze, Italy*, edited by Anna Maria Tammaro, 2002

7. *Recenti acquisizioni nei disturbi del comportamento alimentare*, a cura di Alessandro Casini, Calogero Surrenti, 2003

8. *Proceedings of Physmod 2003 International Workshop on Physical Modelling of Flow and Dispersion Phenomena*, edited by Giampaolo Manfrida e Daniele Contini, 2003

9. *Public Administration, Competitiveness and Sustainable Development*, edited by Gregorio Arena, Mario P. Chiti, 2003

10. *Authority control: definizioni ed esperienze internazionali: atti del convegno internazionale, Firenze, 10-12 febbraio 2003*, a cura di Mauro Guerrini e Barbara B. Tillet; con la collaborazione di Lucia Sardo, 2003

11. *Le tesi di laurea nelle biblioteche di architettura*, a cura di Serena Sangiorgi, 2003

12. *Models and analysis of vocal emissions for biomedical applications: 3rd international workshop: december 10-12, 2003 : Firenze, Italy*, a cura di Claudia Manfredi, 2004

13. *Statistical Modelling. Proceedings of the 19th International Workshop on Statistical Modelling: Florence (Italy) 4-8 july, 2004,* edited by Annibale Biggeri, Emanuele Dreassi, Corrado Lagazio, Marco Marchi, 2004

14. *Studi per l insegnamento delle lingue europee : atti della prima e seconda giornata di studio (Firenze, 2002-2003),* a cura di Mar a Carlota Nicol s Mart nez, Scott Staton, 2004.

15. *L Archivio E-prints dell Universit di Fir enze: prospettive locali e nazionali. Atti del convegno (Firenze, 10 febbraio 2004)*, a cura di Patrizia Cotoneschi, 2004

16. *TRIZ Future Conference 2004. Florence, 3-5 November 2004*, edited by Gaetano Cascini, 2004

17. *Mobbing e modernit : la violenza morale sul lavor o osservata da diverse angolature per coglierne il senso, definirne i confini. Punti di vista a confronto. Atti del Convegno Firenze, 20 aprile 2004*, a cura di Aldo Mancuso, 2004

18. *Lo spazio sociale europeo. Atti del convegno internazionale di studi Fiesole (Firenze), 10-11 ottobre 2003*, a cura di Laura Leonardi, Antonio Varsori, 2005

19. *AIMETA 2005 Atti del XVII Congresso dell Associazione Italiana di Meccanica Teorica e Applicata, Firenze, 11-15 settembre 2005*, a cura di Claudio Borri, Luca Facchini, Giorgio Federici, Mario Primicerio, 2005

20. *Language Teacher education and Training: Italy and Europe. Educazione e formazione dei docenti L2: Italia e Europa. CEFTrain Day - Giornata CEFTrain. Trans-european Contributions — Contributi trasnazionali. Firenze, Italy, 7 may 2005*, edited by Elizabeth Guerin, 2005

21. *Models and analysis of vocal emissions for biomedical applications. 4th international workshop. october 29-31, 2005 : Firenze, Italy,* edited by Claudia Manfredi, 2005

University of Florence
Department of Systems and Informatics Distributed Systems
and Internet Technology Laboratory

Axmedis 2005
Proceedings of the 1st International Conference on Automated Production of Cross Media Content for Multi-channel Distribution

Volume for Workshops, Industrial and Applications Sessions

Florence, Italy, 30 November—2 December 2005

Edited by

Paolo Nesi
Kia Ng
Jaime Delgado

Firenze University Press
2005

Axmedis 2005 : proceedings of the 1st International Conference on Automated Production of Cross Media Content for Multi-channel Distribution : Volume for Workshops, Industrial and Applications Sessions : Florence, Italy, 30 November—2 December 2005 / edited by Paolo Nesi, Kia Ng, Jaime Delgado. — Firenze : Firenze university press, 2005. (Atti, 22)
http://digital.casalini.it/8884533546
Stampa a richiesta disponibile su http://epress.unifi.it
ISBN 88-8453-354-6 (online)
ISBN 88-8453-355-4 (print)

658.514 (ed. 20)
Produzione - Automazione

Universit degli Studi di Firenze
Firenze University Press
Borgo Albizi, 28, 50122 Firenze, Italy
http://epress.unifi.it/

Printed in Italy

AXMEDIS 2005
www.axmedis.org/axmedis2005

Convitto della Calza, Firenze, Italy
30 November — 2 December 2005

Sponsored by

Distributed Systems and Internet Technology Lab, Department of Systems and Informatics, University of Florence (Italy)

Accademia Nazionale di Santa Cecilia Fondazione (Italy)

ACit GmbH (Germany)

Associazione dei Fonografici Italiani (Italy)

Comverse Ltd (Israel)

Consorzio Pisa Ricerche (Italy)

Dipartimento di Italianistica, University of Florence (Italy)

Ecole Polytechnique Federale de Lausanne (Switzerland)

Eutelsat S.A. (France)

Exitech S.r.L. (Italy)

Fraunhofer Gesellschaft zur Foerderung der Angewandten Forschung E.V. (Germany)

Fundació Universitat Pompeu Fabra (Spain)

Giunti Interactive Labs S.r.L. (Italy)

Hewlett Packard Italiana S.r.L. (Italy)

On Demand Distribution Ltd (UK)

Sejer Representing Bordas and Nathan (France)

CRS4 – Centre for Advanced Studies, Research and Development in Sardinia (Italy)

European Commission
Directorate-General
Information Society and Media Content
Knowledge and Content Technologies

International Association of Interactive Multimedia
MUSICNETWORK
www.interactivemusicnetwork.org

Conference Organising Committee

Programme Committee

M. Arnold (Fraunhofer-IGD, DE)
A. Badii (University of Reading, UK)
D. Bainbridge (Univ. Waikato, NZ)
J. Barth lemy (IRCAM, Fr)
P. Bellini (University of Florence, DSI, I)
F. Bixio (AFI, I)
H. Brandl (Infineon, DE)
W. Buhse (CoreMedia, DE)
A. Camurri (University of Florence, Genoa, I)
D. Carboni (CRS4, I)
G. Cesaretti (CPR, I)
B. Co asnon (IRISA/INRIA, F)
D. Crombie (FNB, NL)
R. Dannenberg (CMU, USA)
J. Delgado (UPF, E)
U. Dietz (Vodafone, DE)
C. Dorai (IBM, USA)
I. Fujinaga (McGill University, CA)
D. Fuschi (ILABS, I)
S. George (University of South Australia, AU)
M. Girer (LISTESSO S rl, F)
R. Grimm (Technical University of Ilmenau, DE)
F. Hartung (Ericsson Research, DE)
K. Hirata (NTT, J)
P. Hobson (Motorola, UK)
P. Kafno (HD Thames/University of Nottingham, UK)
T. Kalker (HP, USA)
A. Karydis (ANALYSIS Systems Research High-Tech S.A., G)
D. Keislar (CMJ & Musclefish, USA)
P. Kruger (TISCALI, I)
E. Lambert (Silver Lining Project, I)
A. Lindsay (University of Lancaster, UK)
S. Llorente (UPF, E)
M.S. L pez-Lago (Fundaci n Alb niz, E)
C. Marangoni (HP, I)
M. Mattavelli (EPFL, Sw)
P. Nesi (University of Florence, DSI, I)
K. Ng (ICSRiM, University of Leeds, UK)
R. van Noetsele (Chello Broadband, NL)
J. Nuetzel (Technische University of Ilmenau, DE)
L. Pearce (XIM, UK)
M. Russ (BT Exact, UK)
M. Schmucker (FHGIGD, DE)
P. Stollenmayer (Eurescom GmbH, DE)

Table of Contents

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

Workshop of MUSICNETWORK on MPEG Symbolic Music Representation

Panel on the Role of Collecting Societies in the Digital Era

Organiser: Isabella Longo, International coordinator of AFI Italian Association of Phonographic Producers

Workshop on Virtual Goods Technical, Economic and Legal Aspects of Business Models for Virtual Goods

Message from the Chairs

Current developments in the cross-media domain require innovative and new technologies to meet the challenges of the market. The AXMEDIS conference aims to promote discussions and interactions among researchers, practitioners, developers and users of tools, technology transfer experts, and project managers, to bring together a variety of participants from the academic, business and industrial domains in order to address all relevant technical and commercial issues. Particular interests include the exchange of concepts, prototypes, research ideas, industrial experiences and other results. The conference focuses on the challenges in the cross-media domain (which include production, protection, management, representation, formats, aggregation, workflow, distribution, business and transaction models), and the integration of content management systems and distribution chains, with particular emphasis on cost reduction and effective solutions for complex cross-domain problems. The conference is supported by the AXMEDIS Consortium which consists of digital content producers, integrators, aggregators, distributors, information technology companies and research groups involved in content production, protection and content distribution via a variety of different channels including interactive TV (i-TV), DVBT, DVBS, personal computer, kiosk, mobile phone, PDA and others.

This is the first AXMEDIS conference and it has inherited the experience and the community from the WEDELMUSIC (Web Delivering of Music) international conference series and the MUSICNETWORK Open Workshops that have been successfully held for some years. With the combined effort, cross-fertilisation, and expansion, the new AXMEDIS international conference series continue to grow, to improve, and to enlarge the scope and the communities in size, depth and breath. The program committee has received a large number of paper submissions for research, industrial, poster presentations, and many workshop, panel and tutorial proposals. The selection has not been easy due to the high quality of submissions and the limited time-slots available. This has created a very dense and interesting technical programme. It starts with the keynote lecture of Dr. Leonardo Chiariglione (the *father* of MPEG, DMP and many other challenging international activities) and includes a large number of scientific and industrial presentations, together with workshops and panels of discussion. For example, a workshop on MPEG Symbolic Music Representation organises by the MUSICNETWORK, a workshop on Role of collecting societies in the digital era or ganised by Associazione Fonografici Italiani, a workshop on Technical, Economic and Legal Aspects of Business Models for Virtual Goods, and a panel on European Accessible Information Network , etc. This volume of proceedings is devoted to these activities and to industrial presentations.

We are very grateful to many people without whom this conference would not be possible. Thanks to old and new friends, collaborators, institutions, organisations, and the European Commission, who have supported AXMEDIS. A special thanks to Prof. A. Marinelli, and Dr. P. Vigevano for opening the conference, and to Dr. Leonardo Chiariglione for his opening speech. Thanks to members of the Program Committee for their invaluable contributions and insightful work. Thanks to IEEE Computer Society Press for the organisation of this proceedings, and many thanks to the people behind the organisation of the event, including Dr. S. Ceglia, Nicola Mitolo, and many others. Last but not least, many thanks to all participants of AXMEDIS 2005. We look forward to seeing you at AXMEDIS 2005 and all the future AXMEDIS and MUSICNETWORK activities.

General Chair: Paolo Nesi, University of Florence, Florence, Italy
Program Co-Chairs: Kia Ng, ICSRiM, University of Leeds, Leeds, UK
Jaime Delgado, Universitat Pompeu Fabra, Barcelona, Spain
Atta Badii, IRC, University of Reading, UK
Claudio Marangoni, HP, Italy
Laurence Pearce, xim Ltd, UK

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6 © 2005 Firenze University Press

AXMEDIS 2005

Industrial and Application Papers

Coordinated by

Paolo Nesi, University of Florence, Italy

Kia Ng, ICSRiM, University of Leeds, UK

Jaime Delgado, Universitat Pompeu Fabra, Barcelona, Spain

Atta Badii, IRC, University of Reading, UK

Claudio Marangoni, HP, Italy

Laurence Pearce, xim Ltd, UK

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

Section:
Cultural Heritage

AN ARCHIVE OF MULTIMEDIA OBJECTS ON-LINE USING WEDELMUSIC: THE EXPERIENCE OF ARCIPELAGO MUSIC STUDY CENTER

Dr. ROBERTO LONOCE
Director Multimedia Music Library
Centro Studi Arcipelago Musica
roberto@lonoce.com

Abstract

The experience of Arcipelago Musica Study Center, a non-profit organization located in Milan, Italy. Arcipelago Musica has a Mediateque, specialising in Nineteenth Century Music. We have specific agreements with several music publishers for using our content in the mediateque. In 2002 we decided to migrate to WEDELMUSIC for its innovative functionalities and characteristics. Here we describe our experiences with the "200 Italian composer" project, an online archive of multimedia objects.

To distribute scores online, Arcipelago Music has stipulated a contract with the principal Italian editors and music publishers, bound to the project "200 Italian composers: 1950-2002" ,for the free supply and use. This is possible, thanks to the guarantees of protection offered by Wedelmusic. Despite the important innovations of the Wedelmusic package, some careful evaluations for improvements and future developments have been suggested.

1. Introduction: What is Arcipelago Musica.

Arcipelago Musica is a non-profit organization located in Milan, Italy, in the Fondazione Enrico Mattei of ENI, corso Magenta, Palazzo delle Stelline.

We have been active since 1998 in promoting cultural activities in the Musica area and in managing a multimedia music mediateque. We organise events and meetings where modern and contemporary music is promoted.

Our main funding comes from several sources, including: Regione Lombardia, Provincia di Milano, Fondazione Cariplo, Fondazione Stelline, HP Italia, Philips Italia, Fondazione Enrico Mattei, International Society for Contemporary Music, Conservatory of Milan, Museum of Musical Instruments in Castello Sforzesco in Milan.

In 2000, Arcipelago Musica inaugurated the Mediateque of Music, specialising in Nineteenth Century music. On some multimedia stations it was possible to browse the catalogue of the inserted authors, to choose the composition and to listen the audio tracks while reading the score. The mediateque was supported by a WEB based solution for the fruition of multimedia music content. We have specific agreements with several music publishers for using our content in the mediateque: BMG Ricordi, Rugginenti, Sonzogno, Curci, Suvini Zerboni, Warner Carisch and others. In 2002 we decided to migrate to WEDELMUSIC for its innovative functionalities and characteristics.

In 2003 another important initiative, in the context of the didactic activities, is the "Coro dei Ragazzi della Città di Milano", born in collaboration with the Comune di Milano - Assessorship to the Social Services and with the Civic Schools Foundation in Milan. The Choir has already begun a new and alternative musical experience compared to many existing boy choirs.

In the current year, Arcipelago Musica has an agreement with the Museum of the Musical Instruments of the Castello Sforzesco. In the museum it is possible to follow guided tours for the public of all ages, and it is possible to listen to live performances of the musical instruments.

2. Innovations and potentialities of WedelMusic's package

In the beginning our mediateque library included about 100 works of Italian composers with the "opera omnia" of Anton Webern. This music catalogue was managed by a software package, which allowed viewing of the score and the possibility to listen to the related audio. The software was limited in that it did not evolve or permit an Internet connection.

Our desire was to make our music content available for many users. We needed a software package, which allowed the protection of the objects without violating copyrights of the composers, editors and record companies.

Wedelmusic is an innovative idea to bring the music to the Internet era, completely respecting publishing rights and protecting them from copyright violation. It is possible to hear the audio tracks and view the score, respecting the rights of author and copyright laws. However the scores and the audio tracks cannot be copied or printed.

The techniques for the protection of music by Wedelmusic refers to tools that allow the safe transmission of Wedel objects and the insertion of identifiers into the music in order to demonstrate the owner, the musical area and the distributor. Wedelmusic enables publishers and users to manage their music interactively protecting their intellectual property rights. It is not possible to fully trust every user or control his or her applications on a computer. With the Wedelmusic Server, the basis of the Wedelmusic package, music publishers such as Arcipelago Musica, can store, manage and finally distribute music on the Internet. In fact the Local Distributor, another tool within the Wedelmusic package, can connect to the Server Database via Internet, browse the music catalogues, select the Wedel object to study and finally download them, all strongly protected with security mechanisms.

3. The "200 Italian composers" project

In 2003-2004 the multimedia library of Arcipelago Musica was broadened with the project "200 Italian composers from 1950 to 2002, an anthology of Italian classic music from the last half century ". Five important Italian classical musicians, Mario Ancillotti, Bruno Canino, Giuseppe Garbarino, Enzo Porta and Gabriella Ravazzi have selected 200 composers from a list of all Italian composers who had worked from 1950 to 2002. Each of these 200 composers were contacted (in the case of defunct composers, the family was contacted) and asked to choose a representative composition of the musical language and creativity.

Wedelmusic is a digital object, which can include several different components covering all aspects of a music piece such as audio files, music sheets in symbolic format or images of music scores, documents, videos, animations, and images in any chosen format. Each Wedelmusic object of the "200 Italian Composers" project is made up of an Italian Classification, MP3 audio file, the score in Acrobat pdf (only visible with Acrobat Writer) the composer's biography and photo. Users who do not have the Acrobat Writer application installed, can see the score with a viewer that browses the tiff images of the score. To access to the various components it's sufficient to click on the component of the object.

When you open the editor and ask for a component, the program verifies the registration online through a connection with the WebServer. Only when the verification is successful, is it possible to visualize the components. In this way the safety of data and the fruition facility is guaranteed. The unloaded components can be seen again, without having to repeat the download, after the usual verification of the user's personal account.

Thanks to permission received from the music publishers, the scores may be visualised only. Printing is not permitted, even in low resolution. In the same way the audio files can be listened to, but they cannot be saved.

The end-user who downloads a component of Wedelmusic object cannot share it with other users who are not registered.

The Catalogue with Local Distributor is aimed to give end-users tools to browse and search for the desired objects. A first-time user can browse the Catalogue (by Composer, by Genre, or even by Publisher).

This first way of access to the Catalogue can give him/her a general idea of the content. To access the Catalogue, every end-user must be registered. If an end-user is not registered, he/she must ask the Administrator of the Library, who is the only person authorized to create end users' accounts.

When accessing the Catalogue, a login and a password will be requested. From the main Catalogue page that is accessed after a successful login, end-users can browse the Catalogue, search in the Catalogue, and access personal functions.

The Catalogue is composed of objects, each of these with "classification record", that is to say, information, which enables an easy search and retrieval process. "Classification record" lists

information such as title, name of the composer, date of composition, ISMN number, and much more.

Each object is also composed of one or more components. These components can be, for instance, the score itself, but also an audio record, a text file, and image file and so on. By clicking on the “Editor” button beside the title, the end-user can access the content of the object himself, that is to say, the score, audio files and so on.

Browsing the Catalogue is the main way to access objects stored in the Database.

End users may browse the Catalogue either by Composer, by Publisher, or by Genre. Each of these finally leads to a list of objects stored in the Catalogue, with details on each objects.

In the classification of all object, like genre, we have inserted "200 Italian composers 1950-2002". Despite the potentiality of the Local Distributor as a search engine, we have inserted the data in alphabetical order by composer’s name, which we evaluated to be much simpler for users considering the relatively small amount of information involved.

In our website, www.arcipelagomusica.it, to click on " 200 Italian composers" you enter a html page, which allows you to unload a small manual in Italian for the registration and, subsequently, "to enter" the library. Nine pages, in alphabetical order, show all the available objects for the download. The table is structured as follows: Composer Name, small photo, date of composition, Wedel object, publisher and record company. To click on the grey button with the name of the composition opens the editor and allows the user to browse digital contents.

4. Problems with the WedelMusic Editor and proposals of improvement

Despite the great innovations of the Wedelmusic package previously described, some careful evaluations are to necessary for improvements and future developments.

- The recipients of Wedel objects are usually musicians, music students and researchers, none of whom are computer experts! This requires various considerations. Initially the complexity of installation. Many end users have reported difficulty in the registration. Although there is a small manual available online, many users stop at the initial web pages without unloading the editor and so on.
- Another problem is the difficulty with English in the installation phase, in the use of editor and Local Distributor, the downloading windows etc. For future developments the possibility of multi-language interfaces is of paramount importance.
- As previously mentioned, the main users of the package are musicians, who frequently use Mac OSX platform. Currently Wedelmusic doesn't foresee an Editor for Mac, therefore for future developments and it is necessary to make Wedel objects available to these users.
- The scores of compositions are in Acrobat pdf format. Not all users have installed Acrobat Writer, which allows the user to browse the pages. In alternative a Wedel display device is present, which visualizes the sequence of the tiff images of the score, components insert to purpose; however it doesn't allow all the functionalities of the Acrobat program. A more efficient display device and with greater functionalities is desirable.

For future projects that will extend the musical contents to digital video contents, it will need to develop an extension of the metadati and the study of a suitable graphic interface.

5. Rights and permission of Music Editors and Record Company owner- Legal aspects related to digital content

The consultation of paper and audio documents in electronic format inside a library, in a place adjacent to the physical archive that collects such original documents, doesn't cause particular copyright problems. In fact the electronic copy of the library’s patrimony - if it is consulted in intranet - is considered by current laws in Italy to be a simple tool to facilitate and consult and simultaneously helps the conservation of the document.

The diffusion outside the library of the electronic format of paper and audio documents sets several and more complex problems of authors’ rights. Arcipelago Music has been able to offer to the consumers Internet’s consultation of the anthology" 200 Italian composers: 1950-2002" only after several agreements with publishers, record companies, performers, composers and with the Society of Italian Authors and Editors (S.I.A.E.), and thanks only to the guarantees of protection offered by Wedelmusic. In fact the software allows the complete consultation of the documents by the registered user, but it prevents any download and print.

To distribute scores online, Arcipelago Music has stipulated a contract of free supply and use with the principal Italian musical publisher, bound to the project" 200 Italian composers: 1950-2002" and to the

guarantees of protection offered by Wedelmusic. Up to now BMG Ricordi, Suvini Zerboni, Carisch, Sonzogno and Rugginenti have signed. The management of all scores' rights is in fact delegated to the publishing house.

The management of rights of audio documents for diffusion in Internet is much more complex, because it involves more entities: the record publisher or the performers (in case of unpublished document) and the S.I.A.E.

Therefore Arcipelago Music has the permission of every involved record publisher or, if the document was unpublished, of the performers (or the director in the ensemble case). This permission is also bound to the project" 200 Italian composers: 1950-2002" and to the guarantees of protection offered by Wedelmusic.

Because of the crisis that currently involves the whole music industry, classical music in particular, it is very difficult to get this permission, despite the fact that in this case the project promotes Italian contemporary music in the world.

To distribute online audio documents protected by the copyright (compositions by living composers or dead from less than 70 years) a S.I.A.E license is necessary, obtained through payment of a monthly fee. The Society Italian Authors and Editors has set in the autumn 2004 a new type of license, which is more suitable for projects with the aim of culture's promotion and excluded from any economic profit, that however foresees the payment of an annual forfait of around 550 Euro!

Arcipelago Music has asked the Society of Italian Authors and Publishing to sponsor the initiative, granting consequently a free license.

Naturally every composer or his heir has also been invited to personally join the project with a written declaration, including his best work, too.

6. Two synthetic characteristics of the "200 Italian composers" project

- With this project Arcipelago Musica has begun establishing important relationships with the musical world. In this field there is a particular need for this type of activity because there are not many tools available to get to know contemporary classical music.
- In Arcipelago Musica there are different people involved, each one specialised in a different aspect: musical composition, music history, copyright, relationships with the music world and music publishing industry, management, formulation of projects that will be proposed to public and private sponsors, use of technology tools and the choice of software and hardware.

The knowledege and experience of the numerous people who have worked together, as well as the decisive help of Wedel Music, have made this project possible. This project is just the beginning and for some months we have been working on extending it to other fields connected with classical music of the last century and the present day. We hope that this can help people to enjoy multimedia digital objects in different ways and places

European Projects for Art Schools: The MultimediArt Project

Elisabetta Delle Donne
Pixel
E-mail: eli@pixel-online.net

Abstract

The MultimediArt project is a successful example of the effective use of ICT in the field of art teaching and learning at secondary school level. The project promoted the development of new technical skills which were acquired by art teachers and students.

1. Introduction

European fine arts are among the best known all over the world and constitute one of Europe's main assets. Each European country has its own artistic heritage, some of them are well known, others are less known in the rest of Europe.

It is important to promote access to European artistic heritage first of all to Europeans themselves. This can be done effectively through the implementation of ICT based solutions.

In the last few years the European Commission funded three projects promoting the sharing of information about European art between schools.

The projects were financed in the framework of the Socrates and elearning programmes, both aiming at the promotion of an effective use of ICT in school education.

2. The Context

The MultimediArt project started from the positive experience gained from a previous project, entitled ARTE (http://socrates-arte.net), which was financed by the Socrates Programme ODL action (now Minerva). The objective of the ARTE project, which was concluded in December 2001, was the exchange of information and material about contemporary art in Europe among secondary schools of 6 countries. The sharing of sources took place through the Internet.

Some of the teachers involved in the ARTE project tried experimenting with the use of technology not only for communication and research (in line with the objectives of the project) but also for artistic creation and they verified that the number of artists who are turning to new technology and to multimedia for research and for creating art is growing across the world. The Internet, new technology and multimedia are therefore contributing to the adoption of a new method of art expression.

From here, the idea at the heart of the MultimediArt project was born: technology can serve a teacher of art, providing them with new tools for artistic expression.

The MultimediArt project, promoted by Pixel, is financed by the European Commission in the framework of the Socrates Programme Minerva Action.

The MultimediArt project started in January 2002 with two main objectives:

- To inform art teachers in secondary schools about new forms of artistic expression based on the use of technology
- To promote the use of new technology in the art creation process, training, at distance, secondary school art teachers.

3. The main project activities

The main activity of the MultimediArt project was the distance training course on the theme of the application of new technology in the process of artistic creation. The course was accessible via Internet on the project website.

The course was addressed to secondary school teachers.

The course was structured in three modules:

- The new teaching technology
  - Internet as an aid for the teaching of art, research and socialising
- Art and new technology
  - Technology as an aid to traditional artistic expression
- Technology and new arts
  - Technology as a means of alternative or traditional artistic expression

As a support to the distance training, a Forum section was available in which it was possible to leave

questions for the trainers of the course or to initiate topics for discussion. Furthermore, as a support to the distance training, monthly on-line meetings were organized, in which the teachers involved in the project could participate and they were able to meet virtually with the teachers of the course.

The pedagogical methods used in the distance course have put the creation of multimedia artwork at the centre of the teaching. Departing from an introduction of the employed techniques, with reference to previous "traditional" art, so as to arrive at a series of examples, practical tutorials and pieces of information on the software used.

Learning how to use one or more specific software was therefore not the objective of the training, which would mean putting something forward that already exists and that one can download, usually for free, from the web. The objective was rather to make the artistic potential of such software known, whilst only studying a few aspects in depth. It was then up to the teacher to decide which of the suggested tools was the most congenial to their needs for artistic expression and therefore which of the suggested tutorials they decide to study in depth and download.

5. Problems and solutions

The schools involved in the project experienced a number of problems in the carrying out of the project activities. Here are some of them together with the solutions identified:

- Language barrier: being involved in the two European projects, the teachers had to communicate in English. Few art teachers could speak, write or understand English. The solution adopted was to involve also the English teacher in each of the schools together with the art teacher. This also fostered an inter disciplinary approach to the project.
- Technical divide: the technical tools available in each of the schools involved were very different as well as the technical skills of the participating teachers. We discarded the original idea to find an homogeneous solution for all and encouraged the teachers to try and make the most effective use of the technologies they were familiar with. As a consequence, the less experienced teachers learned from their colleagues who were more technically skilled.
- Management of distance virtual meetings: the first distance meetings were quit confusing because the participating teachers were contributing all at the same time and therefore it was difficult to follow the line of the discussion. The solution adopted was the creation of step by step guidelines for the participation in distance meetings. This resulted in a disciplined and efficient participation in the meetings.

6. The main project results

The results of the MultimediArt project were:

- An online database of the schools involved containing, for each school, a presentation of the institute with all contact details and photographs of the teachers and students working on the project.
- The Multimedia Art Gallery divided into two sections:
    - Artists (with the publication of the works and interviews from the authors)
    - Students (with the publication of their finished works and details of the techniques used)
- A virtual library that gathers information on the multimedia art in Europe with links to art galleries and museums across Europe and with teaching materials for artistic education developed by the teachers involved in the project.
- Distance courses on the theme "Art and New Technology"
- Forum and Chat, which provide points of contact between teachers and students of art in Europe.

5. Benefits for the schools involved

The art teachers involved in the MultimediArt project benefited from the project because they:

- Discovered multimedia art;
- Were trained on the use of new technology for realizing artistic multimedia expression;
- Collaborated with other schools across Europe to make exchanges (of methods, materials and people);
- Contributed to creating the multimedia artists of tomorrow.

New Services for the Public in a Technology-related Approach: the AXMEDIS Project Inside Accademia Nazionale di Santa Cecilia

Annalisa Bini, Roberto Grisley, Tiziana dell'Orto
Accademia Nazionale di Santa Cecilia
a.bini@santacecilia.it, r.grisley@santacecilia.it, tdellorto@yahoo.it

Abstract

European musical institutions started the process of storing their musical content in digital formats. This huge European heritage has the common need to be managed, distributed and valorised.

The containment of content sale prices is a key element when setting up a business venture in the digital cross media content as well as the increase of accessibility to the contents is a key element to create a better exploitation of the digital heritage.

This paper presents a brief introduction to the AXMEDIS project an European funded project that will create an innovative technological framework for the protected automatic production and distribution of multimedia contents and discusses the new applications and exploitations in the library and museums context. For further details see the project website www.axmedis.org.

1.Introduction

Music content is abundantly stored in European musical institutions in digital formats.
In particular Libraries and museums are fast moving to the digitalisation of their contents and to the use of the information technology. The world of libraries has been the pioneer in the introduction of the ICT: creating standards for cataloguing first, and, more recently, with the creation of digital collections.
These institutions have collections which include: music sheets, audio, documents, videos, images, etc. They have hundreds of thousands of digital items that could be exploited for commercial purposes. For example, this digital content covers more than 90% of the needs of musicians that buy music scores and musical. Now digitization is no more just an activity for preservation: this cultural European heritage has the common need to be managed, distributed and valorised.
The lack of adoption of a suitable technology and business model is slowing down its valorisation and exploitation.

The containment of sale prices is a vital key when setting up a viable and sustainable business venture in the digital cross media content. On the other hand the increase of accessibility to the contents is a key element to create a better exploitation of the library's and museum's heritage.
Possible solutions to this challenges could be found by automating, accelerating and restructuring managing and delivering processes, and providing solution to the content protection. Such solutions will enable the managing and delivering processes to be faster and cheaper, while at the same time providing new capabilities to support safer distribution.

2. The experience of Accademia Nazionale di Santa Cecilia in introducing the ICT

In 1997 Accademia Nazionale di Santa Cecilia (ANSC) started the digitalization of its heritage to create a multimedia library. The Multimedia Library of Accademia Nazionale di Santa Cecilia holds a huge collection of invaluable heritage contents, from late XVth century to the present day. The historical archive contains many different forms of content including documents, audio recordings, photographs and others. The library preserves many original manuscripts (particularly from the XVIIth-XXth Century) and printed editions.

Within the summer of 2005, 120.000 pages of the contents and a part of the audio-video sources will be digitised. The digitalization process and the use of the ICT has already improved several internal activities of the library.
Within the end of December the multimedia catalogue of the digitized contents will be available on internet.

The Museum of musical instruments and musical iconography of Accademia Nazionale di Santa Cecilia holds an important collection of ancient and modern instruments.
Since its inauguration in 1895, the so called "Museum of Historic and Modern Instruments" has been characterized by the extreme diversity of both cultured and ethnic European and non-European

instruments it houses.
In 1900, there were 77 instruments in the collection, some purchased by the Accademia and others donated by antiquarians and collectors. Today, there are 270 instruments and about 150 other items (pictures, musician portraits, curious etc.). Three groups in the collection attract particular attention for their importance and interest:

1. the 1926 legacy of Queen Margherita of Savoy (26 examples, mainly plucked instruments, some extremely rare and preciously decorated);
2. two stringed quartets and other single instruments entered in the national competitions for stringed instruments makers organised by the Accademia between 1952 and 1956;
3. the donation of Gioacchino Pasqualini, violinist and researcher in acoustic physics who was museum curator in the 1960's and founder of the Associazione Nazionale Liuteria Artistica Italiana (ANLAI). The collection contains numerous items including important bowed stringed instruments.

The Museum's best-known and most important item deserves special mention: the Stradivari violin from the Mediceo Quintet (1690), known as "Il Toscano", which was purchased by the Accademia in 1953.

Beginning in 1993, the collection has been systematically catalogued and, where necessary, carefully restored. The technical drawings, pictures of the instruments, the images used for the restoring (x-rays, ultraviolet pictures…) and the other items (pictures, musician portraits, curious etc.) has been digitised.

The next steps to a full appreciation of the collection is its exploitation through the use of the ICT (see later on) and its exposition in the new Accademia Nazionale di Santa Cecilia premises (scheduled in 2006).
In 2002 the Accademia Nazionale di Santa Cecilia moved the concert seasons together with the museum's collection, the archives (historical documents, photos, ethnomusicological collection) and the library to a new residence: the new Auditorium of Rome. Built on a project by Renzo Piano, the new Auditorium has three concert halls, several rehearsal halls, exposition spaces, rooms to host the ANSC musical instrument museum and the ANSC multimedia library, shops and restaurants.

A big opportunity for ANSC to exploit the digital content of the museum and the multimedia library came from the participation to the AXMEDIS project.

3. AXMEDIS

The AXMEDIS initiative is funded by the European Commission to create and explore innovative technological framework for automatic production and distribution of cross-media contents over a number of different distribution channels (e.g. networked PC, PDA, kiosk, mobile phone, i-TV, etc) with DRM. In the context of the museums market, AXMEDIS aims to offer innovative solutions and tools to:

- manage and distribute and share digital content, such as audio-visual materials (video/film), images, documents, games, and others, in a protected and verified manner, over many different distribution channels including Internet, mobiles devices, PDA, PC, i-TV, satellite and others;
- increase the visibility and accessibility of content with the realisation of tools for content sharing among content owners. This allows the content to reach distant users with access to larger markets;
- offer additional and relevant sales channels that can simplify content distribution at a reasonable cost for end-users;
- increase both the safety and reliability with the protection models to ensure verifiable and protected delivery the objects to content producers and distributors;
- increase the accessibility of European audio-visual content;
- provide new international business opportunities to all the related SMEs in the areas of cross media content production, aggregation and distribution;
- allow end-users to gain access to the contents at a reduced costs. This will be realised by exploiting the AXMEDIS infrastructure which will open paths for new services for industrial content exploitation and for both public and corporate clients (archives, schools, museums, etc). It will also create low cost distribution chains of digital material for entertainment, education, e-commerce, etc. At the same time, this will accelerate the process of digitisation of contents for archives with reduced production costs, and enhance the value of the cultural heritage by facilitating the exploitation of the archives in digital form.

4. AXMEDIS Consortium and Potential Users

The AXMEDIS consortium consists of leading European digital content producers, integrators, aggregators, and distributors, together with information technology companies and research groups. The consortium has important resources and complementary skills which will have an effective impact upon the industry. It will also demonstrate the value of the project outcomes and the reliability and effectiveness of the project results to a wide range of potential users, including:

- museums, archives, institutions, schools and content producers;
- associations of content producers;
- Publishers and digital content providers;
- Content integration and design, audio and video;
- Networks, broadcaster and their technology providers for i-TV, PC, etc.; Mobile distributor for GSM cells or UMTS, etc.;
- Content distributor operators and technicians towards PC on internet.

5. How AXMEDIS works

The AXMEDIS Framework manages objects. In this context, every object here is a digital container for some digital content. Depending on the ownership, each museum has the right to produce licenses which are modelled as profiles for the use of the content (i.e., print, play, save, time limited use, etc., to control the access and proper usage). On the base of the profile, each museum can issue licenses and establish relevant fees.

AXMEDIS is a complete framework for the normal processes required in the Museum or library domain, including management, control, processing, distribution, transaction (selling and buying), etc. With AXMEDIS, the objects are stored in a database within the institution (reachable through IP address), or in a Kiosk, and the process of digital contents transaction can be improved in several different contexts:

- in normal day-to-day operations;
- new possibility of complete/share content collections (virtually), with access to digital contents from other museums/archives and, at the same time, widen the accessibilities and availabilities of the contents.

The sharing of the content will increase the single market and will create a wider market for all of the musical institution. The exploitation capability and potentiality of their content can increase its value when the critical mass in terms of quantity and quality is reached. This permits AXMEDIS to create a very attractive service for digital objects by putting together digital components coming from several different sources.

Music collections present music in its several forms: music sheets, audio support, music related documents, videos and images, etc. These aspects cannot be separately treated when the goal is music content valorisation. Music collections are extremely relevant parts of our cultural heritage, which presently have not been exploited. AXMEDIS will permit the exploitation of music by using emerging technologies. The interactivity and the new multimedia models satisfy the consumer needs, and the content is enriched by other experiences which could be suitable for edutainment, infotainment and e-learning. Communication and interactivity are embedded in the use of the new technologies such as for example WEB, interactive-TV, PDA, Internet communication.

6. AXMEDIS for ANSC

For the ANSC case, one of the key benefits offered by the AXMEDIS framework is the functionalities and capabilities to process and manage combinations of contents and create complex digital objects.

ANSC plans to explore usages involving:

- Raw objects, which contain just one or more digital items of the same type, like digitised photos, audio files, descriptive records, connected only by means of metadata;

- Complex objects, e.g.:
  - 2 different instruments made by the same maker, coming from 2 different museums. In this case each museum has their own licensing model
  - Catalogue: a UNIMARC (or XML) file with the descriptive record of an instrument and digital samples of the content
  - UNIMARC FILE + original instrument + modern copy

Since the Accademia has a Multimedia library and a musical instruments and musical iconography museum, there are a wide range of available contents in different formats, including archival documents. As an example, a typical Accademia complex object can be an entire archival record of a single musical instrument, and the object contains:

- The XML file of the descriptive record of the instrument containing the data on the

maker, restaurer, date, place, measurements etc…;
- The pictures of the musical instrument (e.g. in JPEG);
- The audio recording of his sound (e.g. in MP3). if available;
- The technical drawings (e.g. in JPEG, or CAD)
- Catalogues of exhibition in which the instrument has been exposed (e.g. in PDF)
- Press reviews related to exhibitions or concerts in which the instrument has been involved (e.g. in PDF)
- Maker's and owners biographies (e.g. in ASCII TEXT)
- Archival documentation related to the instrument (e.g. in JPEG)
- Restauration documentation (e.g. in JPEG and PDF)
- Documents and portraits of players (virtuoso) and instrument makers (e. g.: for Carlo Mannelli detto "del violino", member of ANSC, the portrait and documentation on his violin collection he purchased to ANSC)
- The digitised copy of a printed edition of music composed for that instrument.

With AXMEDIS, the process of the creation of a complex object could be automatic. On the other hand the content delivery process is optimised by means of different distribution channels, including PC (or kiosk), mobile, i-TV, PDA.

With AXMEDIS, the customer can go through the whole process online and receives the contents requested in real time. The ANSC staff has only to check the results of the process and does not need to manually perform all the time-consuming individual sub-tasks.
The ANSC museum or library could also provide to its customer an object made of digital contents coming from other content providers. In this case the AXMEDIS framework will automatically provide to all the content-owners their revenues in accordance to the licence agreed and contract with the museums which produced the objects. AXMEDIS will ensure also that the content distributor will receive a percentage of the income (when agreed) if the content is acquired through a distributor. All these activities are managed in a transparent manner and accessible independently from the different partners of the value chain. Thus each value chain partner may access to the AXMEDIS certifier and supervisor to enquire and receive information on the consumption of any functionality of any object.

The combinations is huge and wide ranging, even considering only musical instruments, and this is why AXMEDIS is important in supporting cross-media to allow optimised processes for museum related domain.

7. Technology-Enhanced New Services and New Possibilities

With the new possibilities resulted from the AXMEDIS framework, ANSC and European musical institutions will have the possibility to promote, manage and distribute their content on a global scale with less effort. The new technology-enhanced business model will be able to support the growth of the European content industry and to enhance the accessibilities and increase the availability of a significantly increased quantity and quality of multimedia content globally.
Possible solutions to these challenges can be found by automating, accelerating and restructuring the management and delivering and distribution processes, together with the application of content protection solution. These approaches can enhance the management and delivery processes by offering faster and cheaper services, while at the same time providing new capabilities to support a safer and protected distribution and sharing of digital content.
AXMEDIS will permit the customisation of digital objects according to different editorial and presentation formats and their distribution by using multi localized channels (such as Local Distributors to reach Personal Computers at home, Satellite Multimedia Broadcast, kiosks in relevant institutions and PDAs). All of them are mechanisms to make them more attractive and much more interesting for exploitation in different ways.

In conclusion,we can imagine different kind of museum or library activities that at present are possible and additional ones that AXMEDIS can realise reducing cost and accelerating the process.
For example, on a B2C (Business-to-Customer) scenario, the museum/library can make use of the AXMEDIS environment to support the sale of the documentation or the merchandising objects owned by the museum/library to its own customer.
The framework can also provide the sale of the contents from other institutions to its own customer.
What happens for example if a visitor of the ANSC museum wants to study and deeply compare the

different violas signed by Antonio Ciciliano? We know several Ciciliano's violas, the one in the ANSC Museum, one in the Bruxelle's collection, one in the Bologna "Museo della musica" and four in the Kunsthistorisches Museum. With AXMEDIS installed in each one of these museums, a visitor can have the complete documentation on each of the violas, photos, technical drawings etc., and he could buy copy of every document he is interested on .
The sale can happen before, during or after the visit of the customer to the museum through the use of a PDA given to the customer during the museum visit, or through the museum kiosk. The customer could also decide to buy additional documents once home, using internet.
On the other hand AXMEDIS framework will automatically ensure the correct revenue to each museums.
In addition the museum can make use of the AXMEDIS environment to realise a customised guide of the museum obtaining multimedia contents from its own or from another heritage (with the possibility to see additional documentation, hear the instrument playing…)
On a B2B (Business-to-Business) scenario, the AXMEDIS environment can be used to support the sale of the contents from its own museum/library to another institution, or to support the sale of the contents from their own museum/library to another business user.

With AXMEDIS, the new possibility will stimulate better value-for-money digital content due to effective and automated processing, production and delivery of the content using latest network technology to enable optimum interconnection and transactions between B2B and B2C, with DRM.

8. AXMEDIS Support

AXMEDIS can offer assistance and technical support to the musical institution interested in using the platform and adopting the AXMEDIS solutions. This support action will be provided through activities such as training, management, assessment and evaluation, dissemination and demonstration at conference and fairs, and affiliating them to AXMEDIS. Furthermore, the AXMEDIS consortium will grant the sum of 1 million Euro distributed by means an European competitive call to companies and research institutes interested in developing real solutions by exploiting AXMEDIS technologies.

9. Conclusion

We believe that the AXMEDIS solution will encourage not only the creation of new digital archives (based on international standards of cataloguing and descriptions (metadata)) but also stimulate the exploitation process for a wider range of digital media over many different distribution channels. AXMEDIS can introduce a new vision for the digitalisation process, encouraging the creation of digital archives for heritage preservation, as well as providing wider and better access to the important contents of the museums such as books (in electronic form) and all other types of audio-visual materials. We hope that AXMEDIS can also encourage the creation of networks of museums, libraries and archive with the framework where it will be possible to buy and sell (free or otherwise) digital contents between all partners, significantly increase the points of entrance to the contents of the museum, on a Business-to-Business model.
It is easy and beneficial for all to gain access to the AXMEDIS technologies. Over the course of the project, some didactic events will be organised to provide better understanding of the AXMEDIS technologies with further information about the potentialities of AXMEDIS. Business delegates can attend these events in order to participate in the project and bring AXMEDIS technologies to their company. Special training sessions and courses will be held for managers, content managers, content producers and integrators, and digital content distributors. Workshops and courses will be organised in several venues in Europe. To provide better understanding of the new solutions, AXMEDIS is providing a forum for discussion with technologists and experts who are ready to assist with any AXMEDIS related queries. Further information, events and calls are available online at the project website, www.axmedis.org

Acknowledgements

The authors would like to thank all AXMEDIS project partners including the Expert-User-Group and all affiliated members, for their contributions, supports and collaborations. The authors would also like to thank the EC IST FP6 for the partial funding of the AXMEDIS project.

European Projects for Art Schools: The Art-Net project

Elisabetta Delle Donne
Pixel
E-mail: eli@pixel-online.net

Abstract

The Art-Net project is a successful example of the effective use of ICT in the field of art teaching and learning at secondary school level. The project promotes the development of e-learning skills acquired by art teachers.

1. Introduction

The Art-Net project idea was developed within the framework of a previous art project which was funded by the European Commission to train art teachers in the use of new technologies for creating works of art.

Some of these teachers started exchanging information about teaching sources in the field of art and new technologies. This led to the creation of an on-line database of art teaching sources in the specific area.

The Art-Net project responds to the need to expand this database of teaching sources for all artistic subjects.

2. The Objective

The Art-Net project was financed by the European Commission in the framework of the elearning programme with the aim to create a transnational database of electronic sources for art teaching, accessible online.

3. The main project activities

The Art-Net project is currently developing the following main activities:

- Teacher training on the development of multimedia courses for the teaching of art subjects. The courses are held by expert professors from the Brera University of Fine Arts of Milan.
- Creation of multimedia courses on art subjects chosen by the schools, developed with the support of the project tutors
- Collection and analysis of e-learning products for teaching art subjects using a common evaluation form.
- Transnational seminars for students (based on distance training methodology) which will be organised on a monthly basis, involving all the European schools involved in the project. The seminars will be available via Internet and will focus on a different art topic each month. As an integration to the seminars, virtual meetings between the European schools will be organised with the aim of promoting discussion on the seminar topics.
- Teachers' workshops organised on a monthly basis as an integration to the distance training seminars
- Participation in the Lov'Art competition. The objective of the competition is the production of an artistic product which must be developed in collaboration (at distance) between two students of different nationalities.

4. The main project results

The results realised in the framework of the **Art-Net** project are accessible on the ArtNet portal (http://www.elearning-art.net). The portal consists of the following two main sections:

- Database. This section collects together material for the teaching of art subjects which have been created and/or evaluated in the framework of the project. The Database allows the user to carry out a simple search using key words or to carry out a more advanced search using the 4 search mechanisms of the Database: Theory (focusing on artistic movements e.g.: Impressionism); Chronology (focusing on

artistic periods eg: Modern Art); Typology (focusing on art typologies eg: Sacred Art); Techniques (focusing on Artistic techniques e.g: Drawing).

- Art Teaching courses. This section contains e-learning based courses developed by the teachers involved in the project. Each teacher identified a specific subject and developed his/her multimedia course. It is interesting to note that the technical solutions adopted are quite different and reflect the existing technical knowledge of each teacher involved, however the contents are presented quite clearly despite the different technical supports used.
- Tools. This section collects the IT tools (software) and the necessary instructions for carrying out the project activities (e.g. how to develop e-learning courses; how to download the necessary tools such as shareware programmes etc.)

5. Benefits for the schools involved

The schools involved in the **ArtNet** project benefit in many ways. The schools can:

- Access a database of electronic material for the teaching of art subjects, reviewed by teachers in Europe.
- Access the online training courses on art subjects created by art teachers throughout Europe.
- Acquire technical and methodological skills for the creation of multimedia courses and artistic subjects.
- Publish their own electronic teaching material in a virtual library for European art teachers.

Section:
Content Modeling and Gathering

Identifying and Defining Customer Orientated Best Practice Business Models within a Cross Media Environment – The INCCOM Project.

Juha Vallanen
juha.villanen@finpro.fi
Kritina Modee
Kristina.Modee@finpro.fi
Finpro Munich, Finland Trade Center Schleissheimerstr. 2, D-80333 München, Germany
Lisa K. Brown
L.K.Brown@surrey.ac.uk
School of Management, University of Surrey, Guildford, Surrey, GU2 7XH, UK

Abstract

The INCCOM project aims at gathering experience's from selected organizations and research projects within a cross-media technology environment, in order to propose new customer oriented business models in the anticipation of stimulating commercial exploitation of innovations as well as to focus future research activities. The growing interest in the forthcoming FIFA World Cup can be seen as an excellent opportunity to exploit the technological possibilities within digital cross-media content segment. As a direct result, the INCCOM consortium has researched mobile-tv trials, mobile-tv technology and end-user expectations and developed an integrated model which combines content, technology, business and user perspectives, whilst allowing the development of a hypothesis for expected revenue over time., This paper details this research which is in progress and in particular outlines the benchmarking process of the INCCOM project and the subsequent emergence and development of the INCCOM model to date. Preliminary research findings are also discussed.

1. Introduction

The INCCOM project examines the latest technical achievements as well as existing business models, and opportunities within a football framework in order to addresses user relevance and potential. The focus of INCCOM (as a Co-ordination Action) is the promotion of the take-up of cross-media content and services based on national and international delivery models of football related content. Scheduled over a period of 28 months and divided into three phases, the first phase was completed at the end of 2004 and early 2005. In the already active second phase, the focus is on the facilitation of co-operations within the digital content value chain which will enable further knowledge aggregation. To achieve this objective, the extensive network of the Consortium Members has been utilized as well as the network created through the implementation of workshops.

The INCCOM consortium subsequently embarked upon a detailed research and aggregation exercise to identify not only priority issues dominating the theme of the project (Network Convergence and Multimedia Distribution and Sports Experiences) but also to identify relevant research projects and completed mobile-tv trials, mobile-tv technology projects and initiatives. This has required the identification and implementation of a comprehensive benchmarking and best practice methodology. The results of which are briefly highlighted within this paper.

Late 2005 will see the exploitable knowledge of the project being transferred with the dissemination of information being executed in the form of a number of workshops held by Consortium Members, with participants representing organizations from the digital value chain. This will also serve as the first validation of the initial models and as a dissemination channel. The consortium, together with the workshop participants in the regions, will utilize the public interest in football partly related to the FIFA World Cup in their concerted effort to finalize the Integrated Cross-media Customer Oriented Models.

2. The INCCOM Benchmarking Process

Benchmarking within the context of the INCCOM Project is the search for best practices within cross-media business models that have lead to superior performance. Best practices, relating to business models and user experience were identified within the operative/working groups (Network Convergence and Multimedia Distribution and Sports Experiences), whom have analysed a number of projects and commercial activities in order to increase the INCCOM know-how and to develop initial commercial models (i.e. business models or service models) and to extract the best practices among them. The objective of this exercise was to understand and evaluate the current position of a business or organisation in relation to "best practice" and to identify areas and means of performance improvement within the cross-media environment.

Best Practice may be defined as 'a technique or methodology that, based upon experience and research, has proven to reliably lead to a desired result' (www.pemcocorp.com, 2005). Best practice applies to every aspect of business, from recruiting staff to how the business is marketed or uses new technologies. It involves keeping up-to-date with developments sectors and measuring performance against market leaders. Best practice is based on the principle that the best way to learn is through the experience of others. One way of doing this is through benchmarking. For the purpose of this paper, benchmarking I defined as 'outstanding results in another situation and that could be adapted to improve effectiveness, efficiency, ecology, and/or innovativeness in another situation' (www.ichnet.org, 2005).

Benchmarking allows the comparison of a business with other successful businesses, to highlight areas of improvement whilst illustrating how to implement such improvements by sharing best practice methods. It was deemed necessary by the INCCOM consortium that best practice should; already exist; should have clearly identifiable aims and objectives; be user-friendly and accessible to other relevant projects; be adaptable and transferable and be capable of being continuously improved. Many projects aim to identify what is best practice. However, it may be said that few projects actually achieve this and do not effectively analyse potential improvements that may be introduced into the organisation management to facilitate the identification and implementation of a best practice.

Benchmarking examining how others achieve their performance levels and to understand the processes they use. In this way benchmarking helps explain the processes behind excellent performance. When the lessons learnt from a benchmarking exercise are applied appropriately, they facilitate improved performance in critical functions within an organisation or in key areas of the business environment. The bottom-line benefit of benchmarking may be said to be improved competitiveness and an overall increased value to customers. However, benchmarking should not be considered a one-off exercise. To be effective, it must become an ongoing, integral part of an ongoing improvement process with the goal of keeping abreast of ever-improving best practice.

The investigation undertaken by INCCOM into best practices required the consortium to embark on various activities and processes which included; the identification of what is actually to be benchmarked; to identify comparable projects/research initiatives; to determine the data collection methods for each working group and collect the data; to also determine the current performance gap between existing cross-media business models; to project future performance levels and emerging cross-media business models; to communicate the benchmarking finds to a wider audience (European and National) in order to gain support and acceptance; for the INCCOM consortium to establish future functional goals for the project and develop action plans for the project and lastly; to implement any specific actions and monitor progress and recalibrate the benchmarks throughout each phase of the project.

On-going project benchmarking has been performed as a, systematic process for measuring and comparing the business models. It has also provided an external standard for measuring the quality and cost of the business models and the identification of where opportunities for improvement may reside. It has also fuelled the development of the INNCOM model which turns a proposed value constellation into a business potential. Figure 1 below;

Service	**Potential service proposition(s)**	**Usage and Revenue**
Mobile-tv service requires an infinite number of roles fulfilled by participating companies who make choices with the technology or content they endorse and how they relate to each other.	The service proposition is defined a sum of the contribution of the participating companies. INCCOM derives a specific service proposition as proportion of the ideal	The analysis of the service forms a basis for identifying target segments, propensity to pay, revenue potential over time & business case for all participating companies

Figure 1. The INCCOM Model

3. Benchmarking within the Working/Operative Groups - Success Indicators for Benchmarking

To each of the priority selection groups, the process of benchmarking and best practice is a discovery process and a learning experience that requires observing what the best practices are and projecting the performance for the future. Information has been gathered which will permit the setting of performance goals that are realistic in the context of cross-media business models and ensure that the more applicable and proven best practices are incorporated into the INCCOM project within the next phase.

The below provides guidelines for what the INCCOM project foresee as successful indicators of Benchmarking to be [1] An active commitment to benchmarking by consortium members and affiliated research projects [2] A clear and comprehensive understanding of how the identified research projects for best practice works, and how they are conducted as a basis for comparison with other cross-media business model best practices [3] A willingness to change and adapt based on benchmarking findings [4] A willingness to share information with other benchmarking institutions and research projects [5] A realization that situations (and competition) is constantly changing [6] To focus first on best practice in cross-media and second on performance metrics [7] Concentrate on projects with a cross-media focus and with functionally best operations that are recognized amongst its strata.[8] Maintain openness to new creative ideas and the innovation of their application to existing processes [9] Make benchmarking a continuous effort throughout the duration of the INCCOM project.

4. Benchmarking Methodology

A critical element of a successful benchmarking program is following a thorough process that requires a profound understanding of the process being studied and of the benchmarking process itself. This practice often involves finding and collecting internal knowledge and best practices, sharing and understanding those practices so they can be used, and adapting and applying those best practices in new and existing situations to enhance performance levels (APQC, 2005). This is illustrated further in the INCCOM project think tank approach to the establishment of best practice and whose development has been assisted through the projects interaction at plenary meetings and workshops. The stages of the think tank process for best practice can be defined (adapted from APQC, 2005) guidelines and methodology) as;

Strategy - During this phase the specific study focus area, key measures, and definitions are established and clearly documented.

Identify - the data collection tools are refined and finalized, and research is conducted to identify the best-practice research projects to study

Accumulate - This phase has two distinct objectives: 1) collect qualitative data and 2) learn from the best. The project analysis template is administered.

Evaluate - Key activities during this phase include analyzing trends and identifying practices that enable and hinder superior performance. The consortium presents a report containing key findings and insights knowledge transfer workshops. At workshops, participants discuss the key findings in-depth and have an opportunity to interact with each other and the best-practice organizations through systematic networking activities and presentations. The INCCOM consortium facilitates participants' initial action plan development to adapt and implement what they have learned.

Revise – Revision and improvement resulting from the best practices identified throughout project that occur after the INCCOM consortium and related research project take the findings back to their organizations.

It is important to remember that there is no single best practice for a cross-media business model and one model may not always be best for everyone. The best practice must demonstrate through evidence that it is better, faster, and cheaper. The Working groups have taken a complementary, but different approach on how to identify good practices, whereas as one focuses on multiple media, the other focuses on the service concepts themselves.

5. Preliminary Findings

Here, best practices and cases focus on the services, content and applications that are seen as good practices for trial business development. For INCCOM, the case comparison and best practices' main function is to identify gaps in the commercial environment, that can be utilized by the "early innovation". The aim of the Network Convergence priority topic is to provide a comprehensive overview of the status of ongoing accelerated network convergence between broadcast and telecommunication media in Europe. It places focus and seeks to support mid and long term market forecasts for the sector and strategies for newly emerging co-operations within it. Another focus is to research and establish differing views of manufacturers, telecommunications, content

providers/broadcasters and end users so as to allow for a more targeted development of services that are expected to be used and paid for because they are seen as adding value to the customer. Hence, the best practices identified within this priority topic provide an indication on how pricing, content and applications are successfully implemented.

An analysis of the end user and what attracts them (based in part on the findings of the M-CAST project and on a relevance ranking), users stated they were most likely to be attracted by News, followed by local content, business & finance, music & videos, TV content, jokes, movie trailers, shopping, user content and gossip. The least attractive was erotica. Based on the INCCOM project analysis and the acknowledged initial commercial models, the group identified good practices as to how to combine old content with new media and vice versa. Since mobile TV is the main focus of this working group, commercial launches and pre-commercial trials were selected as investigative best practice cases, that were analysed in order to observe how "cross-media" is today, and what opportunities it might lead us to and how large R&D projects or pre-commercial trials might want to approach the services and content;

Case 1–Offers "hybrid" approach allows for a valuable combination of TV channels with additional services: **Case 2**-Is an interconnected service that could be enhanced with more value added offerings in order to create an even broader choice for the mobile user: **Case 3**–Offers simultaneous live streaming of existing TV-channels has been proved relatively simple to achieve, e.g. with the offer of Vodafone Live! in Germany. Challenge here, is that mobile TV consumption will be very much focused on short formats/items. **Cases BMCO, Virgin Mobile and PTK Centerel** – Add on services to Mobile/TV and a good combination of technologies and brands to enable successful launch and publicity. PTK's case with 2.2 million new subscribers to the portal, combining old media with new content and vice versa.

The Multimedia Distribution and Sports Experiences working group focuses upon the social and broadcast media of the football experience and its relative importance in shaping the football experience for a particular fan type; Casual, Active or Lifestyle based. The working group has developed a generic, user centric approach on what services need to provide to the user. The football/sports environment was then analysed utilising the attributes of an "ideal service". The emphasis here is to investigate how the football experience and fans involvement can be enhanced by the deployment of cross-media. This involvement is centred on the before, during and after the experience as well as the location of the fan when involved in the experience, classified as being home, on the move or in the bar. A merger of these factors leads to the propagation of cross-media that may be deployed within each context to enhance the experience and to facilitate and above all identify best practices within this area.

In terms of what the 'interviewed experts' external to the consortium and based on the findings of other research institutions, in that all most three quarters of those interviewed believed that sports programmes will be the media watch on a mobile platform. The working group's analysis further focused on complementing the technological approach of the Network Convergence (Mobile TV), this Working Group analyzed the sports environment as a potential application area for the pre-commercial innovation related to the mobile TV. The benchmarking was based on the service concept development. By using the initial criteria of who were the actors for the service and who where the services for, several large sport brands and clubs were evaluated to estimate the potential and gaps for the technology and content, and to identify piloting partnerships for the WP4-WP6. The following were identified as either innovative or potential: FIFA; Mobile Lounge; Sportal. De; Bundesliga; FIFA World cup 2006; FC Barcelona and ContiFanWorld respectively.

6. Conclusion

The Working groups concluded that within the football (sports) environment, the requirements for a successful service are extremely challenging. In conjunction with well known brands, service needs to address as large number of needs of a fairly homogeneous target group. If again, the conclusion of an "ultimate" service is compared with the results relating to what are users are prepared to pay for, we can then get an initial indication of how innovation and business might be combined and the focus for the next research phase.

7.0 References

[1] APQC (2001) 'Mobile Virtual Network Operators and the Future of Mobile Business: off o the Races or Back to the standing gate?' Proceedings from the Silicon Valley world Internet Centre Think Tank Session, 26th September, 2001.

[2] APQC Process Classification Framework (PCF) - Version 3, June, 2005. Retrieved from; http://www.apqc.org/portal/apqc/ksn/APQC_PCF_June_2005.pdf?paf_gear_id=contentgearhome&paf_dm=full&pageselect=contentitem&docid=12138.

Focuseek searchbox for digital content gathering

Nicola Baldini
focuseek (Italy)
nicola@focuseek.com

Michele Bini
focuseek (Italy)
michele@focuseek.com

Abstract

The "search engine" approach to content management can be used successfully in many applications that need to acquire and manage unstructured, heterogeneous and geographically distributed digital contents. focuseek searchbox is a complete content gathering solution widely used in large enterprise environments where scalability, performance and easiness of management are key points. searchbox is also currently used in AXMEDIS european project as main Content Gathering and Information Retrieval platform.

1. Introduction

Gathering data from original sources is one of the main problems in digital content integration and delivery. A very typical scenario is when you have to gather information from many, heterogeneous digital sources that are geographically distributed too. Owners of such digitals sources are focused on their original mission of content production and usually do not provide a standard way to access their archives by the means of other applications. This situation is due to many factors but it is easily understandable that information is the main value of a content provider and he/she desires strict control on how it is delivered. As results of this status in many cases content providers do not really care whether the user wants to use other applications to access their information through standard protocols and formats. This situation is not the ideal one from the point of view of the user who has many content providers to interface with because he/she is forced to setup and maintain a custom communication channel with each. Such channels are characterized by custom user interfaces and are often very hard to be integrated in other applications.
A possible solution to this kind of problems comes from a custom declination of the approach that is currently used by Search Engines for Web plus the Web Service technology.

2. The Search Engine perspective

Web Search Engines cannot influence in any way how web sites publish their information so that if an engine wants to build an index of the content provided by some site it must access the web site on his own. The method used by Search Engines to accomplish that task is called "crawling" or "spidering". A web crawler is a software agent that simulates a real user accessing a web site and read all the information contained in it. In order to succeed with this task a crawler must have a toolbox with any possible "adapter" able to match all access protocols and document formats available on the web. After not so many years after its birth, the WWW begun to support other protocols than the original HTTP and many other document formats other than HTML. Formats like PDF, DOC, Flash and protocols like NNTP, FTP and ODBC (some of which actually predate the HTTP over HTML web standard medium) forced Web crawlers to adapt themselves to the new situation. The basic assumption of a typical Web crawler is that any information source must be treated like a "black box" with no way to contact the webmaster to ask him/her to adapt content for a specific usage. From the Web source point of view a Web crawler is like any other normal user that visits the site.
This particular approach is very powerful because it has zero organizational and technical impact on the information sources and for this reason it has been successfully adopted in the enterprise environment too. In any large company or public administration the goal of aggregating content from different and heterogeneous sources (even if they are located and managed by the company itself), is really hard to be accomplished. Exporting data from an existing database means that either or both the organizations providing and using the content has to obtain the necessary authorizations, writing some software and thus allocate some human resources. All those reasons are serious potential point of failure for any content integration project. In this type of scenario a crawling technology can enormously simplify the integration task because the crawler acts exactly like any other authorized user whose accessing procedures are already defined and accepted by all departments of any company.

3. The bridging brick

An interesting way to visualize the content gathering problem is to imagine that in order to acquire information we have to setup a channel connecting the content provider and the users. Using the already discussed "search engine" approach a possible solution is to create a

system able to aggregate many different information sources and provide some standard application services to access it. In this way users will only need to know the standard application interface provided by the gathering system.

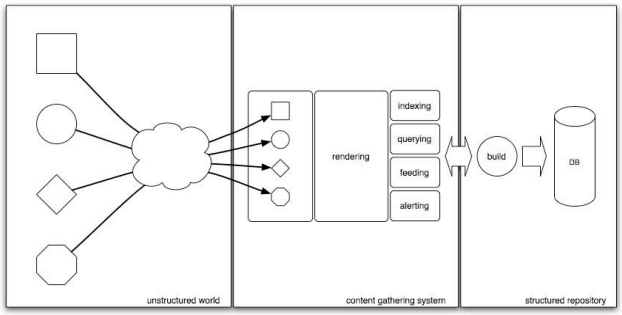


Figure 1. The Content Gathering component as bridge between content providers and client application.

At the left side of the above picture the heterogeneous world of content provider is sketched. Different shapes represent the different protocols and formats used to access to the content. At the opposite side there is a structured repository that needs to be filled from contents coming from content provider. The middle component is the content gathering module which choose the right adapter to gather information from any content provider and exposes some standard services:

The Indexing/Querying Service
Is able to retrieve any piece of information in the repository through a query composed by words or metadata separated by the AND, OR, NOT and NEAR operators typical of any search engine. The indexing is implemented using a full dynamic indexing service in order to take in account when a new content is added to the repository. No index rebuild is needed.

The Feeding Service
Used to automatically feed newly acquired contents through a channel. A very common standard like RSS can be used for this purpose.

The Alerting Service
Generates events to notify that something is changed in the repository. Alerting methods use email messages, Instant Messaging, SMS and Web Service calls.

The above services can be used by a client side component to build any kind of structured object based on the original “raw” information gathered from content providers. Obviously any type of structure provided by the content provider itself will be preserved and indexed too.

4. The focuseek platform

focuseek is a multimedia content gathering and indexing platform whose main goal is managing huge collections of data coming from different and geographically distributed information sources. The focuseek architecture has been conceived as an appropriate layer to construct information retrieval services for large enterprises, government institution, and Internet vertical portals. The focuseek platform, which collects information from different sources, implements a way to analyze the gathered content and provides a very flexible and high performance dynamic indexing for content retrieval.

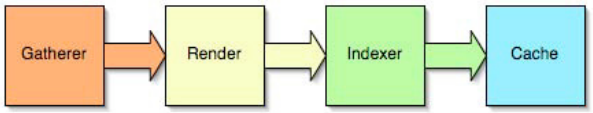


Figure 2. Main building blocks of the focuseek platform.

In Fig. 2, the main building blocks of focuseek are shown. The *gatherer* is the coordinator of a pool of gathering agents whose task is to acquire new data from an information source, as soon it is available. For instance, a noticeable example of a gathering agent is a focused Web crawler, which starts form a set of initial Web pages (seeds) and performs intelligent navigation on the basis of appropriate classifiers (see e.g. [1]). The gathering activities of the focuseek platform, however, are not limited to the Web, but operate with other sources, like remote databases, Web services, news servers, WebDav, IMAP folders, file systems and other proprietary sources. The gatherer module is fully programmable and customizable using appropriate plug-ins for the specific source.

The *renderer* is a central component in the focuseek architecture. focuseek indexing and retrieval system does not work on the original version of data, but on the "rendered version". Any piece of information (e.g. a document) is processed to produce a set of features using an appropriate algorithm. For instance, the features extracted from a portion of text might be a list of keywords, while the extraction of features from a bitmap image might be extremely sophisticated. Whenever possible, the extraction of text by appropriate OCR engines is very important for the information they provide. Even complex sources, like video, might be suitably processed so as to extract a textual-based labeling, which can be based on both the recognition of speech and sounds. All extracted features are then compiled in an internal XML format and passed to the indexing module. As shown in Fig. 3, the extraction

process of the renderer component is done by a pipeline of plug-ins, which provides the compilation of the final XML representation. focuseek currently makes some default plug-ins available for most common contents and an API to write customized plug-ins.

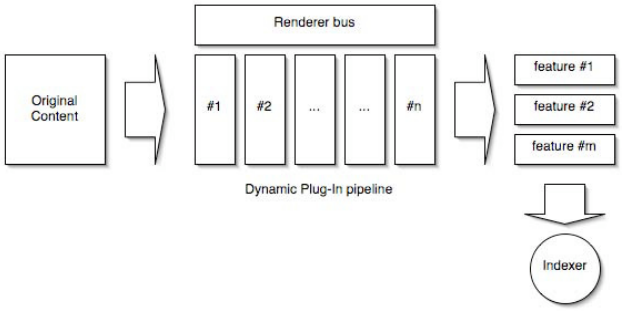


Figure 3. The structure of the renderer module: Different features can be dynamically added into the render.

The *indexer* creates the index of the collection of information gathered from multiple sources, while the *querying module* offers a complete query language for retrieving original contents. The index is fully dynamic in the sense that any indexed content is almost-immediately available for queries. This is a crucial feature when the system is used on highly dynamic sources. The focuseek indexer module can manage any feature that a specific renderer plug-in is able to extract from original raw content. All of the extracted and indexed feature can be combined in the query language made available by the query interface of the indexer module. focuseek provides default plug-ins to extract text from most common types of documents, like HTML, XML, TXT, PDF, PS and DOC. Other formats can be supported using specific plugins. Finally, a *cache* is available. The cache is highly coupled with the index and creates a copy that needs to be locally mirrored. It is a multilevel cache and can be used to store and index multiple versions of the same content. The possibility to "historicize" different versions of the same document is a relevant practical feature, which turns out to be especially interesting for the implementation of the *watch* searchbox concept. The overall focuseek architecture is shown in Fig. 4. focuseek is a component based platform completely developed in C++ and available for Windows, Linux and Mac OS X operating systems. All focuseek features are accessible through a complete SOAP API that is compliant with the latest Microsoft .NET standard. Finally, searchbox comes with a complete suite of administrative tools, both graphical and command line. Further details can be found in [3].

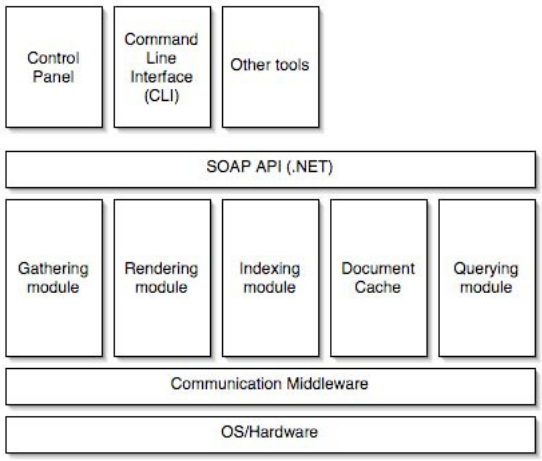


Figure 4. The overall focuseek architecture.

5. searchbox basic concepts

In contrast with other commercial search engine platforms, focuseek and the searchbox client/server application were also conceived as a end-user tools. In order to be configured by a non-expert user, searchbox exposes some basic concepts that can be combined together in such a way to configure the system in many different ways.

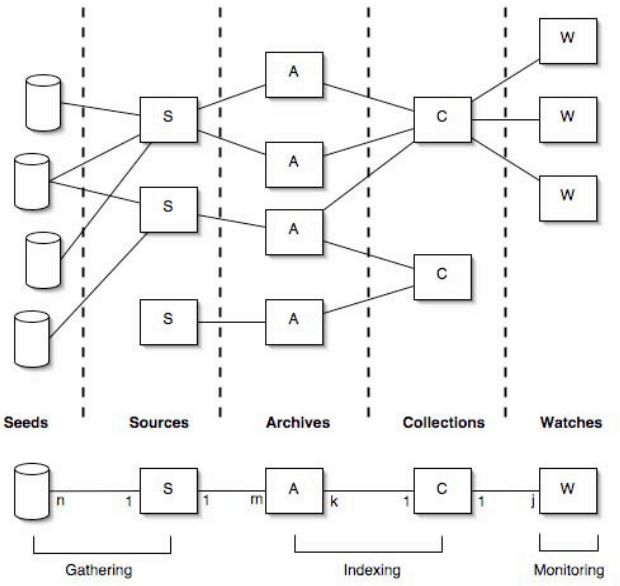


Figure 5 - Searchbox main concepts: Seeds, sources, archives, collections, and watches.

In Figure 5 these concepts are properly organized in three different groups:

Gathering group. The source is the central gathering concept. The goal of a source is grouping a certain number of seeds and configuring a suitable access protocol. A seed can be a database table or a Web page, a complete Web site, a specific portion of the Web or a fully custom data repository. The source natively supports access to seeds with digital certificate, password, cookies etc. A seed can be shared by many sources and multiple seeds can be used by a single source.

Indexing group. An archive represents an index and a repository of contents coming from a given source. Multiple archives can be connected to a single source, since every archive can have different configuration rules for its creation (i.e. caching or not, different access credentials for different users, etc.). Finally, in order to create indexes from different kinds of sources, many archives can be grouped together to form a collection. The collection represents a way to aggregate sources that are heterogeneous from the point of view of seeds, but that are homogeneous in terms of topic (i.e. all the Italian newspapers). Both archives and collections are query-able objects.

Monitoring group. searchbox can also be used as a monitoring tool. Watches contain a set of static filters on the information stream coming from a collection. A watch can be used to implement a customized view on any information stream originated from a group of dynamic sources through the corresponding collections. Watches support subscriptions from any client application that needs to be alerted as soon a specific condition is matched.

6. The searchbox client

In order to implement the *searchbox* application, all the exposed focuseek concepts have to be manipulated by the users. *searchbox* is a client/server application in which the client side software is a stand-alone thin-client application running on any Windows/Linux/OSX compatible operating system.
As shown in the screenshot depicted in Figure 6, the client looks like a standard three-pane application. At the left side, sources, archives, collections and watches are shown as a tree together with the corresponding configuration tabs in the central pane. Selecting the query tab for an archive or a collection, the user can submit a query to the system in the classical search engine way and obtain the results list through the simple built-in internal portal. A simplified modality where only watches are shown is also available. When selecting a watch the result of its persistent query is automatically shown and refreshed as soon as new results are available.

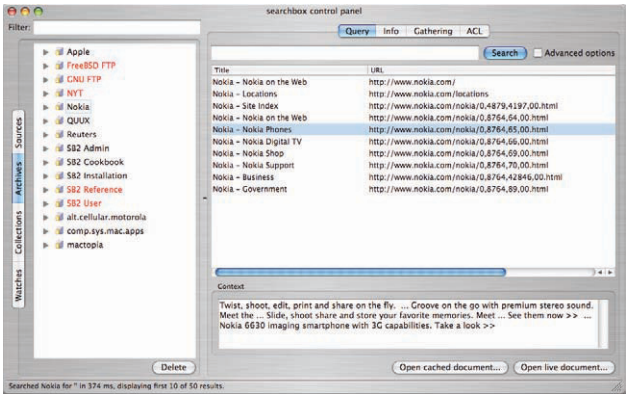


Figure 6. User's interaction at client level.

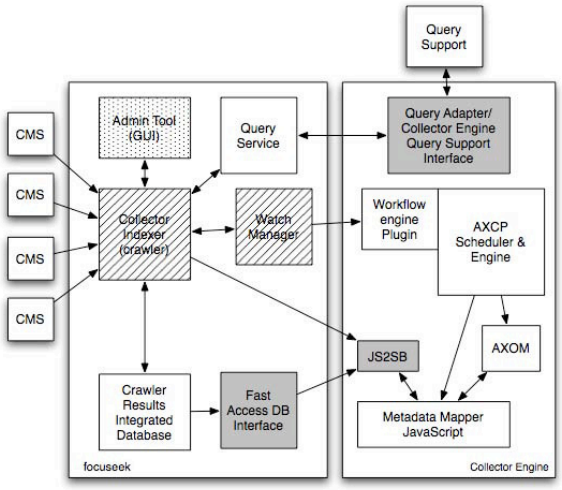
7. searchbox for AXMEDIS

Project AXMEDIS [2] aims to meet the challenges of digital content market demand by:

- reducing costs for content production and management by applying Artificial Intelligence techniques for content composition, representation (format) and workflow;
- reducing distribution and aggregation costs in order to increase accessibility with a Peer-to-Peer (P2P) platform at Business-to-Business (B2B) level, which can integrate content management systems and workflows;
- providing new methods and tools for innovative and flexible Digital Rights Management (DRM), including the exploitation of MPEG-21 and overcoming its limitations, and supporting different business and transactions models.

The AXMEDIS consortium (consisting of leading European digital content producers, integrators, aggregators, and distributors; and also information technology companies and research groups) is to create the AXMEDIS framework to provide innovative methods and tools to speed up and optimize content production and distribution, up to the *production-on-demand* capability, for leisure, entertainment and digital content valorizations and exploitation in general. AXMEDIS format can include any other digital formats and will exploit and improve other formats such as MPEG-4, MPEG-7, MPEG-21, as well as other *de facto* standards.
Focuseek searchbox provides to the AXMEDIS system the necessary capability to gather information from many different content providers. In the following picture a block diagram of the "focuseek for AXMEDIS"

subsystem is shown. A brief description of each component follows.



The above picture shows the subsystems focuseek and Collector Engine and the interoperation between each component:

Admin Tool. It is the standard "Control Panel" application distributed with all versions of searchbox application. This component has been customized for AXMEDIS purposes in order to manage the special way of using Watches into the AXMEDIS system.

Collector Indexer. It is the standard focuseek gathering (crawling) module that can be customized by using special fetching plug-ins in order to gather contents from non standard CMS provided by project partners.

Watch Manager. Watch management subsystem. It has been modified because it must notify the presence of new gathered contents to a Web Service provided by AXMEDIS. Thanks to a special SOAP method, the Workflow Engine plugin will receive the list of document IDs returned by the Watch together with the watch configuration (the query)

Crawler Results Integrated Database. It is the standard searchbox internal storage subsystem used by AXMEDIS as a document cache when needed.

Fast Access DB Interface. This custom module is used to transfer huge documents from the searchbox document store bypassing the standard SOAP interface that is not efficient for this kind of tasks.

Query Support: AXMEDIS component responsible of the query management.

Query Service: Searchbox service devoted to answering queries in a focuseek proprietary format.

JS2SB: C/Javascript bridge library that can be used to gain access to the searchbox SOAP API and to the Fast Access DB Interface module.

AXCP Scheduler & Engine: AXMEDIS component used to build Mpeg21 objects to be archived into the AXMEDIS database.

Workflow Engine plugin: This AXMEDIS component is a Web Service that can be called by Watch Manager through a specific SOAP method.

Query Adapter/Collector Engine Query Support Interface: Query converter from focuseek to AXMEDIS format and vice-versa.
Metadata Mapper Javascript. JavaScript rules collection for AXMEDIS object building. Such scripts can gain access to searchbox functions using the JS2SB library.

CMS. Content Management Systems integrated into AXMEDIS using crawler.

8. Conclusions

In this paper, we have briefly described the content gathering approach to content management and how it can be successfully used in all situations where there are many heterogeneous information sources out of our administration control. Also, we described the basic components and features of focuseek searchbox, a high-end commercial Content Gathering platform currently used by the AXMEDIS projects as "bridge" between content providers and the AXMEDIS system itself. Finally we gave some details on how AXMEDIS and searchbox integrate together.

Acknowledgments

We thank the AXMEDIS consortium for making the integration of focuseek searchbox into the AXMEDIS system possible. We are also grateful to a number of people involved in the project and especially to Prof. Paolo Nesi and Dr. Pierfrancesco Bellini for their valuable support.

References

[1] AXMEDIS Project web site: www.axmedis.org
[2] M. Diligenti, F. M. Coetzee, S. Lawrence, C. L. Giles and M. Gori, "Focused Crawling Using Context Graphs", Proceedings of 26th International Conference on Very Large Databases, VLDB, pp. 527-534, 10/9 - 12/9 2000
[3] focuseek: Introducing the searchbox searching platform (www.focuseek.com)

Jackson: DJ Software Powered by Musical Metadata

Francis Van Aeken
VAN AEKEN SOFTWARE bvba
francis@vanaeken.com

Abstract

We have built an application that facilitates the creative mixing of recorded songs. The power of our application comes from musical metadata embedded in the recorded songs. We use a proprietary format for the metadata but plan to make it conform to the MPEG-7 standard. We believe that musical metadata will play an increasingly important role in the future. It will power not only creativity tools, but also virtual DJs, personal radio and intelligent music compilations.

1. Introduction

In the hands of a disc jockey, a recorded song is not an end product but a starting point. By building up a meaningful sequence of songs and skillfully interweaving them, a new work of art is created.

Many professional DJs use vinyl because turntables allow for a meticulous control of tempo and timing. Only recently CD decks are gaining popularity because they now mimic the tactile control of vinyl decks.

DJ software has existed for some time, but those packages typically emulate the traditional DJ setup. To properly use those programs one needs a physical controller and the same skills as a traditional DJ.

Computers offer new possibilities. Starting from scratch, our company has taken a approach different from emulation. Our product is the first dedicated DJ software that implements sequencer-style mixing and that exploits the power of musical metadata.

The application automates the less creative but sometimes difficult aspects of mixing. For example, beat-matching, the process of rhythmically aligning two different songs is automatic and perfect. This opens up the art of DJing to a larger public without affecting its creative side.

The underlying technology also allows for the creation of virtual DJs. Products like portable mp3 players can be made more attractive by incorporating these software agents that mix songs according to the preferences and the mood of the listener.

2. The DJ software 'Jackson'

Over the last three years Van Aeken Software has been building the DJ tool 'Jackson' that allows for the easy manipulation and mixing of songs. The application, developed in C++, runs under Microsoft's Windows XP. A demo can be downloaded from [1].

2.1. Playing and manipulating songs

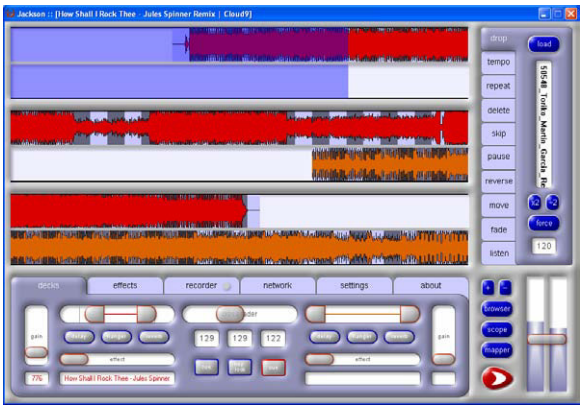


Figure 1. Jackson: main window

Figure 1 shows the main window of the application. Jackson supports two decks. The songs played on deck 1 and deck 2 are visualized as waveforms in different colors. The timeline is divided in three consecutive parts, one on top of another, much as in a music partition. The part of the songs already played is covered by a transparent colored rectangle. This rectangle grows with time until the complete first part is covered. At that point the two lower parts shift up and the 3rd part is replaced with new data. The rectangle disappears and starts growing from the left again.

The users can zoom in and out to see either the details of the waveform or a global overview of the sequenced songs.

Songs are automatically beat-matched and are therefore always rhythmically aligned.

On the right side of the waveforms window there are a number of tabs that cover different tools to manipulate songs and to control the mix. Jackson allows DJs to alter the structure of songs while playing them: parts can be repeated, skipped, paused and reversed.

Below the waveforms window one finds tabs covering the real-time mixer, the effects, the recorder, the network functionality and the user settings.

By using the filters in the mixer window the DJ can fade in and out, bass first or treble first. This way, parts of the spectrum of one song can be combined with parts of the spectrum of another.

The effects window features delay, flanger and reverb effects to spice up songs. All effects are automatically synchronized to the beat.

Through the recorder window the DJ can record the set he or she is playing to hard disk. This music file can then be burned on CD or published on the Web.

Through the network window, the user can configure the networking functionality of the application. Different computers running Jackson can be synchronized over TCP/IP. Several DJs can jam together while all played songs are automatically aligned to the beat.

The application also interfaces to standard midi controllers and custom controllers based on Measurement Computing's PMD-1208LS and Silicon Labs' C8051F320 controllers. Most DJs favor this type of controllers over the mouse. Not only are they better adapted to the type of control needed but they also allow the DJ to change multiple parameters at the same time.

Jackson supports the use of two different sound cards simultaneously. The output of one soundcard is directed towards the audience. The output of the other card is connected to headphones. Using such a configuration, one can cue like a traditional DJ or listen to parts of the mix in the future.

2.2. Selecting songs

A DJ is above all a selector. When mixing, nothing is more important than the selection of the songs. The browser, shown in figure 2, assists the user in this essential task.

The browser supports different virtual crates, corresponding to directories in the file system. Songs from a crate are displayed and ordered according to different criteria like title or tempo.

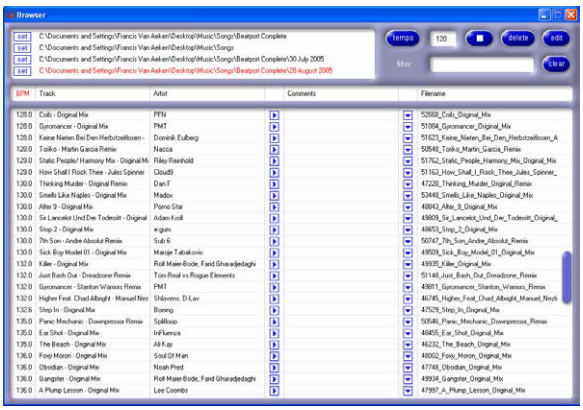


Figure 2. Jackson: browser

The browser features a previewer which is shown in figure 3. The previewer shows the rhythmic structure of any song and that lets the DJ listen to any part of it.

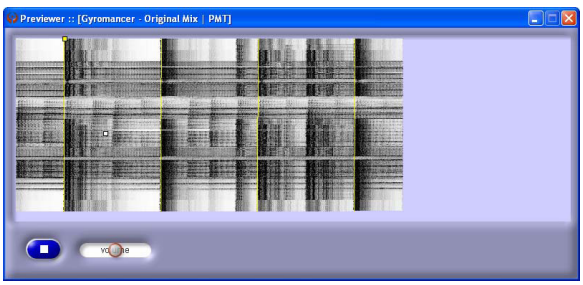


Figure 3. Jackson: previewer

The previewer shows a song as a grid of pixels of which the intensity corresponds to the energy of the signal in time. Black pixels indicate high energy while white pixels indicate the lack of energy, i.e. silence. Each row of the grid represents one measure of music. Time goes from left to right and from top to bottom.

Thanks to the visual representation in the previewer, the DJ can immediately see the rhythmic structure of the song and its evolution in time. An experienced user can readily identify the song in Figure 3 as a fixed-tempo break-beat song having a major breakdown after one 3rd of the song. Indeed, the beats vertically align and form straight vertical lines, meaning that each measure has exactly the same length. Also, the kick drums and snare drums (black elements) do not divide each measure in four equal parts. The pattern is more irregular, suggesting a break-beat rather than a four-to-the-floor song. Consecutive measures (lines) that have no black elements constitute breakdowns.

2.3. Analyzing songs

The key feature of Jackson is that songs are automatically beat-matched and therefore always play in sync. Musical metadata makes this possible.

Before a song can be played it must be rhythmically analyzed. Our application comes with a sophisticated beat-mapping tool that makes this an easy and instructive process. Figure 4 shows the beat-mapper's window.

The beat-mapper uses the same visual representation as the previewer. The system initially estimates the tempo. The user can then adjust the tempo or add markers on the onset of beats to take into account changes in tempo.

For electronic dance music the automatic tempo estimation is almost always 100% on the mark. In these cases, the user only has to put a marker on the first beat of a measure for the system to have complete information about the rhythmic structure of the song.

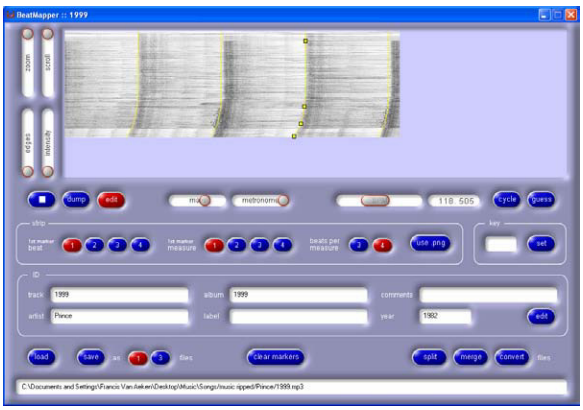


Figure 4. Jackson: beat-mapper

Different from electronic music most songs played live feature changes in tempo. Figure 5 shows the beat-map of the song Gigantic by the Pixies. Small squares indicate the position of markers. As one can readily see, steady tempo is not the trademark of this fine group. Indeed, in this case the beats no longer form black vertical lines. Beat-mapping such a song requires placing many more markers than the single one needed for electronic songs.

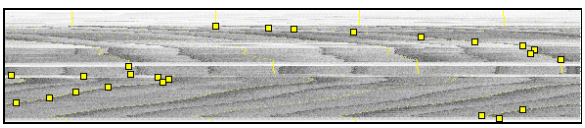


Figure 5. Pixies: Gigantic

Once a song is analyzed, the musical metadata is embedded in it and the DJ can mix and manipulate the song in all freedom.

This approach works with all styles of music, whether the tempo is fixed or not. Such styles often pose problems for traditional DJs. Thanks to a time-stretcher built into the application, songs in different styles and tempos can be combined into one rhythmically and harmonically consistent whole.

Apart from rhythmical metadata, one can also specify harmonic metadata describing the musical key of the song. Editorial metadata such as the name of the songs or its label can also be edited and embedded with the beat-mapper.

3. Musical metadata

Our metadata describes the musical structure of the song. We store the position of the individual beats and specify how they are grouped in measures and groups of measures. We also store a visual representation of the song, editorial and harmonic information, settings related to the beat-mapping process and (if applicable) the positions of mp3 frames.

When we started building the application, we were focusing on practical results rather than on standards. MPEG-7 still was fairly academic and few people were using it.

Given the market a couple of years ago, we focused on the WAV, MP3 and WMA formats. MP3 (or rather ID3v1 & ID3v2) and WMA have some support for metadata, but do not define fields for all metadata we use. Moreover, the two metadata formats are not interchangeable. Given all those factors, we decided to design our own format. We encode our metadata in a temporary file using a format structured much like the excellent PNG file format. This file is then embedded in the audio file. Optionally one can embed a second file that offers an alternative visual representation of the song.

4. MPEG-7

The more the metadata is accessible to different applications, the more its power can be leveraged. As standards mature, we must adhere to them as much a possible.

MPEG-7 is currently the most general standard to describe multimedia material [2][3]. Different from other MPEG standards, it focuses on metadata and not on the encoding of the data itself. MPEG-7 uses

Descriptors to describe low-level features and Description Schemes to describe higher-level features.

MPEG-7 specifies the Audio Waveform Descriptor that maps directly on the part of our metadata that visually describes the song.

On a higher level MPEG-7 specifies a Segment Description Scheme that represents the spatial, temporal or spatiotemporal structure of the audio-visual content. We can use a hierarchy of segments to describe the musical structure of a song. A song can be described as a sequence of sections composed of measures. Each of these is built up of beats. Unfortunately there seems to be no way to explicitly specify the musical meaning of this hierarchy (although the segments can be associated with low-level Descriptors and the relations between segments can be detailed). The mismatch seems to stem from the fact that the main goal of MPEG-7 is to facilitate searching and querying of material rather than their manipulation. This is an issue that we have to investigate further.

At the systems level, MPEG-7 metadata is written as XML, but can be embedded as compact BiM.

5. Other applications of musical metadata

There are relatively few DJs in the world. A DJ has to invest in equipment, has to acquire skills and has to continuously update his or her music library.

Technology like ours can substantially lower the threshold for people to start mixing: the investment in equipment is minimal if the person has a computer already. Also, the DJ no longer has to acquire purely technical skills like beat-matching. Still, not everybody might feel the burning need to become a DJ.

However, most people do love to listen to music and musical metadata can enhance today's listening experience. For example, it can be at the basis of virtual DJs that build musically meaningful sets according to the taste and mood of the listener.

5.1. The changing listening experience

To listen to music, people have been going to concerts and parties, switching on the radio or television or putting on a cassette, vinyl album or CD.

In all these cases user control over the experience was limited. In the case of cassettes, albums and CDs the format of the physical carrier determined the duration of the experience and the sequencing of the individual songs.

Electronic delivery of music over the Internet is doing away with the limitations of a physical carrier. People typically download individual songs rather than albums and sequence them the way they want to.

This new freedom implies new responsibilities: people have to be their own DJs now. Luckily, in the near future they can choose to delegate this responsibility to virtual DJs.

5.2. Virtual DJs

The technology we have developed can be applied to the development of virtual DJs that produce mixes automatically. We expect to find virtual DJs in desktop PCs, hi-fi equipment and portable audio players. Radio stations also will appreciate this technology.

We are currently building a first version of a virtual DJ on top of Jackson. Following a tempo trajectory and a virtual crate specified by the user, the DJ will create a custom mix.

The system will be driven by rhythmic and harmonic metadata so that the mix will be technically flawless. For an optimal listening experience, however, we will have to add metadata about the cultural and emotional aspects of the songs.

5.3. Personal radio

A virtual DJ does not need to be embedded in a hardware device. Nor does it need to run on a personal computer. It can just as well reside on a server that people can connect to for a personal radio experience.

In this business model people pay per time unit rather than per song. It is clear that the quality of the virtual DJ must be high for such a system to have appeal. Again, apart from metadata on the rhythmic and harmonic level, metadata describing the mood and the (sub)cultural identity of songs will drive these systems.

5.4. Intelligent compilations

We also see the potential of intelligent compilations of music (on CD-ROM for example) in which a collection of songs and one or more virtual DJs are combined.

Depending on the preferences of the listener, including the choice of DJ, a different mix will be produced. This new listening experience will be interactive and dynamic rather than passive and static.

6. Distribution of musical metadata

Currently no online distributor of music sells songs having musical metadata embedded in them. It is up to the end user to add the metadata to the files. This typically means musically analyzing the song or retrieving the metadata from another place (like a central server). Our application comes with an excellent tool to analyze music, but songs having an irregular tempo can take some time to analyze.

It is obviously in the interest of the user that this metadata is included at the source. Given the wealth of applications that can benefit from musical metadata, we expect many distributors in the future to embed metadata in the songs they sell. Standardization will be key to the success of musical metadata. It is therefore important that companies like us work closely together with competitors and standards organizations like the MPEG-7 Consortium.

7. Conclusion

Computer technology offers us great opportunities to enhance our consumption of music. The application 'Jackson' demonstrates that recorded songs do not have to be final products, but can be raw material to play with. Musical metadata is the key to this functionality. The MPEG-7 standard offers us a common language to write the metadata in. Many other applications can benefit from this metadata and we hope that music publishers and distributors will grasp this opportunity sooner rather than later.

8. References

[1] http://jacksondj.com

[2] François Pachet, “Knowledge Management and Musical Metadata”, *Encyclopedia of Knowledge Management*, Schwartz, D. Ed. Idea Group, 2005.

[3] José M. Martínez, *MPEG-7 Overview (version 9)*, ISO/IEC JTC1/SC29/WG11N5525, Pattaya, March 2003.

Multimedia Content Management and Streaming Services Platform

Alberto Canadé
Datamat Spa
alberto.canade@datamat.it

Abstract

The Multimedia Content and Streaming Services platform is a real life solution implemented at a tier one Mobile Service Provider in Italy.

Driven by market's needs it enables the provider to increase its ability to deliver new streaming multimedia services and applications, in near real time and time-shifted.

Key functionality is the automation of several steps in the multimedia content management chain, from real time extraction to enhanced digital processing capabilities (digital filtering, cropping, clipping, archiving) to encoding into selected and multiple streaming format till delivery to the streaming node.

Besides traditional scheduling, the Video Sequence Detection functionality further automate clip creation. Efficiency is achieved reducing time, costs, lowering errors for repetitive tasks. Multiple input feeds in parallel are allowed.

Advanced post processing functionalities also allow to easily test new handset's streaming capabilities as well as to support designing new multimedia format for its streaming services.

1. The scenario for a business case

As a basic principle of marketing would state, in mobile services market the end users' demand is steadily forcing Mobile Service Providers to reshape their offering in order to meet ever increasing expectations for multimedia and streaming services. Text-based information services pushed on customer's handsets through SMS are already evolved into streamed video news, video-communicating is taking place more and more frequently tying customers to this new form of getting in touch. That is accomplished also thanks to improved and more convenient business models and a broader coverage of 3G networks at least in high populated areas.

On the other hand, offering is based on the availability of comfortable handsets making easy and appealing consuming new services and applications.

The Mobile Service Provider has to be able to deliver innovative services before its competitors or at least at the same pace they do to stay competitive. All that facing shrinking budget, shorter time to market and eventually shortening staffs.

2. A positive integration

That was the scenario at a tier one Italian Mobile Service Provider when we were asked to propose a solution.

The Provider was already delivering innovative streaming and multimedia services to its customers and was facing all the difficulties outlined above. Reduced time from service designing to service roll out, new content formats needs to stay ahead of competition, new handset models testing requirements to make services and applications more compelling.

Upon analyzing processes and technologies in use at the Provider's it emerged a lack of integration in existing processes which was the outcome of recent and reckless projects aimed at quickly satisfying specific objectives: content format shaping, handset testing, multimedia content management and archive.

That opened a real opportunity to boost efficiency.

To tackle this situation Datamat proposed a multimedia content and streaming services platform enabling the envisaged integration.

The design of the proposed platform encompassed both the choice of functions to be implemented by the solution and the scouting of best of breed technology components to integrate together.

The first task was accomplished bearing in mind which processes was to be supported; that required a tight collaboration with the Provider' staff.

Hardware and software selection basically came out from requirements in terms of end to end processing time and streaming formats supported.

First step in the overall multimedia service chain is the acquisition of content from several sources: satellite, digital and analogic terrestrial feed, offline content (dvd, cd, betamax).

The requirement for the extraction phase was to enable the content publisher (the publisher is in charge of content creation through cropping, clipping and digital filtering) to identify in near real time the interesting events to be directly streamed or to be selected in order to be processed for service and content creation and archiving.

The extraction shouldn't disrupt original video flow integrity.

Following the capturing phase a two-fold path originates to enforce, as mentioned, a live streaming chain and a processing chain.

The first one is the simplest enabling the encoding of the extracted video segment into multiple streaming formats (Real, 3GPP, Windows Media Format). This chain actually implements the handset testing functionality in a straightforward manner.

Encoded video feeds a streaming service node ready for delivery.

For the live streaming chain the extraction can be triggered via a scheduled timetable in alternative to the on demand mode.

The second chain is more sophisticate since it includes additional functionalities and steps.

Upon extraction an MPEG-2 video segment is available for processing before service content creation. The processing phase is pretty articulated and it's where the selected components thoroughly come in.

Automated steps preceding clip creation are accomplished (a clip is a short video segment to be used in streaming services, content previews, browsable video lists and so on).

As video content encoding is completed the media analysis begins to get the semantic of the encoded segment: the I & B frames (in MPEG standard), colors, motion, faces, luminescence, shots, scenes, stories, objects, audio, speech, closed and open caption. The analysis process results in meta-data creation stored into an internal database; such data represents the base of video indexing, attained through a keyframe generation process. The video index is populated by thumbnail-sized keyframes automatically linked to browsing interfaces available to the video content publisher.

Another requirement from the Operator stated that keyframes were to be generated in size and formats suitable for different network and playback devices.

This objective has been met enabling the delivery of contents on GPRS, UMTS and broadband WEB (for which an higher resolution was preserved).

At this stage the video content, captured and analyzed, is represented in a non–linear way. The publisher can search and browse the video content randomly to identify and retrieve significant video segments and thus proceed for clip creation.

The web-based GUI has been designed to ease the search, browse and playback processes, including meta-data, video description, timestamps and actual video assets (the content itself).

The video review process is thus shorted by over 50% providing the publisher with a greater efficiency when editing, for instance, streaming services in near real time for particular event (sport events, music concerts, religion events and so on)

Clip creation has not only been facilitated but also enriched by advanced cropping and archiving capabilities. Zoom-in and zoom-out functions are easily integrated and accessible to the publisher through the web interface boosting the quality of the edited video content.

Before final publishing stage, clips are tagged with additional meta-data for enabling further search capabilities and are stored in the platform video archive. Multiple clips may be edited and merged to create new clips.

Automated digital filters may be activated to improve video quality.

Finally, multiple encoding processes take place in order to generate encoded versions of clips: these are in turn transferred to the streaming node from where they can be streamed in the supported formats (Real, WMF, Quick Time, 3GPP).

3. Platform architecture

The platform in its simplest implementation is represented by two parallel chains enabling live streaming and content and service management capabilities, as described above.

This two chains share the streaming node implemented by the RealNetworks Helix Universal Server – Mobile.

Additionally the live streaming chain includes a live capture and encoding node which is implemented by a Windows machine equipped with a Hauppauge WinTV Go capture card. Encoding is performed through RealNetworks Helix Mobile Producer.

The processing chain includes a capture machine and a processing machine (performing processing, publishing and encoding tasks) before the streaming

node. The capture and processing nodes are implemented through Almedia Gateway and Almedia Publishing software by Anam Software. The Gateway machine is equipped with an Osprey capture card.

The platform contemplates a Gigabit Ethernet connectivity.

The modularity of the platform easily allow to scale in order to support multiple parallel capture nodes and encoding machines. Server clustering, raid configuration of storage arrays and redundant network connections ensure high availability and load balancing requirements to be met.

3. Meet the objectives

Main objectives of the projects have successfully been met enabling several processes to be performed through an integrated content management and streaming services platform, with very minimal requirements.

Push and pull (on demand) services are both enabled: live streaming, messaging, video searching, video browsing, video playback.

Costs reduction and shorter content editing and creation time ensured a quick ROI for the Provider.

The centralized web interface eased the task of the publisher while the process automation in the video management workflow ensured a lower probability of errors.

4. Beyond the project

Further analysis of the implemented platform are in progress to extend its use also inside the Provider's intranet. The modularity of the design easily allows to preserve ensure appropriately higher resolution and video quality for users connected through a LAN, or DSL or broadband Wireless connections.

Content management capabilities are well suited for setting up video assets archive where content can be easily be searched and browsed.

An MMS composer has since been introduced to enable publisher to easily create short media presentation compliant for playback over the web, emails and GPRS/UMTS handsets.

Finally the creation of services and streaming contents can further be automated by the Video Sequence Detection which was not used in the project and which represents an alternative to the traditional scheduling interface for video content extraction process. Through this powerful functionality the publisher can simply specify the initial video sequence triggering the automated extraction and the overall duration for the clip to be generated (this is particularly suited to automate the extraction of clips from content like news or meteo, whose initial sequence is well defined). Before transmission to the streaming node clips can be reviewed through direct playback on the web interface. The non disruptive extraction process also allows the publisher to regenerate clips wherever required, accessing to the original video segment.

Additional encoding format can easily be supported by the platform just integrating appropriate codecs into the video content creation workflow, which requires very minimal configuration and integration work.

Section:
Distribution and Reporting

mReplay: Mobile Sports Replay and Fan Democracy

Patrick F. Riley
Fulbright Fellow
University of California, Berkeley
Berkeley, California, U.S.A.
patrick@sims.berkeley.edu

Erich Schubert
Ludwig-Maximilians-Universität München
Center for Digital Technology & Management
München, Germany
erich@vitavonni.de

ABSTRACT

mReplay (short for "mobile replay") is a solution to an intriguing problem: individuals attending sporting events have less information than those watching the same game on television, where commentators attempt to explain events and sometimes replay events repeatedly, showing the viewer a certain play. mReplay is an information system that provides on-demand instant sports replay functionality to most mobile devices, (for example non-3G mobile phones, PDAs, and even the new Sony PSP) including those mobile devices without video playback. mReplay also allows users to vote during the sporting event: for example, for their favorite play of the game, or on whether an officiating call was accurate, or for their favorite player, all from their mobile device.

Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

1. INTRODUCTION

Imagine yourself to the eighth inning of Game 6 of the super-charged 2004 American League Championship Series between the rival New York Yankees and Boston Red Sox. The Yankees lead the series 3-2, but the visiting Red Sox lead the game 4-2. While at bat, Alex Rodriguez hits the ball to the pitcher. The pitcher fields the ball and in his attempt to tag the running Alex Rodriguez, the ball comes out of the pitcher's glove. Derek Jeter scores from first base. Rodriguez is safe on second with the tying run. But the umpires huddle and discuss for four minutes, giving no indication of what has happened or what they are discussing. When they emerge, they reverse the call. Rodriguez is out and Jeter is sent back to first.

Figure 1. Shown above is the play that initiated 25,000 mobile phone calls from confused sports fans at the stadium calling individuals at home for information.

The important occurrence during this game was not whether the runner, Rodriquez, was out: it was that over 25,000 mobile phone calls were made by fans at this Yankees-Red Socks game during this four minute period [1]. Based on interviews and surveys following the game, we discovered that a significant majority of the calls were initiated by fans with the intention of finding out what happened during this play. Major League Baseball (MLB) rules prohibit replays of "potentially controversial" calls in the stadium, so no replay was given to the many confused fans. However, the television channel broadcasting the game on television replayed this particular play numerous times for the at-home viewer, from different angles, with expert commentary, and at slower speeds.

There are several information management problems that are exposed by this baseball game. First, fans in the stadium have less access to information than those fans watching the same event on the television. Secondly, fans in the stadium currently have few ways of "interacting" with other fans except through SMS and phone calls. mReplay is a multimedia processing and analysis system designed to address both these information problems by providing instant sports replay in mobile phones.

2. PRIOR WORK

Until recently, much of the research on sports fan interfaces has been devoted to backend systems of computer vision, most notably Rees et al [2] and Tovinkere et al [3]. These improvements, however, have not led to significant progress in replay interfaces, especially those that could be made available to mobile sports fans. Similarly, despite many intensive searches, we have found no similar research on portable devices being used by sports fans, or research with a focus on the information discrepancy between fans at the stadium and fans watching sporting events on television.

3. DESIGN GOALS

Our primary design goal was to design a backend system and a portable application that accesses this system allowing a typical mobile phone user on watch replays, and vote using current mobile phone technology. There were several difficulties it achieving this goal, most notably limited usability functionality on current mobile phones, limited bandwidth in current 2G (Second generation) mobile phone technology, and few preexisting automated sports replay systems. As we see in the example of the Boston-Yankees game, there is a preexisting system for sports fans at the game: contacting via voice or SMS to those watching the game on television, or possibly accessing the Internet via a WAP (wireless application protocol) browser to search for textual information.

We were certain, however, that an alternative system could be developed to provide sports fans with a fast, interactive, and rich experience via the mobile device. After several usability tests, mReplay provides the smooth and effortless experience that sports fans crave. The design of our backend system and "capture algorithms" also evolved through the development process, eventually providing highly reliable and automated annotations for sports plays, offering the typical sports fan the opportunity to watch the sports plays immediately following the particular occurrence in the game. During the design process of both the front and backend, and following the usability testing of the user interface, we considered many choices:

- **"Fan Democracy":** Give fans the ability to control what they want to watch, and vote on aspects of the game.
- **Consistency and Reliability:** Users should be given consistent and reliable annotations of the replays in a recognizable format that allows for prompt decisions.
- **Highly Dynamic:** Replays should be offered seconds following a specific play, as the value of a replay drops significantly over time.
- **Rich experience:** Replay image quality should be high despite current bandwidth limitations. A highly accurate annotated playlist should be presented for effortless information retrieval and a play-by-play history of a sports game.
- **User Control and Freedom:** The application should be flexible and efficient to provide basic and advanced functionality for a diverse user base. Allow fans to watch replays from 10 seconds ago, or even months ago. Offer easy features to interact, such as voting for their favorite player, or if a play was controversial or not.
- **Hardware and Platform Independent:** Provide a highly rich environment without 3G (Third Generation) mobile connectivity or hardware. Build an application that can be used on almost every preexisting mobile phone or other wireless mobile device.

4. mREPLAY MOBILE APPLICATION

Our application starts with the user selecting the team and game they would like to access. Following their selection, they are offered a textual chronological list (Figure 2) of the replays available from that specific game (i.e. "Cal TD, Cal 49, VT 35"). These descriptions of the plays are automatically generated by the mReplay backend, which we will cover later in this paper. If the selected game is in progress, this list is refreshed and updated with new plays soon after they occur.

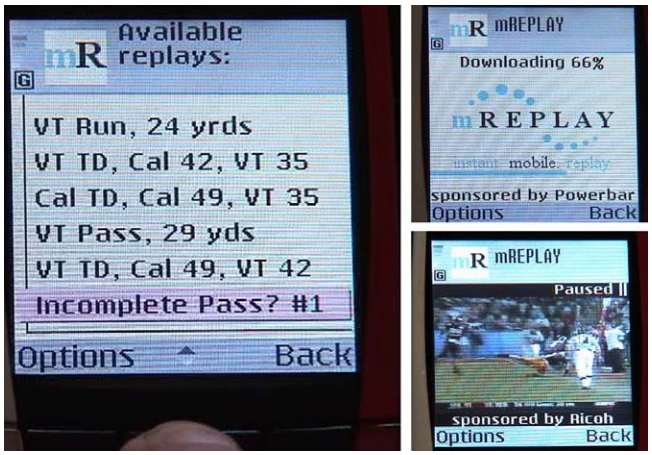


Figure 2. The mReplay "playlist", downloading and playing of a football replay on a Nokia 7610.

On this "playlist", the user is also offered "Last 15 sec" or "Last 30 sec," allowing the user to watch any part of game, not only those replays that were significant enough to have been automatically captured, annotated and listed in the mReplay system. Once a user selects a particular play, the application downloads the particular images to compose that replay. The replay then automatically played on the mobile phone, and the user has the ability to rewind, fast forward, or pause the replay frames. System status is displayed during a change of state (i.e. "Pause"). During our beta tests of varying sized replays, the average time for downloading a replay was 13.6 seconds. At any point during the process, a user can select "back" to any previous screen in the mReplay application, or access their other mobile applications such as phone or messaging functionality.

During the occurrence of a game, an mReplay user also has the ability to vote on particular polls regarding the game, such as favorite player, or the "play of the game", or even if they think an officials' call was accurate or not. The user will be offered these voting opportunities built into the same chronological playlist. Users are then offered to vote and the results of a particular poll are available at any time. During our beta test, we had an average of one voting opportunity every twenty minutes, and over 92% of the beta users used this voting functionality.

5. mREPLAY "AUTOPILOT" BACKEND SYSTEM

In order to provide the dynamic features that the mReplay offers users of mobile devices, it was essential that the

backend system be design to intelligently parse the barrage of incoming information. The backend system must not only capture the television signal and convert it into a rich format that is compatible with mobile phones: the backend system also has to know precisely what it is capturing. We designed what we call the mReplay "Autopilot" that takes a cable television signal, captures the imagery, analyses the images and other information. The "Autopilot" computer than provides all the users of the mReplay application the "playlist" of available sports replays. Figure 3 shows the many of valuable information "mediums" that our system can use to recognize what is occurring in a sporting event.

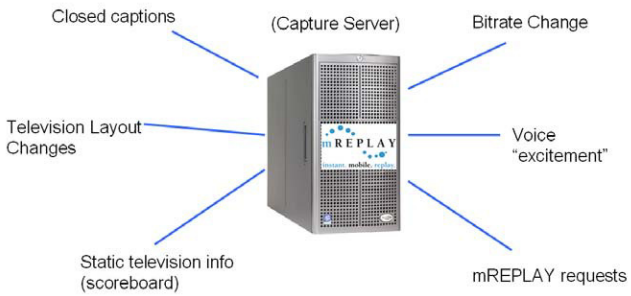


Figure 3. The analysis of the mReplay backend system.

5.1 Analysis of "Non-imagery" information

Many of the following individual procedures, for example, slow motion detection, are based on past work in computer vision and HCI. However, we believe our design and algorithms for consolidating all this information in order to provide immediate sports replays to mobile phones is quite original. This "non-imagery" information is essential for interpreting what is occurring in the game so the system can then provide an accurate and concise description of the replay for the user on the "playlist." Let us first look at how the backend system captures and examines the "non-imagery" information:

- **Closed captions:** We derive much textual information from capturing the closed captions to look for keywords and semantic relationships of words. Play-by-play commentary, game score information, player name information: all of extremely value for our system to provide a factual explanation of what is accruing in the game.
- **Audio recognition:** Voice recognition is still quite unreliable and in our case, unnecessary, since most television broadcasts of sporting events offer closed captions. However, particular information can be determined to be more creditability with the combination of audio values. For example, during a game, the word "home run" may be found by our closed caption scanner. Naturally, this does not necessarily mean a home run occurred: perhaps a sports commentary simply said, "There has not been a home run all game long." However, audio values of both the sports game commentary and the spectators are available though the television signal. If "home run" is found in the closed captions relatively simultaneous with an audio value consistent with "excitement", then our system has a highly accurate process to determine what is valuable in the game and when a "replay" should be offered to users.
- **User Requests LOI (Level of Interest):** Since our system always offers a replay of the "Last 15 seconds" or "Last 30 seconds", requests from users also can be used to establish the LOI (level of interest) [4]. If, for instance, our backend system does not automatically discover a valuable part of the game, but a significant percentage of the users request the "Last 15 seconds" or "Last 30 seconds", our system immediately determines this increased level of interest and offer this popular replay as part of the regular "playlist."

5.2 Analysis of "Imagery" Recognition

Our system also heavily relies on specialized forms of multimedia analysis to assess what is valuable to the user and what is occurring during a sporting event.

- **Layout Changes:** During our research, we found that during televised sporting event, the television network changes the screen layout during their replays. As shown in Figure 4, when the television broadcast replays an occurrence in the game, they remove the "scoreboard" and other onscreen information to give the viewer a clean view of the replay. Our backend system watches for this change and can assess it as a replay occurrence that may be desired by mReplay users.

Figure 4. The layout change during slow motion replays.

- **Static Information:** Typically during a sporting event, the network displays an onscreen "scoreboard" and other factual information about the game in progress. Although many changes have different formats of how this "scoreboard" is displayed, our system can recognize this information and contribute these values with the additional information assessed from the system.
- **Slow motion detection:** Based on the research of Kolba et al [5], our system uses the macroblock, motion vector and bit-rate information to accurately determine when a slow motion replay is occurring on television.

By consolidating all this information, our backend system is designed to systematically determine which sections of a particular sports game are mostly likely to be desired by sports fans. In addition, there are several features that

provide for customized use, such as watching the last 15 or 30 seconds of any portion of the game, and voting about particular aspects of the game, all in the same application. The backend updates the mReplay application in a mobile device via a WAP or 802.11 connections, offering the user a list of all the important replays of a particular sports game.

6. IMPLEMENTATION

Despite bandwidth limitations of mobile devices, variations in operating systems, variations in hardware, especially in terms of mobile phones, the overwhelming positive feedback we received from our beta tests was very encouraging. In order to optimize the speed in which the system captures and analyses the sports events, we used Gentoo Linux on our backend server. This allowed us to configure the kernel, access and utilize the hardware as needed, easily add applications specific to system needs and design iterations, and optimize compile time.

Our backend analysis computer was a dual Intel Xeon 2.8 GHz with a Hauppauge Win PVR 250 TV capture card and an Intel GigE capture card connected to our webserver. The mReplay application was installed on a significant variety of mobile phones, but we used several Nokia 3650, 7610 and Sony Ericsson 610 Series during our development and design iterations. All the components comprising the backend computer were composed in C, and the mReplay application was developed in J2ME [6].

7. LESSONS LEARNED

We learned from our beta test that fans were indeed eager to have the control to watch replays whenever they want with their current model mobile phones. We were satisfied with the design and quality of our replay system, and feel it offers a unique interaction for sports enthusiasts. The following are lessons that will continue to inspire us to create future versions of mReplay:

- **Speed is demanded:** Most replays are significantly "less valuable" over even a short period of time. This is what makes mReplay valuable. Fans can find and watch replays promptly following the particular occurrence in the game.
- **Automatic annotations are critical:** Despite having the "Last 15/30 seconds" replay option, most users still base their decision to watch a particular replay on the annotation of the replay that our backend system creates. Using only the information from the television signal, our system provided highly accurate descriptions of the important replays.
- **Interaction was popular:** Users enjoy being able to interact with the sports event using the voting functionality. We think there are significantly more features that could further enhance the sports fan experience in this manner.

7. SUMMARY

Instead of thousands of fans using their mobile phones to call individuals watching the same sporting event on television, mReplay provides sports fans a reliable and highly dynamic application for instant sports replay. Naturally, manual annotation or analysis of these sports games would be far too expensive and arduous task for either an individual or company to do. And current mobile TV services such as MobiTV [7] and Orb [8] provide only the television signal, without any way to store or offer a selection of the best plays. By analyzing the many attributes of a television signal, we can automatically detect and annotate sports games, and can give individuals at a sporting event the at-home television experience. One can have the "bragging rights" (boasting about attending a sports game) and enjoy the atmosphere of attending a game, while still having significant user control and freedom with Tivo-like features in the palm of their hands.

The mReplay system consolidates information to offer sports fans true "Fan Democracy": the ability to chose, watch, or pause sports games with their mobile phones at anytime, whether seconds or months following an event, and the ability to interact with other sports fans in a democratic fashion by voting on sports related issues such as the best play of the game or whether an umpire's call was correct.

8. HARDWARE REQUIREMENTS

The mReplay mobile application can be demonstrated on one of the hundreds of mobile phones that are J2ME-enabled. At the conference, the authors will project the screen of a Nokia 7610 during the presentation to demonstrate the mReplay system.

9. REFERENCES

1. Fry, Jason. Real Time, Dow Jones, (Oct. 25, 2004).
2. David Rees, Johnson I Agbinya, Nick Stone, Fu Chen, CLICK-IT: Interactive Television Highlighter for Sports Action Replay, Pattern Recognition, 1998.
3. V. Tovinkere , R. J. Qian, Detecting Semantic Events in Soccer Games: Towards A Complete Solution, Proc. ICME 2001, Tokyo, Japan, Aug 22-25, 2001
4. Nair, R. Calculation of an Aggregated Level of Interest Function for Recorded Events. In Proceedings of the 12th annual ACM international conference on Multimedia (MM 2004; New York, USA).
5. Kobla, V.; DeMenthon, D.; Doermann, D., Detection of slow-motion replay sequences for identifying sports videos, Multimedia Signal Processing, 1999 IEEE 3rd Workshop on , 1999, Page(s): 135 -140
6. Java 2 Platform, Micro Edition, Sun Microsystems.
7. MobiTV, http://www.mobitv.com.
8. Orb Networks, http://www.orb.com.

Digital Items: Event Reporting Integration with Company Business Models

F. Fioravanti, M. Spinu and M. Campanai
Exitech, Italy
ff@exitech.it, ms@exitech.it, mc@exitech.it

Abstract

Together with the organization business model evolution, companies adopt business information management systems; to this end, also reporting systems are used. This paper approaches the two systems integration by defining the rules, the types of data and the practical mechanisms to put them together.

1. Introduction

A large effort has been spent in the past and is currently employed in formalizing how to report events related to the manipulation of Digital Items (DI), in order to track user actions. User action tracking is the primary source of information for digital right owner in order to bill the user (both business and consumer) that performed the action: event reporting is obviously the basis for tracking user actions.

Several ISO documents [2] [3] and projects (among them also AXMEDIS [1]) are focused on Event Reporting (ER). Such Events are mainly related to the actions performed by the users and/or business partner but usually do not contains information related to "business related data" such as payment amount, payment type, etc.

It has to be considered that ER is only a part of the information that companies working on DI have to consider.

Other efforts were spend in research for the development of performing and secure financial systems (network security, payment systems, data protection, risk management, etc.).

This paper inspects the relationships among ER system such as those of AXMEDIS and the integration with business information providing a model that can be successfully employed on the company Customer Relationship Management (CRM) side. The model is based on the information collected during the specification of the AXMEDIS system [1] but is usable in all the sectors where DI manipulations are performed.

2. Event Reporting Systems

ER systems, in a general manner, allows to generate a report on actions performed by an user on an DI. These actions are usually authorized by using trusting systems or are managed by licensing mechanism.

The ER System is usually capable of collecting a huge number of information. Among the ER system, the AXMEDIS Certifier and Supervisor (AXCS), collects several different information that are summarized in the following table. These information are not limited by the adoption in a single project since have been collected by interviewing different company types (collecting societies, distributors, DI integrators, and so forth) adopting different business models.

logId	Unique ID of the transaction
objectID	ID of the DI
userID	ID of the user that has performed the transaction
distributorID	ID of the distributor to which the transaction is related
deviceType	Device type used for the transaction (kiosk, portable, mobile phone, etc)
executionTimestamp	Timestamp of the transaction
recordingTimestamp	Timestamp of the recording of the transaction
licenceID	ID of the license under which the DI has been accessed
location	Geographical area in which the DI is used (especially for collecting societies)
operation	Use of the DI (reading, printing, aggregating, editing).

Table 1: Data recorded by a typical ER system

These information are not enough for the billing or accounting process since several business information are missing.

On the other hand managing business models and business information is not the work of an ER System.

3. Business Information to be collected

Starting from the analysis performed in AXMEDIS user requirements and specification [1] it is possible to identify a set of information that are important from the business point of view and that must be considered in the integration among ER Systems, Content Management System (CMS), CRM, and banking systems.

These information can be usually recovered in the company CRM on the basis of some ID reported by the ER System.

A non comprehensive sample of such information is reported in the following table:

transactionValue	The overall value of the transaction
userDetails	Detail of the user (such as name, company, address, etc)
objectDetails	Details of the object (such as version, metadata, etc)
paymentForm	The code describing payment form selected (Cash, Credit card, Coupon, Pre-paid card, etc)
paymentID	ID of the economic transaction for example for credit card or pre-paid cards)

Table 2: Business data not recorded by a typical ER system that need to be merged with ER data

These information can be extracted directly from CRM (such as userDetails on the basis of the userID, or transactionValue of the basis of objectID and licenceID); obtained by external system such as banking or accounting system (such as paymentForm and paymentID); or obtained from the CMS where the object is located (such as objectDetails).

The only data that can bind these items together are the IDs that are contained in the ER System that become central in the business process. Each time an economic transaction is performed, the system can send to the billing or accounting system the logID in order to collect the billing information together with the ER related to an object or to a user.

3. Model for merging information

On the basis of the analysis performed in the previous sections, a general model for the cooperation of all the entities involved in the business process together with the ER System can be provided. This can lead to obtain in each company an entity that collects all these information together named XERAS (eXtended Event Reporting Accounting System).

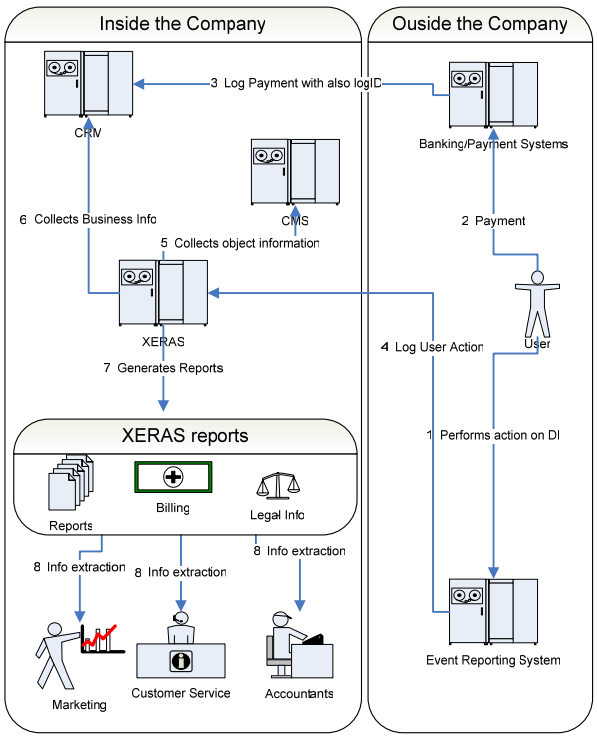


Figure 1: XERAS system that merges ER information with CMS and CRM data.

A simple flow of the action that can typically be performed during an action on digital items are summarized in the previous figure and briefly described in the follow.

User performs action on DI (1) and pay for that operation (2). Banking system and ER System send to CRM (3) and to XERAS (4) their respective information. XERAS, on the basis of the IDs collects object information from CMS (5) and business information from CRM (6). These information are organized in order to prepare reports (7) that can be used in the company at different level (8).

The reports generated by the XERAS system are not limited to the statistical reports that can be generated also by the ER system, but are enriched with all the business information that can help marketing for

strategically planning and accountants in solving contracts problems.

XERAS is therefore a system that collects and merges information from CMS, CRM and ER Systems in order to generate comprehensive reports that can be used by different people in different company contexts.

4. Conclusion

Event Reporting Systems for Digital Items offer the possibility to create a collecting point (the AXCS of AXMEDIS project [1], for example) available for different companies with different business model. This collecting point allows to develop inside the company a complete system for collecting together and merging information on action performed on Digital Items, information related to payments and information extracted from the company CMS. This system has been identified in this paper as the XERAS system that could be a good starting point for discussing with companies about the integration of standard source of information (such as CMS and CRM) with new emerging standards such as Event Reporting system.

5. Acknowledgments

The authors would like to thanks to all AXMEDIS project partners including the Expert User Group and all affiliated members, for their contributions, funding and collaborations. A specific acknowledgment to EC IST FP6 for partial funding of AXMEDIS project.

6. References

[1] www.axmedis.org
[2] www.chiariglione.org
[3] ISO/IEC CD 21000-15 (Information technology -- Multimedia framework (MPEG-21) -- Part 15: Event Reporting)

Keeping Control of Online Assets Using Windows Media DRM 10

Duncan Burbidge
Stream UK
duncan@streamuk.com

Abstract

Windows Media 10 Digital Rights Management protection allows a complete range of business rules to be applied to content.

Protection of the files can be done either live or on-demand.

Stream UK has developed a complete set of management tools for the administration of the rights and marries this to a global content delivery network.

1. Introduction

DRM works by requiring users to obtain a license before they are able to play or listen to a file. The rights to the license are set by the owner of the content.

Stream UK is one of only three Microsoft-recommended DRM providers within Europe (see the site at http://www.microsoft.com/windows/windowsmedia/drm/9series/providers.aspx#live). The complete solution is usually integrated with e-commerce and a front-end that allows license permissions to be set on the fly.

2. Licenses

Licenses are acquired when a user tries to play the file. This means that the media can be distributed through offline means such as CDs, as long as the user is online when he or she wishes to play it.

License permissions allow the commercial exploitation of valuable content. Licenses cannot be transferred between machines and can be delivered invisibly or subsequent to a form-based login.

Once the initial requirement to obtain a license has been built into the media file, the use can be restricted according to:

- Expiry date;
- Start date;
- Number of plays;
- Total amount of time viewed;
- CD-ROM burning;
- Download permissions; and
- Various other commercially exploitable rules.

3. Protection of the files

Protection of the files can be done either live or post event.

Live protection is done by downloading a profile to Windows Media Encoder. This profile integrates with the Stream UK system.

On-demand protection is done through the Stream UK content management system. More complicated business rules are better done through this system.

4. Hosting of the content

Load-balanced and redundant hosting of the content is essential for successful delivery.

This section will touch on the key elements of a global CDN, and the advantages of the Smart Content Delivery system.

4. Delivery to mobiles

The new Play4Sure® technology allows delivery of content to mobiles, complete with a full set of business rules.

5. Case study

Celtic Football Club uses the DRM system to ensure secure protection for all their clients. Revenue generated is sufficient to run all of the web activities at a profit.

Panel on the Role of Collecting Societies in the Digital Era

Organised by

AFI, Associazione Fonografici Italiani

Coordinated by

Isabella Longo, AFI, Italy

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

Workshop of MUSICNETWORK on MPEG Symbolic Music Representation

Organised by

International Association of Interactive Multimedia MUSICNETWORK
www.interactivemusicnetwork.org

Coordinated by

Piefrancesco Bellini, University of Florence, Italy

Paolo Nesi, University of Florence, Italy

Maurizio Campanai, EXITECH. Italy

Giogio Zoia, EPFL, Switzerland

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6 © 2005 Firenze University Press

MPEG Symbolic Music Representation, history and facts

Pierfrancesco Bellini*, Maurizio Campanai[%], Paolo Nesi*, Giorgio Zoia+
* DSI-DISIT, Department of Systems and Informatics, University of Florence
Via S. Marta, 3 - 50139 Florence, Italy, nesi@dsi,unifi.it, http://www.disit.dsi.unifi.it
+ Signal Processing Institute, École Polytechnique Fédérale de Lausanne
EPFL - 1015 Lausanne, Switzerland, giorgio.zoia@epfl.ch
% EXITECH srl,Florence, Italy, mc@exitech.it
http://www.interactivemusicnetwork.org/mpeg-ahg/index.html
http://www.interactivemusicnetwork.org

Abstract

A symbolic representation of music is a logical structure based on symbolic elements representing audiovisual events, the relationships among those events, and how they can be rendered and synchronized with other media types. Many notations have been developed over the years and ages to represent visually or by other means the information needed by a performer to play the musical piece and produce music as composed and imagined by the author. Symbolic music representation generalizes music notation concepts to model the visual aspects of a music score, and audio information or annotations related to the music piece. Symbolic Music Representation overcomes the limitations of MIDI, which is good enough to transport music event information (its main purpose), and it has limitations in producing satisfactory results on the audio and visual representation sides. The evolution of information technology has more recently produced changes in the practical use of music representation and notation, transforming them from a simple visual coding model for sheet music into a tool for modeling music in computer programs and electronic devices in general. As a consequence, symbolic music representation is currently used for several purposes other than sheet music production and music teaching, such as for audio rendering, entertainment, music analysis, database query, performance coding, etc.

1. Introduction

The MPEG-4 technology covers a huge media domain. In particular, the Audio part of this standard offers the possibility to include standard MIDI content synchronized with other forms of coding; it allows structured descriptions of audio content through a normative algorithmic description language associated with a score language more flexible than the MIDI protocol (MPEG-4 Structured Audio). These tools, though allowing to *derive* in someway a symbolic representation out of the information they carry, are to a large extent not enough to guarantee a correct coding of notation as they lack for instance any kind of information about visual and graphic aspects, many symbolic details, a thorough music notation modeling, and many necessary hooks for a correct human-machine interaction through the SMR decoder. MPEG-7 also provides some symbolic music related descriptors; but they are not meant to be a means for coding SMR as a form of content. On the other hand SMR content is a complete symbolic music representation and it may be rendered in synchronization with other audio-visual elements.

The MPEG SMR work item is trying to open the way for all the new applications summarized in Fig. 1 below. Many music-related software and hardware products are currently available in the market and they may greatly benefit from it, since it will foster new tool development by allowing highly increased functionality at reduced cost. Examples of applications that may rapidly be affected by this increased functionality include:

- Interactive music tutorials
- Multimedia music publication
- Software for entertainment (sound, text and symbolic information)
- Play training, performance training, ear training
- Compositional and theory training
- Software for music management in libraries
- Piano keyboards with symbolic music representation and audiovisual capabilities
- Mobile devices with music display capabilities

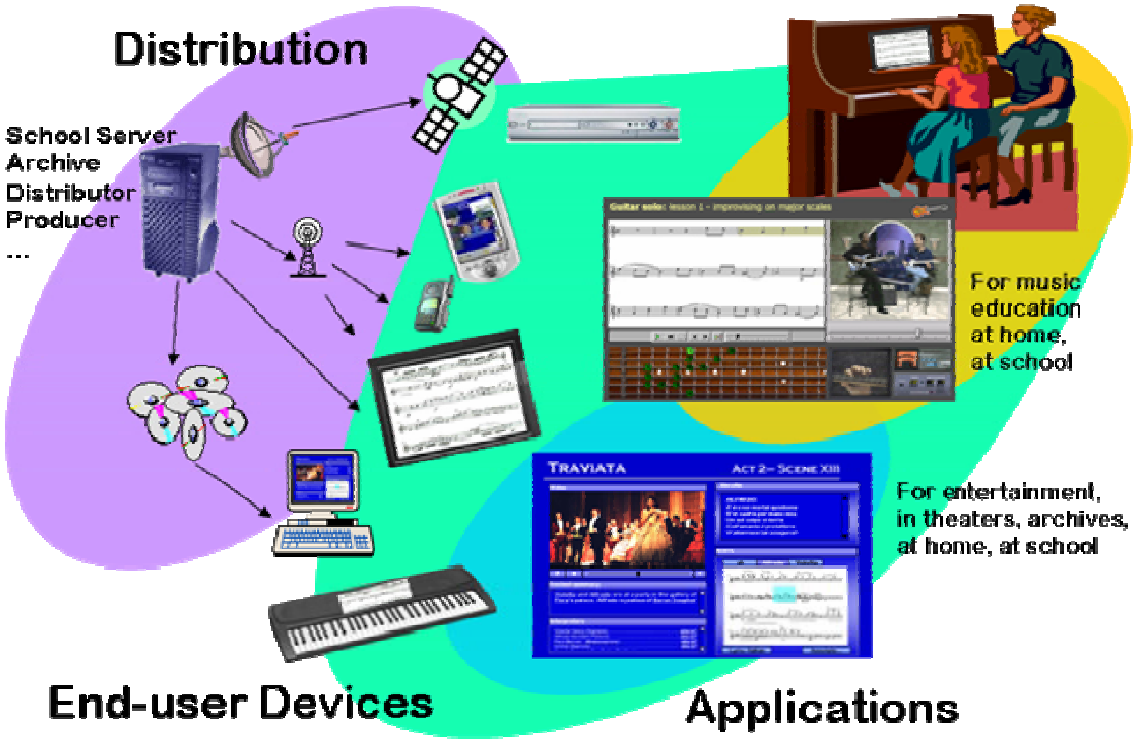


Fig.1 – General scenarios for MPEG SMR in entertainment and education. MPEG-4 with SMR support distribution by means of satellite data broadcast, internet, wireless or traditional communication and storage towards theatres, homes, archives, to devices such as i-TVs, tablet PCs, PCs, PDAs, smartphones.

2. MUSICNETWORK and MPEG

Since the beginning, one of the main aims, if not "the aim" of MUSICNETWORK has been to focus on stimulating the realisation of widely adopted formats for music notation representation. These formats, which must be integrated with multimedia applications and models, will deal with the needs of all the relevant actors (including publishers, music editor producers, copyists, integrators, etc.) involved in the realisation and the distribution of an "interactive" multimedia music piece. Music notation representation is an important issue, and an open standard format allowing exchange and cooperation with other multimedia formats still does not exist. We should not be limited to the applications only related to the printing process. It is clear that music notation needs to be and will be in the future accessed more and more using different kinds of devices, from the PC to Tablet PCs and UMTS terminals, from the classical printed music sheets to the electronic lectern.

The first task was the identification of the requirements for the integration of music notation with multimedia applications.

This work has been performed with the support of several experts at the MUSICNETWORK open workshops by means of the discussion forums and the mailing lists provided by the MUSICNETWORK portal. This analysis has been performed with considerations of the past experiences of several European Commission projects, including CANTATE, MOODS, WEDELMUSIC, IMUTUS, PLAY, PLAY2, etc., that worked in the area of music notation, and in some cases, on the integration of music notation with some multimedia content and features.

The second step was to study the state of the art technology in the area of music, computer music and electronics, to better understand the music notation/representation formats, their integration, and aspects on all the WGs involved in the MUSICNETWORK, and their usages in multimedia applications. These activities have been described in a number of deliverables and reports of the MUSICNETWORK that have been downloaded by thousands of participants from the MUSICNETWORK web site. It is evident that this work has received very

strong attentions and interests and it has been strongly appreciated by all related communities.

This second step has allowed us to understand the state-of-the-art and the real needs of the users and of the companies that produce music and computer music applications for the market, mainly in the areas of education, entertainment, content distribution, archiving and cultural valorisation, electronic consumer equipments (such as i-TV, PDA, cellular phones, and others), etc. At the same time, the experts of the MUSICNETWORK have identified the major problems that are preventing and/or limiting the exploitation of the present technical and technological solutions in those applications.

MPEG is probably the largest standardisation group that works on multimedia coding. It is a working group of ISO (International Standard Organisation, WG 29) and it is the producer of all the MPEG standards: MPEG-1 (coding for distribution at small bitrates, including the core of the mp3 format), MPEG-2 (audiovisual coding for digital TV and higher bitrates), MPEG-4 (multimedia coding for audiovisual objects, to be used also by satellite distribution, etc). At the forum, there are participants from all the major companies including IBM, MICROSOFT, HP, SAMSUNG, SONY, THOMPSON, PANASONIC, YAMAHA, SANYO, PHILIPS, SHARP, etc., under the umbrella of their respective National Bodies. Overall, MPEG includes all major companies involved in Consumer Electronics devices and technologies, and all the major research centers in the area. At each meeting, typically more than 330 partners are represented. These are the main motivations for which MPEG has been considered by the MUSICNETWORK as the best forum to propose and to create a Music Notation/representation standard with multimedia integration, because probably is the only forum in which that task could be performed at that level.

As mentioned before, one of the main limitations for the full exploitation of music notation/representation integrated with multimedia, is the lack of a common standard for representing the notational information. On the other hand, the presence of a standard is not the unique problem since presently there are some *de facto* standards that cannot be used and are not used for solving the above mentioned problem. These include representations from FINALE, SIBELIUS, etc. In fact, in most cases, these *de facto* standards are capable of supporting more than 95% of the global production of music notation pages. However, they remain unacceptable and incapable to be fully exploited in multimedia music applications. Another interesting case is that of the XML-based music notation formats. In the last 5 years, we have seen about 15 different XML-based models proposed by several groups and companies. Among these only MUSICXML of Recordare has gained interest demonstrating a quite interesting interoperability with several applications, including in some measure, FINALE and SIBELIUS. In this case, the MUSICXML remains a subset of their models, and it is not capable of modeling multimedia music concepts. This is confirmed by other efforts in the past in which we have seen other proposed standards as interchange formats for music notation such as NIFF and SMDL (SMDL was an ISO draft, to produce a standard on music notation). SMDL has been canceled several months ago. The same problem is evident from the effort for starting a standardisation process in IEEE, that could be presumably based on MPEG SMR.

The MUSICNETWORK activity aiming at the integration of Music Notation in MPEG started in May 2003, with the elaboration of a joint proposal for the MPEG meeting in July 2003, in Trondheim, Norway. In that occasion the aim was to demonstrate the effective need of a Music Notation in MPEG in terms of applications scenarios and requirements.
The MPEG group has agreed on setting up an Ad Hoc Group, which is, in MPEG parlance, a specific group aimed to study a particular topic. In this case, the Music Notation and its possible integration in MPEG. A mailing list (a reflector in MPEG parlance) has been setup, together with a web site to support this activity. The chairs of this AHG have been designated by MPEG to be Paolo Nesi (also chair of the MUSICNETWORK) and G. Zoia of EPFL, Switzerland. After that step by step and involving more than 60 different experts coming several countries the MPEG AHG on SMR has been grown and was capable of formalizing:

- Requirements for MPEG SMR;
- Call for Proposal, to collect submissions presenting technologies to be integrated into MPEG architecture;
- Assessment model for evaluating the proposed technologies.

This process has produced at the end of a selection the first Working Draft of the ISO standard on MPEG Symbolic Music Representation. That WD is still internal and accessible to all MPEG people and companies, since July 2005. The WD is also accompanied by the source code of a demonstrator that is an integral part of the standard.

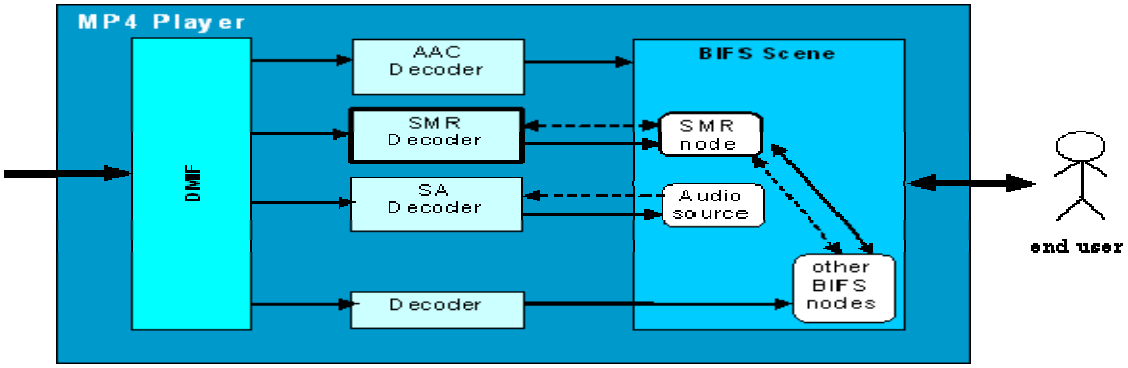


Fig. 2 -- SMR inside an MPEG-4 Player

According to the MPEG mechanisms, all companies interested in becoming compliant with a standard can access to the source code (that is called reference software in MPEG parlance) for creating their applications.

3. Integration of SMR in the MPEG framework

Symbolic Music Representation (SMR) will be integrated into MPEG-4 by:

- defining an XML format for a text based symbolic music representation, to be used for interoperability with other symbolic music representation/notation formats and as a source for the production of an equivalent binary information that may be stored in files and/or streamed through a suitable transport layer;
- adding an SMR Object Type for the delivery of a binary stream containing SMR, synchronization information, and rendering rules; the associated decoder will allow to manage the received information to add the necessary "musical intelligence" for the interaction with humans;
- specifying the interface and the behavior for the symbolic music representation decoder and its relationship with the MPEG-4 synchronization and interaction layer (MPEG-4 BIFS nodes)

The SMR XML content can be produced using appropriate converters and/or a native SMR music editor. Then, an MPEG-4 SMR-enabled encoder tool can multiplex the SMR XML file into an MPEG-4 binary stream (standard XML binarization is also available in MPEG).

The SMR binary stream contains information about music symbols, their synchronization with other media in time and space, and possible rendering rules for formatting music symbols.

A decoder in the user's terminal (player) converts this stream into a visual representation, which can be for instance rendered inside a BIFS scene. Figure 2 reports a simple example of an MPEG-4 Player supporting MPEG-4 SMR. The SMR node(s) in the BIFS scene is used to render the symbolic music information in the scene (this could be performed by exploiting functionality of other BIFS nodes) as it is decoded by the SMR Decoder. The end user can interact with the symbolic music representation (change page, change view, transpose etc.) through the SMR interface node, using sensors in association to other nodes defining the audiovisual, interactive content. User commands are sent out from the SMR node fields to the SMR decoder (dashed lines), which generates a new view to be displayed in the scene.

The structure of the SMR decoder is reported in Figure 3; steps of the decoding process include:

1. the **Binary decoder** decodes the binary stream; the decoder extracts the optional SMR rendering rules and the synchronization information from the SMR access units, loading the SMR Rendering Rules data structure to any SMR Rendering Rules engine, and sending the synchronization info to the SMR Manager
2. the **SMR Model** includes only SMR parameters, while the images, audio, video, etc. (other object types) are simply referred to other MPEG objects
3. the **SMR renderer**, controlled by the *SMR Manager*, uses the *SMR Model* with its *parameter values* and the *SMR Rendering Rules* to produce a view of the symbolic music information in the SMR Decoder Buffer.
4. the **SMR Decoder Buffer** may contain pixels and/or vector graphics information; this may be a solution dependent issue

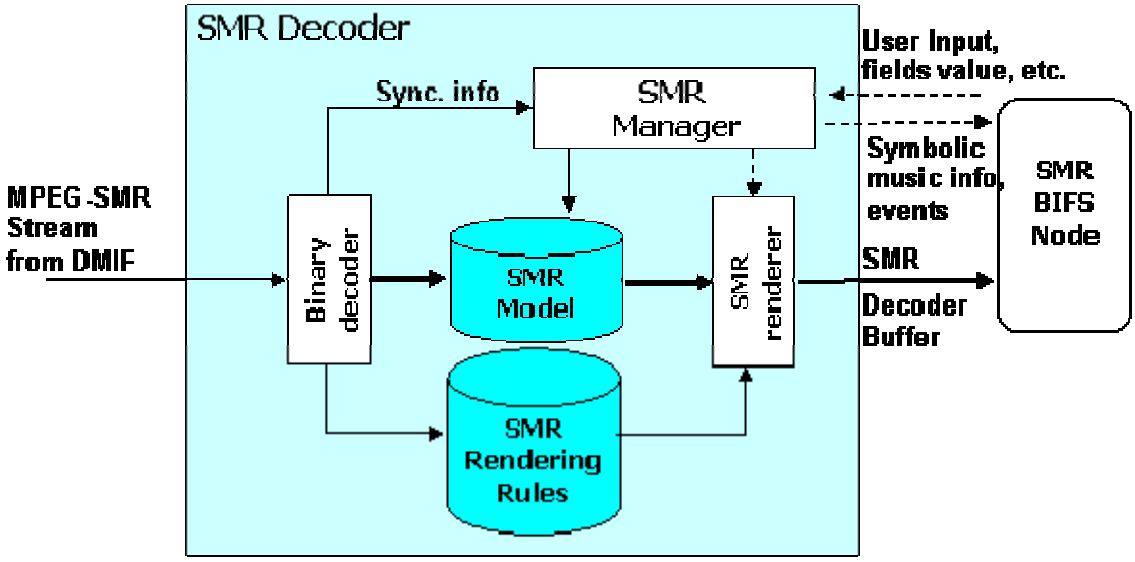


Fig.3 – The structure of a MPEG SMR decoder

5. the **SMR Manager** coordinates the behavior of the SMR decoder; it (i) receives and interprets the events coming from the SMR node interface. According to the command type, it can modify parameters in the SMR Model (e.g., transposition) and/or control the SMR Renderer (e.g., change view, change page, etc.), and (ii) it controls the synchronized rendering using the synch info
6. an **SMR node** and **other BIFS nodes** attach the content in the scene and specify the interface to the rest of the BIFS scene and to the user

A particular care is dedicated by SMR to the relationships with MIDI information; if a publisher wants to use only some MIDI files in MPEG-4 compliant devices (this is possible through the simplest Object Type defined in the Structured Audio subpart) and if these devices support SMR visualization, the specification will permit MIDI files to be automatically converted (through some specific algorithm) into SMR at the client and rendered. Similarly only the SMR may be available and delivered.

In those cases, the MIDI information can be generated at the client from SMR to be used with MIDI compliant devices. This is particularly important to guarantee straightforward adaptation of current devices.

Information about the ongoing MPEG SMR activity can be obtained from the web pages of the MPEG ad hoc group on SMR (http://www.interactivemusicnetwork.org/mpeg-ahg). A large collection of documents containing requirements, scenarios, examples, and links can be accessed easily.

4. Conclusions

Integration of symbolic music representation (SMR) in MPEG is opening the way to the implementation of a large set of new applications in the areas of education, entertainment and cultural valorization. Most of these applications are not available yet on devices accessible to the end user such as interactive TV, mobiles, etc., and those available on PC are not based on standard content formats, thus constraining producers to reshape any functionality from scratch by creating specific tools. This is a strong limitation for the diffusion of music knowledge and for the educational market of music. The integration of MPEG SMR is going to allow to code and distribute new extended functionalities that will be accessible for a larger number of citizens enabling the development of a number of innovative applications, from distance learning, to rehearsal and musical practice at home, and any imaginable form of music enjoyment on any kind of end user devices like those mentioned above.

Further information on MPEG SMR can be recovered from the web pages of the MPEG ad hoc group on SMR [23]: http://www.interactivemusicnetwork.org/mpeg-ahg. A large collection of documents which contain requirements, scenarios, examples, and links can be accessed easily. The MUSICNETWORK is now an international association with a range of partnerships

and memberships and many exciting activities, to continue building on the successful achievements of the project so far. If you are interested in the activities, memberships and services of the Association, we welcome you to join the association, to participate and involve in the activities and development of the MUSICNETWORK Association for the advancements and success of this interdisciplinary domain.

5. Acknowledgments

We would like to express our deep thanks here to all people who took part in the discussions on the reflector of the MPEG ad-hoc group on SMR, and to all those belonging to the MUSICNETWORK WG on music notation and those that attended joined meetings of MPEG and MUSICNETWORK. More in particular (in any specific order): Jerome Barthelemy, Kia Ng, Tom White, James Ingram, Martin Russ, Eleanor Selfridge Filed, Neil McKenzie, David Crombie, Hartmut Ring, Tillmann Weyde, Jacques Steyn, Steeve Newcomb, Perry Roland, Matthew Dovey, and many many others. We apologize in advance for forgetting to mention other names, if any.

10. References

[1] B. W. Pennycook, "Computer-Music Interfaces: A Survey", ACM Computing Surveys, June 1985, 17(2):267-289.

[2] CANTATE project, Deliverable 3.3: Report on SMDL evaluation, WP3, 1994. http://projects.fnb.nl

[3] CAPXML: http://www.whc.de/capella.cfm

[4] CUIDADO: Processing of Music and Mpeg7: http://www.ircam.fr/cuidad/

[5] D. A. Byrd, Music Notation by Computer, Department of Computer Science, UMI, Dissertation Service, Indiana University, USA, http://www.umi.com, 1984

[6] D. Blostein, and H. S. Baird, "A Critical Survey of Music Image Analysis" in Structured Document Image Analysis, (H. S. Baird and H. Bunke and K. Yamamoto, eds.), Springer Verlag, NewYork, USA, 1992, pp. 405-434.

[7] D. Blostein, and L. Haken, "Justification of Printed Music", Communications of the ACM, March 1991, 34(3):88-99.

[8] E. Selfridge-Field (Ed.), Beyond MIDI - The Handbook of Musical Codes, The MIT Press, London, UK, 1997.

[9] F. Pereira, and T. Ebrahimi (Eds.), *The MEPG-4 Book*, IMSC Press, 2002.

[10] Freehand: http://www.freehandsystems.com/

[11] G. M. Rader, "Creating Printed Music Automatically", IEEE Computer, June 1996, pp.61-68.

[12] IMUTUS: http://www.exodus.gr/imutus/

[13] J. S. Gourlay, "A Language for Music Printing", Communications of the ACM, May 1986, 29(5):388-401.

[14] K.C. Ng, (ed), Journal of New Music Research (JNMR) special issue on Multimedia Music and the World Wide Web, 34(2), ISSN: 0929-8215, Routledge, to appear 2005.

[15] M. Good, "MusicXML for Notation and Analysis", In *The Virtual Score Representation, Retrieval, Restoration*, (W. B. Hewlett and E. Selfridge-Field eds.) MT: The MIT Press, Cambridge, 2001, pp. 113-124. http://www.recordare.com

[16] MOODS project. http://www.dsi.unifi.it/~moods

[17] MPEG SMR AHG web page: http://www.interactivemusicnetwork.org/mpeg-ahg

[18] MUSICALIS: http://www.musicalis.fr/

[19] NIFF Consortium, "NIFF 6a: Notation Interchange File Format", 1995.

[20] Notation, http://www.notation.com

[21] OPENDRAMA: http://www.iua.upf.es/mtg/opendrama/

[22] P. Bellini , J. Barthelemy, I. Bruno, P. Nesi, and M. B. Spinu, "Multimedia Music Sharing among Mediateques, Archives and Distribution to their attendees", Journal on Applied Artificial Intelligence, Taylor and Francis, 2003, http://www.wedelmusic.org .

[23] P. Bellini, and P. Nesi, "WEDELMUSIC FORMAT: An XML Music Notation Format for Emerging Applications", *Proceedings of the 1st International Conference of Web Delivering of Music*, IEEE press, 23-24 November 2001, Florence, Italy, pp. 79-86.

[24] P. Bellini, Della Santa, R., Nesi, P. (2001). Automatic Formatting of Music Sheet. Proc. of the First International Conference on WEB Delivering of Music, WEDELMUSIC-2001, IEEE Press, 23-24 November, Florence, Italy, pages 170-177.

[25] P. Bellini, F. Fioravanti, and P. Nesi, "Managing Music in Orchestras", IEEE Computer, September 1999, pp. 26-34, http://www.dsi.unifi.it/~moods/.

[26] P. Bellini, I. Bruno, P. Nesi, "Automatic Formatting of Music Sheets through MILLA Rule-Based Language and Engine", under publication on Journal of New Music Research.

[27] P. Bellini, Nesi, P., Spinu, M. B. (2002). Cooperative Visual Manipulation of Music Notation. ACM Transactions on Computer-Human Interaction, September, 9(3):194-237,

[28] R. B. Dannenberg, "A Brief Survey of Music Representation Issues, Techniques, and Systems", Computer Music Journal, 1993, 17(3):20-30.

[29] R. B. Dannenberg, "A Structure for Efficient Update, Incremental Redisplay and Undo in Graphical Editors", Software Practice and Experience, February 1990, 20(2):109-132.

[30] Sibelius Music Educational tools: http://www.sibelius.com/

[31] SMDL ISO/IEC, Standard Music Description Language. ISO/IEC DIS 10743, 1995.

[32] WEDELMUSIC: http://www.wedelmusic.org

[33] Yamaha tools: http://www.digitalmusicnotebook.com/home/

Accessible Symbolic Music Representations

David Crombie[1], Roger Lenoir[1], Neil McKenzie[1]
[1]*FNB, The Netherlands*
Dcrombie@fnb.nl, rlenoir@fnb.nl, nmackenzie@fnb.nl

Abstract

In answer to the call for proposals for additions to a Symbolic music representation language within the MPEG 4 framework(ISO/IEC JTC1/SC29/WG11 N6689), FNB have started a Core Experiment in order to integrate Accessible Music notation renderers within this new technology. This paper outlines the background for this work, describes the Core Experiment and then looks at some of the intended future work.

1. Introduction

The FNB International Projects Department[7] has been involved in a large number of European Commission funded projects over the last 10 years. These projects have addressed many different aspects relating to design and accessibility of materials and information for the print impaired population throughout Europe

FNB submitted a Core Experiment to MPEG following their participation in the MUSIC NETWORK project[8] . The Music Network is a thematic network, funded by the European Commission Fifth Framework IST Research Program, was being established in the area of music coding. The main aim of the network is to bring music into the interactive media era.

Much of FNB's accessible research is disseminated and carried out through the EUAIN Project.[4] This project aims to promote e-Inclusion as a core horizontal building block in the establishment of the Information Society by creating a European Accessible Information Network to bring together the different actors in the content creation and publishing industries around a common set of objectives relating to the provision of accessible information. For more information about the EUAIN project and FNB's work in Accessibility in general please visit and get involved at http://www.euain.org

2. Accessible Music and SMR

2.1 Accessible Music

The Accessible Music Software Suite (AMS) was designed following the meta-modelling principles found within the MPEG family, with particular attention paid to compatibility with emerging DIA(Document Type Adaptation) infrastructures within MPEG 21. This means that the software can be easily adapted to provide an integrated decoder. A brief description of what is required for this process is given below.

Braille Music conversion takes place based on the International Manual of Braille Music Notation, but due to the ever-changing requirements of this format the user preferencing elements must remain extensible. Talking Music provides a spoken description of the elements of the score. These spoken descriptions are provided in such a way that the information is compressed as much as possible to ensure the spoken elements provide usable information and the descriptions do not become unwieldy. The format has proved very popular with print impaired users, who in the past have either had no access to scores or have had to contend with the logistic problems of traditional production methods associated with Braille Music. Talking Music has therefore become valuable to many users as both a learning tool for music and also as a means of navigating through all the elements found in a traditional score.

Given the ever-changing requirements of music representation, the interfacing with accessibility tools is constantly set back. With every modification of the models that are used for music analysis, representation and synthesis, additional effort has to be invested to synchronise the consumption and production

opportunities for print impaired users with those of the average end-user. However, accessibility should be an integral component of any system, and where such a component is considered integral to the design process, the resulting system benefits on many different levels. For music production and consumption systems, a naturally available transformation and representation feature can replace the 'workaround' nature of traditional accessibility enhancements. Our approach seeks to provide these new opportunities.

It is intended that through additions to the SMR(Symbolic Music Representation) RM0(Reference Model Zero) software that the available software for decoding Accessible Music notations can be improved in terms of extensibility and adaptability for the future. This is mainly through the additions of modules within the SMR core which allow the specification of parameters for user settings to be defined. These design mechanisms will ensure that a means of preference-setting can be available on reduced functionality systems of the future.

2.2 Symbolic Music Representation

MPEG-4 provides a framework for encoding multimedia content through object types and scene descriptions. This builds on the MPEG 1 framework which was designed to as a basic audio and video compression standard, and is built for greater integration with the MPEG-7 tools for content description(As opposed to content encoding).

The MPEG framework does not however provide support as of yet for symbolic music representation. Symbolic Music Representation would provide a standard for encoding music notation ina significantly intelligent manner that many new user applications could be made possible in the areas of entertainment, music sheet production, music teaching, music analysis, content query, provision of enhanced or adapted music for consumers with specific needs, etc.

Symbolic representations of music have a logical structure consisting of: symbolic elements that represent audiovisual events; the relationship between those events; and aspects of rendering those events. There are many symbolic representations of music including different styles of Chant, Renaissance, Classic, Romantic, Jazz, Rock, Pop, and 20th Century styles, percussion notation, as well as simplified notations for children, Braille, etc.

An Ad-Hoc group was created by MPEG on the request of the MUSIC NETWORK group in order that specifications and requirements could be discussed which could lead to a Symbolic music representation language being included as an addition to MPEG

One of the main advantages of integration of Symbolic music representations within the MPEG framework is that SMR elements can be synchronized with audio visual events which are created using existing MPEG technologies.

The integration will allow the interoperability of different music applications. This can take place due to the breadth of MPEG standards for multimedia representation. Music is a rich body of information, which requires a solid framework in order to provide for the use cases generated by the different perspectives taken on music.

3. Accessibility

3.1 Accessible Design

For some time now FNB International Projects Department have been carrying out projects which look at accessible deisgn at a more macro level. This builds on the tenets of Design for all[5] to move towards an environment where Accessibility can be seen as a process rather than a products.

Much of the work of the EUAIN[4] project targets this objective, where the focus moves to the process of accessible information Processing and ensuring interchange of expert knowledge between all the various parts of the processing chain. The aim of the EUAIN consortium is to integrate accessible notions and initiatives at an earlier point in the chain.This requires a review of the notions of Accessibility from Scratch[2] and Openfocus which are covered in greater detail elsewhere.

Accessibility from scratch introduces the concept of building accessibility into frameworks from the ground up. If accessibility is included as a system component within a robust foundation then the advantages of Interoperability, scalability and extensibility are intrinsic to the system.

Openfocus describes the wider scope which has to be taken in order to see the macro level picture of a situation or system in order to build for the process rather than the product. The objects in the system have to be build in such manner that they can be built for future extensibility and adaptability.

3.2 Accessible Music.

A brief description of Talking Music and Braille Music formats is provided in section 2.1 above, but

from a more methodological perspective, it is important to understand some other concepts about the design requirements of Accessible music.

Accessible Music systems(Or any music system) is based on affording the user the choices and preferences for ensuring that they can process, consume or create music for their particular specialist needs. In this way the idea that every user of music systems has a set of specialist needs reiterates the idea that accessibility is about communication and accessible music is just another set of needs to add to the myriad of possibilities.

In order to provide such a rich set of options for the user, it is essential that every option and preference is represented within a structured model and available for query at a system level. In this way the internals of the system provide a foundation for building a usable system. Such a usable system then requires the expert knowledge of use cases in order to present suitable options to the user and not overwhelm them with choices which aren't relevant to their particular set of specialist needs. The information is then processed or filtered in a non-destructive manner to ensure that the information is specialized to their perfection but still suitable for transfer to an information set based on a different set of preferences as the information to do so is present.(Albeit intrinsically)

4. Core Experiment

4.1 Description

Given the ever-changing requirements of music representation, the interfacing with accessibility tools is constantly set back. With every modification of the models that are used for music analysis, representation and synthesis, additional effort has to be invested to synchronise the consumption and production opportunities for print impaired users with those of the average end-user. However, accessibility should be an integral component of any system, and where such a component is considered integral to the design process, the resulting system benefits on many different levels. For music production and consumption systems, a naturally available transformation and representation feature can replace the 'workaround' nature of traditional accessibility enhancements. Our approach seeks to provide these new opportunities.

It is intended that through additions to the SMR RM0 software that the available software for decoding Accessible Music notations can be improved in terms of extensibility and adaptability for the future. It is the intention of this core experiment to prepare ground work and analysis of the RM0 such that talking music and Braille music renderers can be integrated with the SMR framework.

As output from this core experiment, it is intended that changes can be proposed to the RM0 which allow parameters to be set for user preference setting for both Talking music and Braille music renderers.

4.2 Objectives

The objectives of this core experiment are as follows:

- To check if the present problem can be solved with the present SMR technology and solution.
- To define an XML listing for Parameter settings based on the Talking Music and Braille Music requirements specified in ISO/IEC JTC1/SC29/WG11 M11542.
- To enhance supporting documentation and developer documentation for the Accessible Music Decoders.

The experiment will be jointly performed by all participants within the Ad-Hoc SMR group and co-ordinated by FNB. The CE will start after the meeting in Poznan(July 2005) and Results of stages M2 and M3 will be presented at Nice, France(74th MPEG Meeting October 2005). The core experiment will end at the 75th MPEG meeting in Bangkok in January 2006.

A Core Experiment in general last fro around 6 months and performs the work of 1 or 2 developers in order to answer a call for proposals. In this case the call was for additions to the SMR RM0 software. One of the objectives of the Core Experiment is to decide if it is possible to use the SMR to enhance the current accessible music solution. The proponents of the Core Experiment are confident that the two systems can be good bedfellows, as much of the issues discussed above were raised during the specification of the RM0 system.

In answering the call, the main objective is to define what additions need to be made to RM0, in order that the accessible Music suite can be suitable integrated with it. The main additions, which are envisioned, relate to means of specifying rendering hints which can be stored within the SMR representations. These can then be read by a renderer which uses them as defaults for creating accessible music renderings where there is no ability to set preferences and user requirements.

Until now, the accessible music suite, while available has required installation and training in order to add it to the production chain in specialist orgoantions. It is envisioned that with further documentation on both

developer and user elements of the system that the system can be installed in priduction chains with the minimum of technical support.

This element of the Core Experiment will take place through the AccessMusic sourceforge project.[ ]

5. Future work

5.1 MPEG 4 SA

In order to be able to meet the preferences of several niche markets within the (accessible) user groups, every aspect (as far as possible) of the software must be extensible. One simple example of this in the current solution is that of the multilingual software support. The software creates Talking Music independently of an end-language or dictionary, and this means that any dictionary can be used. As a result, in order to add a further language output to the software, the only requirement is to create another DTD (Document Type Definition) file which translates the musical terms required (approximately 200 words). The same extensibility is possible for many aspects of the software, including Braille Unicode definitions; the grammar used in Braille and Talking Music scores; and the meaning and use of various symbols (although these may require a rudimentary knowledge of XML grammars and code).

In order to create accessible software, the designer must endeavour to be as adaptable as possible. The primary output format of the Talking Music software is the DAISY Talking Book, following the NISO standard[1]. This is a SMIL based structure for presenting spoken audio information. Currently the system is being adapted in order that the SMIL structure can be recreated in the XMT-A/O format. This allows the functionality of DAISY to be recreated within the MPEG framework. By way of illustration, we can see the similarities between XMT and SMIL:

```
<smil>
  <head>
    ...other head elements...
    <layout>
      <region id="txtView" />
    </layout>
  </head>

  <body>
    <seq dur="4.024s">
      <par endsync="last">
        <text src="ncc.html#ec001"
id="info0001" />
        <audio
src="please_insert_cd_1.wav" clip-
begin="npt=0.000s" clip-
end="npt=1.456s" id="info0004" />
      </par>
    </seq>
  </body>
</smil>
```

```
<XMT-A>
      <Header>
      …
      Header
      …
      </Header>
      <Body>
            <par begin=”0.0”>
                  <Replace>
                        <Scene>
                        …
                        Scene
                  information
                        …
                        </Scene>
                  </Replace>
            </par>
      </Body>
</XMT-A>
```

Connection and coupling to MPEG 7 (via the Object descriptor framework commands) data description systems would allow future databases to interact with any solution, and the future solutions within the MPEG framework can then be coupled using the capacity of XMT-A to support BIFS.

5.2 Further internationalization

During the design of the Accessible Music production suite it was always envisioned that other organisations who produce accessible music notations would be able to adapt the software to their specific needs.

One of the easiest of these adaotations would be to adapt the software and the format into other languages for use ion other countries. Currently the software is available and in use in both Dutch and English, but through addtion of a set of definitions in a DTD(XML Document type Definition) file.

[1] http://www.daisy.org/publications/specifications/daisy_202.html

It is envisioned that anyone who is interested in adapting the software for their language can do so through the AccessMusic sourceforge project[1]

5.3 Integration with future tools

Currently the SMR technology within MPEG is quiet young, but as it matures there will be further development required to integrate the technologies with future tools.

It will be interesting to see what possibilities arise from these new tools for rendering and producing symbolic music representations with an eye to provide further tools to end users of Accessible Music formats. When these technologies mature, there are well established end user groups who will be pleased to help define use cases and requirements for such systems

One area receiving some attention within the accessibility world is that of metadata. In information processing, metadata becomes the essential tool for deriving information for interchanging this information with other formats. It therefore follows that Metadata is an essential tool for processing accessible information, where many formats rely on context and surrounding information in order to build alternative representations of the data.

The CEN/ISSS Workshop on Metadata for Multimedia Information – Dublin Core was held in 2004, organized by the Dublin Core Accessibility Working Group (part of the DC Metadata Initiative). This workshop put the stress in both multilingualism and accessibility. In its final recommendations, the Workshop advocated for the creation of a new element for Dublin Core metadata, *DC:Accessibility*, to describe accessibility of resources and services.

The aim of the CEN/ISSS MMI-DC workshop (see 3.4 above) was to identify and investigate the ways in which metadata can help achieve efficient and future-proof solutions to accessibility. It is assumed that this encompasses the provision of adequate access to information for people with disabilities and for everyone in a multilingual and multicultural environment. In order to make this perceived information useful, it must be represented within an architecture which allows the accessibility requirements to be questioned in more than one way. Such an architecture must enable both the core system to adapt to new and changing representation requirements, and to allow (theoretically) infinite user requirements.

It is envisioned that following this core Experiment further work can be scoped which ties in some of the tools used by MPEG 7 technologies and tools with some of the standard which are emerging from the accessibility world. This will enhance the user experience by allowing more intelligent searching and exchange mechanism and move towards an environment where Accessible music notations are significantly information heavy that they can be integrated with Content Management systems[2]

6. Conclusion

The work of the core Experiment, although simple will provide a basis upon which accessible music notations can be integrated within the MPEG 4 framework. It is intended that this work is built upon in order to move towards a sustainable and adaptable system for future accessible music production.

If anyone would like to get involved with the core experiment or any of the planned future work surrounding this core experiment, please get in contact with the authors of this paper.

7. References

[1] AccessMusic Project:
http://accessmusic.sourceforge.net

[2] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Accessibility from scratch : how an open focus contributes to inclusive design, Proceedings ICCHP, Lecture Notes in Computer Science, Vol (3118). Springer-Verlag, Berlin Heidelberg New York

[3] The DAISY consortium: http://www.daisy.org

[4] The EUAIN Project:
http://www.euain.org

[5] European Institute of Design fro all: http://www.design-for-all.org/

[6] Fifth Framework program:
http://www.cordis.lu/fp5/home.html

[7] FNB International Projects Department:
http://projects.fnb.nl

[8] The Music Network:
http://www.interactivemusicnetwork.org

Meaningful mapping: Appreciating music through accessible structures

David Crombie[1], Roger Lenoir[1], Neil McKenzie[1]
[1]*FNB, The Netherlands*
Dcrombie@fnb.nl, rlenoir@fnb.nl, nmackenzie@fnb.nl

Abstract

This paper examines in some detail the requirements for understanding and implementing advanced modelling concepts within music applications and emerging multimedia and networked environments. Achieving innovative and flexible modelling techniques for learning, exploring, providing, composing and performing music, requires a high level of adaptability of these musical structures. To allow the emergence of understandable and therefore useful musical structures and musical structure processing logic, we need to enable the users to associate themselves with the musical structures. To be able to re-use these insights, we need to provide entities that represent the components involved in this association process explicitly. Through this representation they can be addressed directly. The modelling layer on top of these entities allows a high level of adaptation, personalisation and specialisation that can be described separately from the lower executive levels of the music representation architecture.

1. Introduction

Talking to one another other efficiently requires a level of abstraction. In order to capture relevant features that are important for *both* ends of the communication line, features of *both* participants should be present in the 'system' that facilitates this communication. Various parallel processes are entwined and efficiently unpacked *if* the participants communicate consciously. All these processes serve the same aim: to get a message across and make sure the message is understood. During a musical performance these communication strategies are elevated to the highest level, especially in the context of improvised music performances. Again there appear to be parallel and entwined processes, each of which regards the central theme - the message in the music - in its own specific structure. If the communication process is performed efficiently and appropriately, the quality of the performance will radiate to the outside world. The *audience* will perceive this and will be enabled to participate in this process as well. This relationship between communication structures is the complexity which makes every performance different, providing the routes are open to interaction.

The problem with interaction is the complexity which develops. How do we conceive a firm yet flexible base of notions that can be used to represent the resources and participants of this process in such a way that it facilitates simulation of radiating performance qualities? Should we base the representation system on the observers only? Only *one* observer? Or more? Which type of observer? Observer as in performer, composer, audience, or transcriber? And which manifestation of the performed or composed content should we use? A score, 'source code', a graphical notation in any format, an audio recording? Or should the representation system focus *exclusively* on the connections between various entities? From our experience in making music more accessible [1,2,3], it is our contention that in order to achieve a representation and modelling system that exhibits sufficient flexibility, a system should contain *all* the components mentioned above. Communication of content of *any* form relies on an interplay between all the entities that are relevant to the communication process. All these entities should be accessible from one another and in this way form a communication network.

2 Modelling Musical Knowledge

As described above, representation systems rely heavily on choosing appropriate representation components. The aim of a holistic representation system is to preserve as much of the initial

communication process as possible. This requires an architecture that limits the filtering of information structures to a minimum.

There are many ways to describe communication processes with a focus on musical communication. Any scenario, other than the interactive performance use case described in the introduction, involves parallel description processes that all require their own structuring paradigms, that is algorithms. Every music analysis algorithm presents its own advantages and limitations. Each specific need with its own specific aim requires its own specific analysis algorithm and with that an environment that allows the execution of such an analysis process. An algorithm that - as an end result - allows parameterisation of only *one* specific musical parameter. Given that a communication process, especially a musical one, relies on the interplay between various parameters, we run into a fundamental problem. This fundamental problem keeps us from using technology, and the singular parameter focus behaviour it exhibits, in a natural and intuitive way.

This problem is particularly noticeable where creativity meets technology. Traditionally, creativity is performed by one set of users in various "communication" languages, and technology has its own set of users and languages. Only in niche markets like accessibility can we find people who traverse the two areas. The problem of interfacing the two domains becomes a communication issue. We believe that the only means to communicate between two (or more) such domains is by using computer modelling techniques in order to find a "common language" or common ground.

2.1 Mapping meaning: Input, representation, output

We can deduce three processing layers that form the basis of any modelling structure, which in themselves are usually systems. There is an input layer and output layer and these centre around a representation layer, as in Figure 1 below:

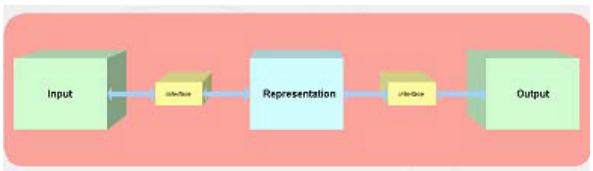


Figure 1: Basic representation model based around three abstract processing stages

It is important to realize that the model is not linear but a loop where like any consumer/producer models, each input can also be an output and vice versa. The basic diagram is theoretical, as most systems will have multiple inputs and outputs, but the fundamental concepts are the same.

2.1.1 Input: perception of musical data

In order to process information, there must be some sort of input of information. This is performed by a perception layer. In the case of a musical system this can be any sort of musical input – OMR [4], midi, music notation files, or more abstractly the perception of music by any cognitive means, such as using the cognitive abilities of a computer or a human [5]. It is important at the input level to include all the possible information, even if it is not relevant for the primary objective of the system. This ensures that the input, which can be the output of another system, does not cause the rest of the system to inherit preconceived losses in information density.

2.1.2 Representation: (re)structuring into musical information

This is the most important layer of any model which is used to restructure the information in a form which can communicate between the various inputs and outputs required by the system. This layer forms the information into a common language which is suitable for communication by both the input and the output. On a more abstract level it could be said that the advance study of any field results in the study of structure: Computing (Object orientation), Biology (Genetics), Musicology (Shenkarian analysis) and so forth [6].

2.1.3 Output: merging information and responses into musical knowledge

At output, the information set out by the structure in the representation layer is chosen and restructured into the user (or next input) requirements set out in the configuration of (this particular) output. The output layer can be seen as an instance of a perception of the represented information. In the case of music, it is a perception suitable for the end users who interact with that particular music. For example, to provide accessible solutions for print impaired users, this output representation of standard musical information becomes Braille Music [2] or Talking Music [7].

2.2 Meaningful mappings: interfacing the three processing layers

In order that the foundations of a system based on the three block model can be utilised to their full potential, it is essential that the three blocks can be sufficiently well interfaced to ensure that there is no loss of information throughout the model. At this point the previously rudimentary and simple system can become abstract and complicated. Interfacing is a communication paradigm which requires implementation throughout the system.

In order to interface a piece of information, a single "interface", possibly viewable as an axon, is responsible for ensuring that the correct viewpoint of the source information is taken in order for the destination "accessor" to understand and accept the information entity into its structure.

It is essential that at all points in such a system the information can be viewed by several interfaces simultaneously. For this to be achieved, the information should not be structured in such a way that it is focused to a specific primary application – as many secondary applications (i.e. accessible solutions and initiatives) are thus ruled out. In order to do this at a coding level, the classes defining the information entities should be completely defined on the source information rather than the destination information. If this is intuitive, it is valuable to extend these classes to ensure that the source information can communicate with interfaces in a multitude of ways, which encourages the idea of multiple interfaces on a single object, as illustrated below:

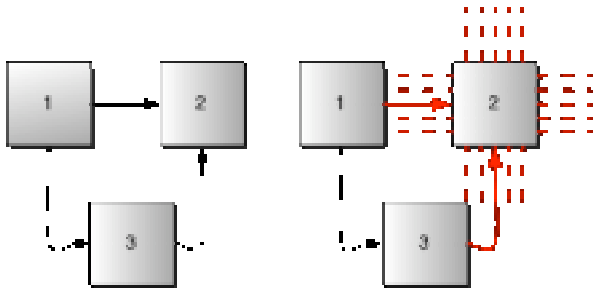


Figure 2: Illustration of multi-faceted interfacing objects

3. Accessing Musical Knowledge

In order to build a system for musical knowledge on the above basis, it is essential that the requirements for input of music material, representation of musical material and user/dissemination at the output level of the system are known. Music is a rich corpus of data, and almost every user of musical information and knowledge uses it in a slightly different way. In order to cater for these specialised uses of the various dialects of musical information, it is essential that a representation framework offers a rich gamut of viewpoints and perspectives on whichever genre or flavour of musical knowledge is catered for to the satisfaction of the user or the interfacing entity. Musical Knowledge can be considered to be represented implicitly and explicitly using procedural descriptions of information entity dynamics and declarative descriptions of facts - or static entities [8,9].

3.1 A user's perspective

In order to cater for the destination requirements of the user information it is first sensible to analyse the requirements of the traditional use and cognition of musical information. The users of music as an information body vary almost as much as the categories of music. Music is used in a variety of forms in every market imaginable, from consumer to research to education to industry.

3.1.1 The composers : specifying musical structures

The first port of call for discussing user requirements is the users of musical notation who create the content themselves. To the composer, musical structure is their canvas, and musical entities (notated or otherwise) are their palette. Traditional composition is extremely structured, and from a modelling point of view falls into simple hierarchies and structures at every meta level: Intervals; Key signatures, Time signatures; Tertiary/Symphonic form; Choir orchestration. Since the twentieth century, musical structure and form has struggled suitably to embrace the modern form of Avant Garde, Serialism.

The composer's imagination is the force that pushes the musical structures onto a imaginary canvas: the score. Once conceived and created a score can then be considered to be a snapshot of the composer's mind, rendering the dynamic and volatile musical concepts, movements and mappings static in the form of

symbols. The composer's imagination permits translations of these static indications into audible sounds, phrases and gestures. The composer may convey additional directions on these manifestations of his or her imagination or they may decide to leave the transformation into the audible range of the imaginary spectrum to the performer of the score. From the systemic perspective the translation from the score -as a transcription of the composer's mind and artistic and/or conceptual intentions- to the performance and the interpretation phase is a requirement. It can be considered to be the vector between two points.

The decision on the extent of freedom a performer has, depends on the composer *and* the performer. The composer may add huge amounts of additional descriptive material (for example Karlheinz Stockhausen), barely leaving space for personal interpretation. Within the information science idiom we would call this an overload of redundant information that leads to minimal compression of the message. The composer may provide a simple score that even gives the performer the space to improvise the musical structure itself. Here the provided musical structure (the score) could be considered to represent the composer's Ursatz, which permits extension by the performer. Alternatively, a performer may decide to completely ignore the composer's wishes and perform and change the musical material as they see fit. Taken together, this represents freedom of interpretation. In all these scenarios we can consider the composer as providing musical structures (organised sounds) that may be communicated through notated score [7]. The exact language or format that will be chosen to communicate the composer's thoughts depends on the composer's needs, alongside the consumer's needs.

3.1.2 The performers : interpreting musical structures

Performers of musical scores *interpret* the musical structures they are provided with building on their frame of reference; frame of competencies; their instrumental skills; and finally, their taste. The combination of these components determines the level of freedom performers can permit themselves to transform the musical structures from static structures into audible dynamic ones. This freedom of interpretation and transformation may be seen as a reflection of the performer's virtuosity. However, for performing virtuosity to emerge, a score - as in a complex of musical concepts in whatever static or dynamic notation - that facilitates the exposure of this process of imagination requires an inspiring and accessible interface.

3.1.3 The distributors : providing interpreted musical structures

Providing scores that include the interface that leads to intuitive performance and consumption can be regarded as an interpretation process itself. Providing space - physically or virtually- to perform or present the score and its performed interpretation involves various processes, such as: performance venue selection; recording techniques; marketing and public relations; target audience selection; catalogue building and so forth. All these 'parameters' are important ingredients in the provision of a useful interface from the composer to the performer to the audience. A high level of intuitiveness will ease the interfacing between the various components that govern the facilitation of an infrastructure that provides transportation and transformation of musical structures in all its incarnations [11].

3.1.4 The audience : digesting interpreted musical structures

Freedom of choice, ease of choice and level of appreciation are the key requirements for the audience. The sense of respect they feel for the composer, their community and their individuality all play a role, as does the level of 'uncoloured' interpretation.

3.2 An entity's perspective

How do computer information architectures respond to and cope with human communication? How do we ease the level of interfacing between the human user and the sterile computer system? Should the user adapt wholly to the computer architecture or should the computer architecture be flexible? Does the computer architecture reflect the requirements for intuitive and respectful communication? Once again, interfacing and communicating are the key factors.

In order to be able to interconnect the applied analysis algorithms (including the visualisation of the musical materials), flexible architectures are required. Not only the architectural components need to be flexible in their use, but the results they yield need to be flexible in their channelling and re-use. We need abstract entities that can be used to represent algorithms and their results and allow groupings of these resources into collections which can then represent theories and/or approaches. A flexible and re-usable methodology for the representation of the above can be found in Minsky's work (*op. cit.*). Building a representation system that builds on the notion of

agents and agencies allows the creation of an extensible framework of components.

4. Understanding Musical Knowledge

Musical structures may seem to be static. There is a score on paper or projected onto the screen. However, musical structures may become WHAT??? because of the involvement of the user's perception. Each particular *use* appears to have its particular way of restructuring the apparently static musical structures and any particular use influences, or is the result of, the user's frame of mind.

Associating the usage scenarios, such as taste, education, history and so forth, could represent the user's experience. Through the representation of the user's experience using this music representation strategy implemented in a software architecture, we can associate all the particular visions on musical structures with one another. They may exist in parallel and may provide a framework with resources that will enlighten the *meaning* of the musical structures by allowing *all* the visions, described in algorithms for analysis and synthesis, to exist.

4.1 Adaptability and personalisation: musical meta-modelling

Achieving innovative and flexible modelling techniques for learning, exploring, providing, composing and performing music, requires a high level of adaptability of these musical structures. To allow the emergence of understandable and therefore useful musical structures and musical structure processing logic, we need to enable the users to associate themselves with the musical structures. To be able to re-use these insights, we need to provide entities that represent the components involved in this association process explicitly. Through this representation they can be addressed directly. The modelling layer on top of these entities allows a high level of adaptation, personalisation and specialisation that can be described separately from the lower executive levels of the music representation architecture.

The meaning is represented implicitly through the perception of the structures and the dynamics between the structures used to represent the musical information. The meaning of music can be seen as an emergent feature that is closely related to the user's experience –it is influenced by the users' perspectives at any and all instance(s) of time. For each individual user type to form through *experiencing* the structures, the architecture that provides these structures is required to provide access to all *possibly* important features of the musical structures. Allowing any system to independently represent, associate and with that model the users' experiences and the information, structures of information and the dynamics between the structures of information, allows modelling of these perception mechanisms. Modelling as in learning from them because they now explicitly exist and modelling as in exploring their meaning and expanding the representation framework where new insights were obtained (see Figure 3).

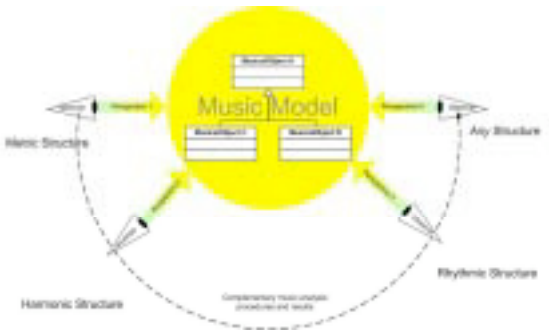


Figure 1: Association of multiple viewpoints onto a music representation framework yields representation of various user views including their preferred (or required) parameters. It also allows complementary use of multiple concurrent musical structures.

The meaning of the music is something that should be detached from the musical representation architecture, since the implementation of such meaning - and with that the computer program, its concepts and the consequences of this interpretation - would be based on the definition of that meaning at that time. The system would be incapable of allowing any user to form their own personal interpretation of the meaning by default. Multiple definitions of meaning should be permitted to co-exist within the same framework. The entity that distinguishes the meanings is the entity that represents a particular user: that is, the eye of the beholder.

4.2 Parameterisation of user or application requirements

Each user perspective may now represent a specific point of view. An algorithmic point of view, since the procedures that define the user's point of view may be described using computer code and can be associated with all the other entities that make up the process of personalised perception of the musical material. By using these procedural descriptions of user requirements

that are based on the specific users' preferences, requirements due to physical incapacities or due to mental incapacities for example can also provide *parameters*. These parameters can be used for addressing and associating musical or other behaviours to the musical structures filtered *through* the user's ambassador in the software architecture.

5. Conclusion

Such structures describing musical information provide a network of inter-related entities that allow attachment of user representation entities to these musical structures without imposing a pre-defined view on these musical structures. Each user and application entity can keep track of its own local history in its own model of time. Each user or application entity can provide its insights and interpretation back to the music representation framework, or any designated module that handles 'opinion events' from the user application module. These perspective-specific decisions and responses may then be channelled back towards the originating user or application, or for sharing insights, stimulating learning or encouraging innovation.

An *Accessible Musical Knowledge Framework* is not only suitable for representation of musical structures. Accessing musical information requires entities that *access* the musical information to be present. These entities may represent 'common' end user requirements, or 'specialised' end user requirements, such as print impairment, dyslexia or age related issues. This set of users may also include composers, distributors, sellers, performers, researchers, scholars, librarians, and so forth. The only difference is the direction of the information flow. Since such a framework aims at interactive relations between entities (as in *bi-directional*), both ends of the communication line are represented in the framework and advanced synchronisation features are a built-in feature of the system.

6. References

[1] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Making Music Accessible: an introduction to the special thematic session, Proceedings ICCHP, Lecture Notes in Computer Science, Vol *(in print)*. Springer-Verlag, Berlin Heidelberg New York

[2] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Designing accessible software for music applications, in Hersh, M., & Johnson, M. (eds) Assistive Technology for vision impaired and blind people, Ch. 16, (*forthcoming*) Springer-Verlag, Berlin Heidelberg New York

[3] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Integrating music representations within MPEG environments, in Proceedings 4th International Conference on Web Delivering of Music, Barcelona, IEEE (*forthcoming*)

[4] Ng, K. C., (2002). Music manuscript tracing. In *Graphics Recognition: Algorithms and Applications*, Springer-Verlag. 2390: 330-342.

[5] Minsky, M. (1987). Society of mind. Simon & Schuster, New York.

[6] Crombie, D., Lenoir, R., and McKenzie, N., (2004) On Scoping Abstraction**,** in Proceedings Extreme Markup Languages Conference 04, Montreal (*forthcoming*)

[7] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Accessible Music Technology**,** in Journal of New Music Research (*forthcoming*)

[8] Desain, P. (1995). LISP as a 2nd language. NICI, Nijmegen.

[9] Desain, P., Honing, H. (1992). Music, Mind and Machine - Studies in Computer Music, Music Cognition and Artificial Intelligence. Thesis Publishers, Amsterdam.

[10] Crombie, D., Lenoir, R., McKenzie, N., and van Peursen, D., (2004) Towards coherent publishing platforms, (*forthcoming*)

MPEG Symbolic Music Representation and Music Education Software

Tillman Weyde
City University
Department of Computing
London, UK
t.e.weyde@city.ac.uk

Abstract

Music education software is an important application area that will benefit from the development of an MPEG standard for Symbolic Music Representation (SMR). Although there are many existing representations, they lack necessary features for educational purposes and are mutually incompatible. This paper gives a survey of the state of the art and MPEG SMR features in relation to music education software, and it discusses perspectives for development based on completed and current projects.

1 Introduction

Developers of music education software face the problem of appropriately and efficiently representing the subject of study. Music is represented on different levels, mainly audio, performance, notation, structural, metadata, and annotations. Music education software needs not only to play back and display musical content, but it needs to represent music on several levels interrelated levels order to process user input as well as exercise and test material. This can be compared to teaching mathematics, where it is not sufficient to know the right result, but to have a model of the the terminology, calculation method, and operations involved.

Currently there is no standardised representation of music education contents. Most available music education softwares are developed as a one-off design. For efficient development of music software, reusability and tools on high abstraction levels are necessary. The MPEG Symbolic Music Representation, as it has been laid out in the Requirements and the Reference Model 0, aims at the integration of music notation and structural information into MPEG standards, thus enabling music education software to make full use of MPEG compliant contents and tools.

2 SMR in Music Education Software

Music education makes almost always use of some form of music symbols to denote musical elements and structures, as it is a central aim of music education to enable written (and spoken) musical communication. Most music education software concentrates on introducing students to the basics of music notation and music theory, and a wide range of technologies is used for this purpose in existing software.

2.1 Existing Representations

Most symbolic music representations are pre-XML alphanumeric codes like Humdrum/Kern** (see [8, 7]), Plaine and Easie (see [6], abc[1], MusixTeX, LilyPond, etc[2]. Their parsing and graphical rendering are demanding tasks, as in addition to the inherent complexity of music notation the syntax and semantics of most formats are not defined formally and require the development of a specialised parser and suitable data structures. In older programs, the music is representation is handled as character strings until the rendering stage. Therefore music manipulation is realised as string manipulation, which makes programming complex and error prone compared to an data structures modelling musical content directly. On the other hand, using specialised object structures entails the need for specialised tools and rendering components, which is a large effort for educational software projects which are usually underfunded.

A popular code in academia is the Humdrum/Kern representation developed by David Huron, which is supported by the Humdrum Toolkit, a set of UNIX command line programs. This approach is useful for research as it allows rapid software prototyping by using shell scripts. On the other hand it limits the feasibility of programs for the end

[1] See `http://www.gre.ac.uk/~c.walshaw/abc/`.

[2] See `http://www.music-notation.info/` for a comprehensive list.

user as most end users use neither UNIX nor the command line. Another well known format is the Plaine and Easie code which is used by RISM archive for storing musical icipits. It is also used by the CAMI Talk authoring system for the development of interactive music education programs (see [2]).

2.2 File Formats

Apart form text based formats there are proprietary binary formats used by commercial software, most notably the Enigma format used by the *Finale* notation software, the specification of which has been published[3], and the binary formats used by sequencer programs like *Cubase*[4] and *Logic*[5].

In recent years, the introduction of XML has fostered the development of several XML formats for music notation. MusicXML is an exchange format for western music notation[5]. The WEDELMUSIC system has been designed as a format for music distribution over the web, including DRM and it will serve the basis of MPEG SMR (see [1]). MUSITECH comprises object model, XML format, and software modules and emphasises integrated representation of different levels for education and research (see [4]). CapXML is the XML format used by *Capella*[6].

XML has the advantage that parsers and other software tools are readily available in most programming languages for all major platforms. XML, especially XML-Schema, allows also the expression of structures representing an object model almost one-to-one and can describe constraints on values and structure e.g. range checks. The main existing standardised representations are MIDI and audio formats. For symbolic information, mainly standard western notation, there is a variety of different codes with different properties and different tools available.

2.3 Extensibility, Integration and Synchronisation

The integration of different levels of information is important for music education, e.g. the connection between symbolic and performance events or timestamps pointing to an audio track, that can be annotated and structured. As music teachers invent new teaching and training methods, they need additional information to be stored, which makes extensibility indispensable. For demonstrations and interactivity synchronisation is essential, e.g. when showing notation or annotations together with audio or MIDI/SASL.

[3] See `http://www.lilypond.org/web/devel/misc/etfformat`.

[4] `http://www.steinberg.net`.

[5] `http://www.apple.com/logicpro/`.

[6] See `http://www.whc.de/capella.cfm`.

2.4 SMR Processing

The efficient development of effective music education software requires tools for processing music on a symbolic and structural level, as symbols and structures are central parts of the music curriculum. Standard functions required are transposition, change of key and time signature, extraction of time segments or voices, etc. In addition, music exercises and lessons require special functions depending on the pedagogical intention, e.g. generation variations within certain constraints for multiple choice tasks. A different set of functions is necessary for the analysis of user input. Here it is necessary to have input modules, e.g. defining notes with the mouse, and to develop algorithms specific for the pedagogical intentions, e.g. checking for enharmonic errors.

3 MPEG SMR Features for Music Education Software

The MPEG SMR offers a chance to fulfill the needs of music education through a number of planned technical features and by the effects of standardisation itself.

Standardised Format and Tools The work on the standardisation of an SMR format within MPEG 4 offers the chance to reach a significant level of interoperability. This is because the recognised organisation ISO/MPEG and the well established standards MPEG 1, 2, 4, and 7 support acceptance in the markets. The adoption of one public format by a large number of hard- and software developers can vitalise a market very much, as the success of MP3 has shown.

An established standard can lead to a community and market for software tools and modules, that allows music software developers to concentrate on their core competences. If standard function like notation and playback are readily available, it will allow developers to concentrate on the musical, pedagogical, aesthetic, and ergonomic aspects of design.

XML and Binary Format The MPEG SMR will be available in XML and binary representation, which can both help the success of music education software. XML is important as it facilitates the development of parsing, storage, and retrieval, greatly reducing efforts compared to non-XML codes. On the other hand, the option to use a binary format supports applications on devices with limited resources in terms of processing power and memory and limited wireless connections. By overcoming technical limitations and MPEG standard are accepted for all types of devices, MPEG SMR will enable music education software to reach a wider audience.

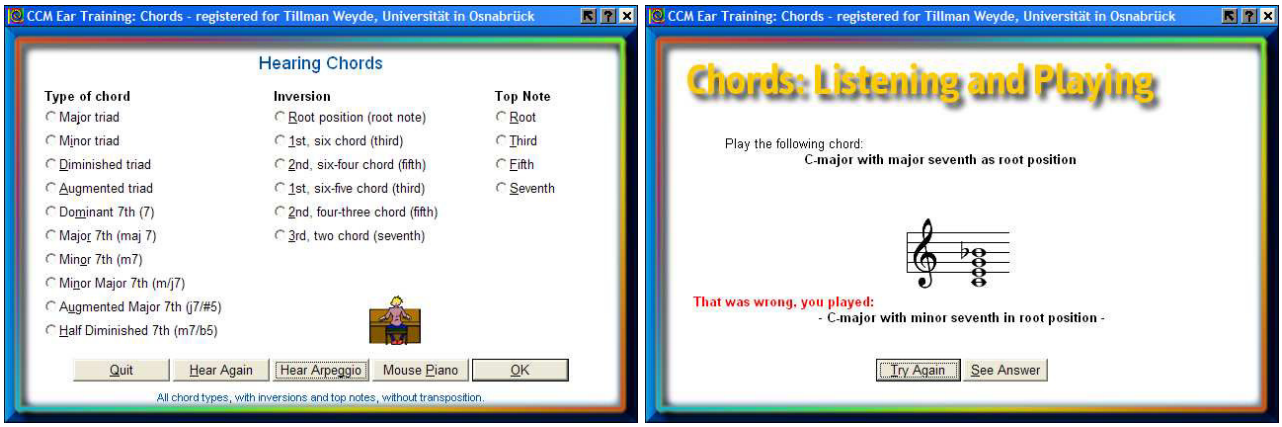


Figure 1. Chord recognition and construction exercises in CCM Ear Training.

Musical Structure The representation of musical structure is essential for MES as it is one of the main topics in music education. The representation of all structurally relevant information like metre, tempo, voice structure, harmonic information, and the flexible selection model allowing to describe arbitrary sets of events are invaluable for designing music exercises.

Annotations and Hyperlinks In addition to the representation of structures, the annotation of events and structures is important for educational applications. Annotation enable the definition of exercise material based on pieces from the literature. Annotation can also be used for the communication between teacher and student, or for marking student works. Tutorials and lectures can use annotated musical pieces. Hyperlinks in the musical material, allow content authors to offer additional help or background information after the completion of a task, thus enriching the application.

Non-Standard Notation The MPEG SMR shall support to some extent notations other than western standard notation and its variants. This gives the opportunity to introduce different notations to the student without the need to develop new technologies.

4 Perspectives

The effect that the MPEG SMR can have on computer based music education shall be discussed here in the context of completed and planned projects.

4.1 Examples of MPEG SMR Potential

The *Computer Courses in Music - Ear Training* (CCM)[3] is a software for music students to improve their aural skills. It has been developed using the authoring system CAMI-Talk which uses Extended Plaine and Easie Code (EPEC). The following typical examples from the CCM Ear Training shall illustrate how MES development could benefit from MPEG SMR.

In the CCM- Ear Training we developed a comprehensive chord recognition exercise for three- and four-part chords (see figure 1). The cord examples types are generated in strings using EPEC. Generating even these comparatively simple structures requires extensive string manipulation, either by of parsing and regenerating EPEC strings or by devising string based functions for specific musical tasks. We used a mixture of both methods, but either method takes considerable effort to develop, test, and maintain.

The situation is quite similar in a related exercise, where the user is asked to play a chord given its name (also figure 1). Here the analysis of the input is very important, to not only be able to recognise the expected chord, but also detect the type of chord that has actually been played. Here we have the problem of pattern matching on the interval structure which is not readily available in the string representation.

Using an object model based on the MPEG SMR format facilitates the implementation of musical logic on the objects, therefore avoiding the re-invention of the wheel. The use of a standardised format allows the reuse of modules by different developers and companies.

Another example is a melody dictation exercise as shown in figure 2. The analysis of the user input needs to be fault tolerant and therefore must recognise some structural properties and similarities between input and presented model. This is especially of interest in respect to the temporal structure. Although the EPEC interpreter we developed also generates timing information for MIDI, there is no straightforward way to link the time stamps to the symbolic informa-

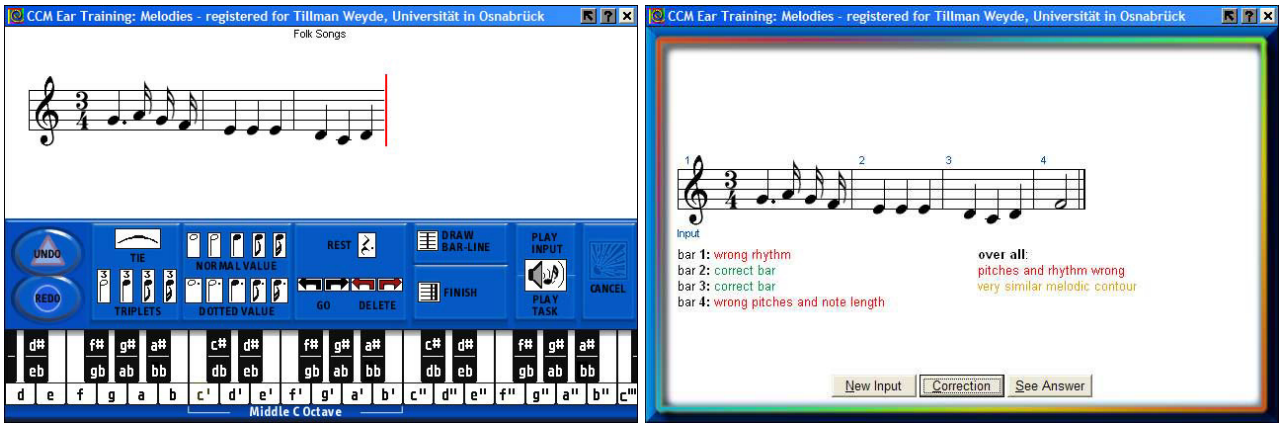


Figure 2. Melody dictation exercise, input and evaluation in CCM Ear Training.

tion. Here the integration of MPEG SMR with MIDI/SASL will improve the situation.

4.2 High Level Abstractions

The use of a standardised format and tools facilitates the development of tools that allow a high level description of musical exercises.

The selection model in MPEG SMR allows for instance to define a series of exercises that use excerpts from a given piece. Details about the presentation and interaction can be saved as annotations to selections. The definition of such pedagogical information can be done manually but also with the support of tools that allow automatic or semi-automatic processing for task like

- excerpts from existing music
- generating variations
- generating new music

In these tasks, the satisfaction of pedagogical constraints is of importance, e.g.:

- appropriate levels of difficulty
- find/vary/generate music according to criteria like
  - harmonic context
  - melody style: chromatic, diatonic, pentatonic, etc.
  - melodic contour
  - rhythmic patterns

Although some of these tasks in general represent unsolved problems of musical AI research, partial or approximate solutions can often be found for specific problems. This is an especially powerful approach in combination with formalisations of pedagogical paradigms and techniques like

- multiple choice
- cloze
- selection/indication
- text input
- construction/composition (visual, keyboard, MIDI instrument)
- performance (real time).

The combination of such tools with the MPEG SMR will be developed in the European project I-MAESTRO to build a modular system for the efficient development and authoring of effective music education software.

5 Conclusions

The development of MPEG SMR offers a chance for music education software developers to benefit from standardisations. The technical qualities of MPEG SMR, as it is planned now, will fulfill many of the requirements for rich interactive and intelligent music education software. This can lead to the development of tools that are useful and available to all developers using MPEG SMR. As experience from previous projects shows, a better representation and tools can allow developers of music education software to concentrate on open problems in their core domain such musical intelligence, human interface design, and pedagogical concepts.

References

[1] P. Bellini and P. Nesi. Wedelmusic format: An xml music notation format for emerging applications. In *Proceedings of the Wedelmusic Conference 2001*, Firenze, Italy, 2001.

[2] B. Enders. Besonderheiten der erstellung von musiktutorials und multimedia-applikationen mit der autorensprache camitalk. In B. Enders and N. Knolle, editors, *KlangArt-Kongreß 1995*, volume 1 of *Musik und Neue Technologie*, pages 227–248. Universitätsverlag Rasch, Osnabrück, 1998.

[3] B. Enders and T. Weyde. *Computer Courses in Music - Ear Training*. Schott Musik International, Mainz, London, New York, 2001.

[4] M. Gieseking and T. Weyde. Concepts of the musitech infrastructure for internet-based interactive musical applications. In P. N. u. M. S. Christoph Busch, Michael Arnold, editor, *Proceedings of the Second International Conference on Web Delivering of Music WEDELMUSIC 2002*, pages 30–37. IEEE, Los Alamitos, California, 2002.

[5] M. Good. Musicxml in practice: Issues in translation and analysis. In *Proceedings First International Conference MAX 2002: Musical Application Using XML*, pages 47–54, Milan, 2002.

[6] J. Howard. Plaine and easy code: A code for music bibliography. In E. Selfridge-Field, editor, *Beyond MIDI – The Handbook of Musical Codes*, pages 362–372. The MIT Press, Massachusetts, 1997.

[7] D. Huron. Humdrum user guide. Online manuscript, Ohio State University, 1999.

[8] J. Wild. Review of the humdrum toolkit. *Music Theory Online*, 2(7), 1996.

The Symbolic Music Representation of Traditional Korean Music

Yong-Soo CHOI, Hyoung-Joong KIM
Kangwon National University, Korea
{ciechoi, khj}@kangwon.ac.kr
and
Yongju Cho, Jae Gon Kim
ETRI, Korea
{yongjucho, jgkim }@etri.re.kr

Abstract

An SMR (Symbolic Music Representation) for Korean traditional music is proposed in this paper. These proposed notations have been recently used for score by several schools of Korean traditional music. Among them a few ornamental notations are quite different from western notations. Thus, to make the western music notation can represent the Korean music notation some ornaments have been contributed to MPEG SMR standard. This paper introduces the standardized Korean ornaments.

1. Introduction

According to the Call for Proposal on the Symbolic Music Representation from the international standardization body (ISO/IEC JTC1/SC29/WG11 N6689) [1, 4] a methodology for symbolic music representation of traditional Korean music has been proposed for interoperable integration of it into MPEG music notation [3]. Traditional Korean music (a.k.a., guk-ak) has followed the traditional musical notation, chong-gan-bo, since the fifteenth century. Both the staff notation and the chong-gan-bo are used in Korea now. Of course, staff notation is far more popular than the latter. Only the part of traditional music is scored by the latter. Even though those two notations are totally different, chong-gan-bo can be written in existing markup languages for western music notations because eventually music is music. However, some typical ornaments of Korean music do not exist in western music. Thus, it is important to include Korean ornaments into western music notations to make these notations rich enough to cover western music as well Korean music.

Korean ornaments were notated differently by musician by musician such as Kim Ki-su, Lee Ju-hwan, Hong Won-gi, Kim Jung-ja, Choi Su-ok, and so on. Their notations and interpretations of the same ornamental performance are different. Thus, it is not easy to standardize the Korean notations. However, for the sake of international standardization, some of typical notations have been contributed to the MPEG [3]. This paper introduces the standardized notations.

2. Korean ornamental notations

Chong-gan-bo is a systematic musical notation first invented in East Asia, which is able to indicate both pitch and rhythm conveniently. “Chong-gan” means a square, and “bo” means the score. The chong-gan-bo uses basically twelve Chinese characters to indicate the twelve pitches in the octave. A pitch is represented by one Chinese character in the square [2]. An empty square indicates sustaining the previous note. One square stands for one beat. Number of squares indicates the number of beats. For example, one square that contains one character in the square with accompanying an empty square stands for two beats. Number of small squares in one chong-gan indicates the number of notes for one beat.

Further marks for ornaments are added next to the notes. Some ornaments in Korean music cannot be expressed by western music notation because they are native to traditional Korean music and subjective to the native instruments, performers, and scales. A few among many unique ornaments for expressions are listed in Table 1. As was mentioned before, the notations are different from musicians. These notations are a combination of different notations.

Based on the ornaments listed in Table 1 and XML notation, chong-gan-bo can be expressed by the common western music notation. Thus, the proposed XML notations can enrich the existing music notation. The chong-gan-bo itself is a unique and extremely systematic musical notation. Therefore, chong-gan-bo needs a unique music representation method. In this document, only 8 music notations listed above that can be marked on the staff notation are introduced. The presented notations are commonly used in playing the music (not in special instrument or vocal features). Nonghyun means vibrating the string in a strict sense. These nonghyuns

will make the music more musically sensitive by rapid fluctuation in the sound level. The important thing is to keep one beat even if how many features are included in the one chong-gan.

The first nonghyun ("increase") vibrates the sound decreasingly with changing the sound from higher to lower, and the second nonghyun ("decrease") vibrates the sound increasingly with changing it from lower to higher. Figure 1 shows the schema of proposed notations. It consists of Korean Music SMR, Nonghyun and Sliding Element.

Table 1. Ornaments for Korean traditional music

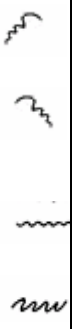
Mark	Explanation	XML notation
	Sliding tone to higher (can visually show how long and deep)	<sliding>up </sliding>
	Sliding tone to lower (can visually show how long and deep)	<sliding>down </sliding>
	Two consecutive sliding tones to higher	<sliding>double-up</sliding>
	Sliding tone to lower and to higher	<sliding>down-up</sliding>
	Nonghyun (vibrato) (can visually show increase, decrease, narrow or wide)	<nonghyun> increase </ nonghyun > <nonghyun> decrease </ nonghyun > <nonghyun> narrow </ nonghyun > <nonghyun>wide </ nonghyun >

Figure 1. XML schema for proposed notations.

Figure 2. XML schema for sliding element.

Above Figure 2 represents all value of sliding element.

Figure 3. XML Schema for nonghyun Element.

Above Figure 3 represents all value of sliding element. These three xml elements can be embedded to express the music score with proposed music notations.

```
<pitch>
  <step>F</step>
  <sliding>up</sliding>
  </pitch>
```

Figure 4. A sample of XML representation

3. Example using XML

An example will be more powerful than hundreds of words. Thus, in this section an example using XML as the language of representation is porvided. The Chinese character in the box in Figure 4 stands for "F" or "Fa" in western music. The symbol below the Chinese character is the sliding up in the Table 1. This chong-

gan-bo notation can be expressed by XML as is in Figure 4.

```
<?xml version="1.0" standalone="no" ?>
<!DOCTYPE score-partwise>
- <score-partwise>
<movement-title>Sample</movement-title>
- <part-list>
- <score-part id="P1">
<part-name>Voice</part-name>
</score-part>
</part-list>
- <part id="P1">
- <measure number="1">
- <attributes>
<divisions>2</divisions>
- <clef>
<sign>G</sign>
<line>2</line>
</clef>
</attributes>
- <note>
- <pitch>
<step>G</step>
<sliding>down-up</sliding>
<octave>4</octave>
</pitch>
<duration>2</duration>
<voice>1</voice>
<type>quarter</type>
<stem>up</stem>
- <notations>
<slur type="start" number="1" />
</notations>
</note>
- <note>
- <pitch>
<step>F</step>
<sliding>down</sliding>
<octave>4</octave>
</pitch>
<duration>2</duration>
<voice>1</voice>
<type>quarter</type>
<stem>up</stem>
- <notations>
<slur type="stop" number="1" />
</notations>
</note>
- <note>
- <pitch>
<step>D</step>
<nonghyun>increase</nonghyun>
<octave>4</octave>
</pitch>
<duration>2</duration>
<voice>1</voice>
<type>quarter</type>
<stem>up</stem>
- <notations>
<slur type="start" number="1" />
</notations>
</note>
- <barline location="right">
<bar-style>light-light</bar-style>
</barline>
</measure>
</part>
</score-partwise>
```

Figure 5. A sample of XML Code

Thus, extension of existing music notations are needed to incorporate the traditional Korean music notations. An example of musical notations including "sliding" and "nonghyun" is given in Figure 5. (See the highlighted lines).

4. Experiments for Notations

This section shows referred waveforms of proposed music notations. From Figures 6 to 9 show four examples of notations in waveforms of each ornament. Note that these waveforms show that they are different from vibrations of western music in many aspects.

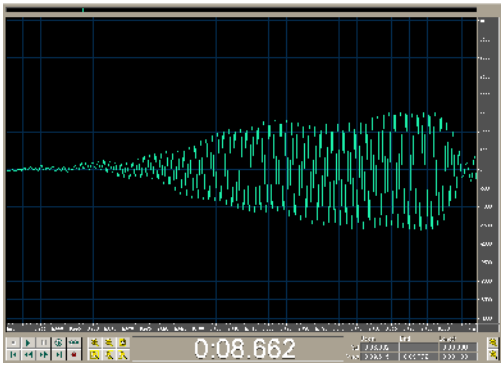


Figure 6. Waveform of "Sliding Up"

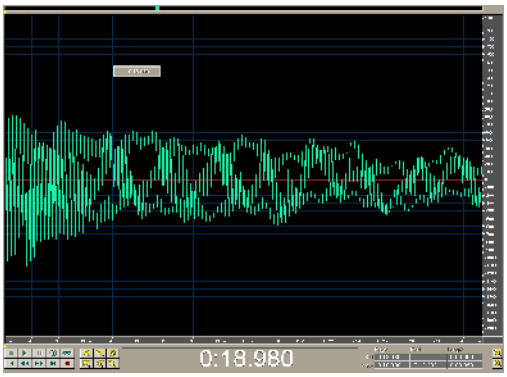


Figure 7. Waveform of "Sliding Down"

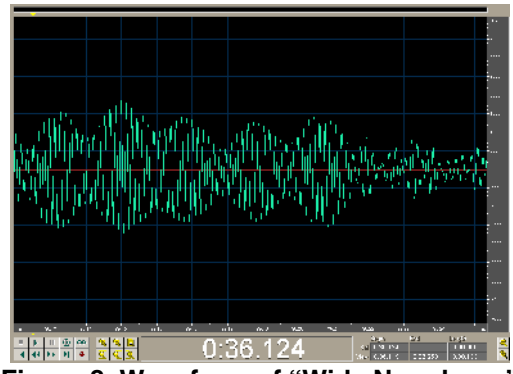


Figure 8. Waveform of "Wide Nonghyun"

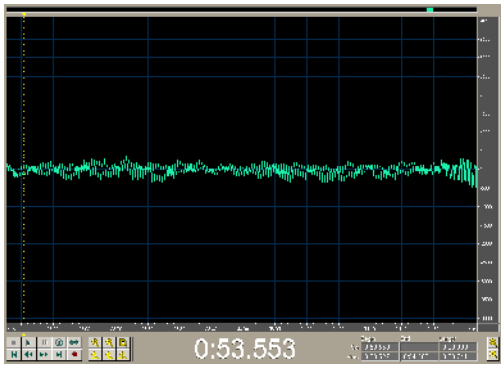


Figure 9. Waveform of "Narrow Nonghyun"

These waveforms are extracted from the real music example (The Sang-Ryung-San played by Yu Cho-Shin). Needless to say, other waveforms can be extracted for associated features suggested in this document. If the proposed notations are applied to existing music score, they can be noted like Figure. 10.

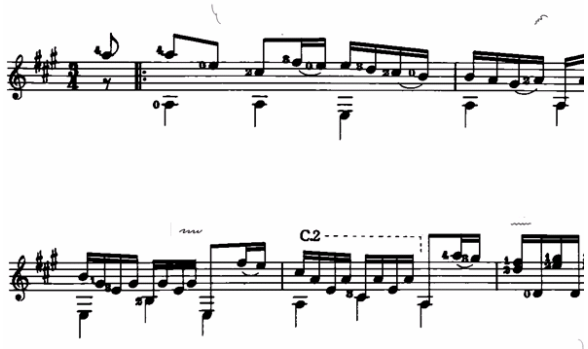


Figure 10. A graphical example of music score containing proposed notations.

5. Conclusion

Chong-gan-bo, a traditional Korean musical notation is introduced. This type of ornaments has a special feature that can only be observed from far-east Asia (typically in Korea, China and Japan). So these types of notations will satisfy the functionality when a music software plays the Asian music. There are more ornamental notations which were not standardized. For the interoperability of the standard symbolic music representation, eight ornamental notations are provided in this paper. This suggests that extensions of existing music notations are further needed to incorporate all the traditional Korean music notations. The XML schema proposed can be embedded into established standard XML representation.

6. References

[1] Audio subgroup, "Draft Call for Proposals on Symbolic Music Representation" ISO/IEC JTC 1/SC 29/WG11, Germany, March 2004.

[2] J.-I. Heo, The Korean Transverse Flute *Taegum* and Its Music *Taegum Sanjo*, Ph.D. Dissertation, Florida State University, 2002.

[3]. H. J. Kim, Y. S. Choi, and Y. Cho, "On the Symbolic Music Representation of Traditional Korean Music," ISO/IEC JTC 1/SC 29/WG11 M11630, February 2005.

[4] P. Nesi, G. Zoia, P. Bellini, and J. Barthelemy, "Music Notation Application Requirements and MPEG Technology," ISO/IEC JTC 1/SC 29/WG11, Brisbane, Oct. 2003.

Integrating music representations within MPEG environments

D Crombie[1], R Lenoir[1], and N McKenzie[1]

[1] *FNB*
Molenpad 2, 1016GM, Amsterdam, Netherlands
projects@fnb.nl

Abstract

By building on recognised Design For All methodologies, systems should be constructed in such a way that the mainstream solution should be easily adaptable and extensible to add functionality for niche markets. Accessibility can be considered to be an interplay of various processing domains. These processing domains relate to processes in production and perception of musical material. Through clear separation of the technological descriptions of music storage and music manipulations procedures, the ideal 'mix' of the interplay between the processing domains – logical, gestural, visual and analytical - can be described. The process of contriving a procedure to interface the various processing levels should be based on use. The representation of the interplay between the various user groups should always remain accessible. If all relevant entities in a representation system remain accessible, creating meaningful mappings is a matter of connecting the appropriate entities.

1. Introduction

This paper outlines the FNB (Dutch Federation of Libraries for the Blind) approach to the incorporation of accessible music within MPEG environments. Much of this work is currently underway under the aegis of the *ad hoc* Symbolic Music Representation MPEG Group. FNB have co-ordinated a number of EU funded music projects (eg CANTATE, HARMONICA, MIRACLE [1]) and their experience in this field comes mainly from previous design of decoders for alternative music representations, rather than from the MPEG framework design perspective (although FNB is a member of MPEG ISO/IEC JTC1/SC29 WG11).

This experience has strongly indicated that XML based approaches to music notation representation are of limited value in this field. XML is not a data representation model: it is an interchange format. As an interchange format it is good at what it does, but naturally not perfect. The usefulness and applicability of the XML formats and processing models involved in using these formats rely on the quality of the analysis of the representation requirements. Description, and with that representation of these properties, involves programmatic description of requirements on a relatively high abstraction level. The sheer number of XML dialects for common western music notation illustrates the problem. The only accessibility related issues in this respect are that when creating an accessible system it is important to synchronise the most complete set of information as possible relating to the content in order to ensure that all the information is provided in alternative formats. Similarly any standardised interchange format must be open source and should also be easily expandable by those working in the field of accessibility.

Integration of accessibility notions into the MPEG family, however, will provide previously unavailable opportunities in the provision of accessible multimedia information systems [5]. It will open up modern information services and provide them to all types and levels of users both in the software domain and the hardware domain. In particular, the work being undertaken by MPEG will provide access to multimedia content to print impaired users. Additionally, new consumption and production devices and environments can be addressed from

this platform and will provide very useful information provision opportunities indeed, such as information on mobile devices with additional speech assistance.

As part of the work of the *ad hoc* Symbolic Music Representation MPEG Group, a framework is being developed which aims to provide a logical and structured description of music. As this framework is being fully defined, a decoder can be constructed. Such a decoder would be able to interface with this musical framework; interpret the various aspects of this description; and transform this representation into alternative formats such as Braille Music.

2. Communicating accessibility

We have previously observed that the relationship between technology and music has always been somewhat problematic, given the tensions inherent in any interface between creativity and technology. From our point of view, most of these problems stem from poor communication, a situation that arises because technologists and composers often seem to use different languages while essentially working towards common goals [2].

This paper also addresses the modelling of user-centered interaction paradigms at a fundamental level. Interfacing can be described as defining and specifying 'connection' points for communication. To this end, a high level of flexibility and accessibility can be achieved by separating the various entities that are of importance in the communication process.

2.1 Accessible music representations

Ever since Louis Braille invented his system for representing music, blind musicians have been able to obtain scores in the Braille music format. There are international guidelines for Braille music notation [3]. Materials in Braille music make up the largest portion of the available alternative formats, and include the standard repertoire for most instruments, vocal and choral music, some popular music, librettos, textbooks, instructional method books, and music periodicals. However, Braille music is produced by a relatively small number of institutions throughout the world.

The international guidelines define a rule set which allows all the musical information to be parsed into its relevant representation in Braille cells. This becomes several Braille cells for more complicated aspects of the music. This results in a much larger bulk of information in Braille than in traditional notation. The following simple example illustrates some of these points:

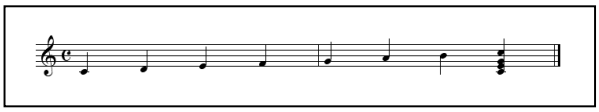


Figure 1: Simple C Scale - Tradition Notation

For a simple two bar melody the Braille Music takes almost a full line. The information is also serial, so the user has to read through the Braille regardless of whether that piece of information is of immediate importance.

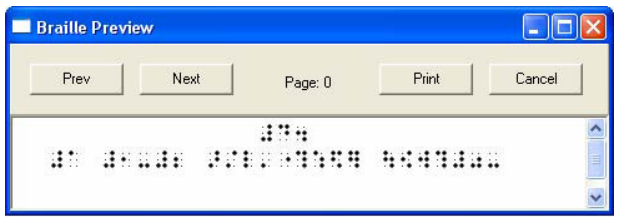


Fig 2. Simple C Scale - Braille Music Notation

For more complex pieces of music, such as the first bar of Chopin's *Revolutionary,* the Braille music requires six lines.

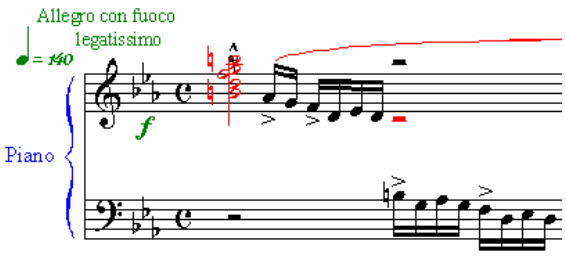


Figure 3. Chopin - Opus 10, 12th etude

Figure 4 shows the Braille output below:

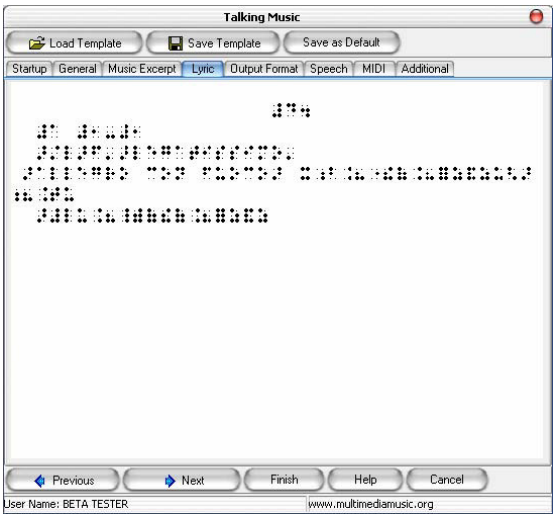


Fig 4. Chopin - Opus 10, 12th etude - Braille Music Notation

The quantity of material accessible in Braille represents only a small percentage of that available to the sighted musician. Sighted musicians can "sight read" an unfamiliar piece while blind musicians (with the possible exception of singers) must spend time memorizing the music which can be difficult to source. Currently, transcription turn-around times can be measured in weeks, months and even years, and this puts the blind musician at some considerable disadvantage [4].

As outlined above, the traditional approach taken to the provision of music for the visually impaired has been largely concerned with Braille Music, and this remains the most widely used method of allowing visually impaired musicians to read scores. Reading Braille Music, however, is a complex task and for most people who become visually impaired beyond the age of twenty it is unlikely they can learn Braille. Alternative formats are therefore needed, and additional approaches include Talking Music and Large Print music [5]. For Talking Music to relay the same level of information to the visually impaired user as the sighted user, everything on the page of a music score must be represented in a spoken format, applicable to all types of music and instruments. Talking Music [6] is presented as a played fragment of music, a few bars long, with a spoken description of the example. The conversion of the musical content to spoken music formats has proved popular with end users, and new technology has been developed to automate production processes. Large (or Modified) Print music is useful for partially sighted end users and could use SVG as a standard where Talking Music uses DAISY [7].

2.2 Accessing music through interactive communication

For most people music is a medium of communication. From the point of view of the composer, music is a means to communicate a thought or idea to wider audience. From the performer's point of view, the music is a means of adding their perspective and interpretation to the thoughts of the composer, and to communicate with a wider audience. From the listener's point of view, the music is a means to escape into your own interpretation of the composer's and or performer's thoughts, while adding your own interpretation of what the music means.

This interactivity is what makes the music enjoyable. The meaning and thoughts which our perspective holds as being contained within a recording or performance are defined when we interact with them. In this way a recording of music can become dynamic, as much of the meaning is defined by the state of the observer. So a static recording becomes a different performance every time it is played if the music or art inspires sufficiently creative thought or interaction within us. The aim of a composer or artist is then to inspire these multiple reactions by encouraging interactivity. In this way notions of 'accessibility' can be further widened.

2.3 Service based on interaction

Bearing these notions in mind, we need to examine the requirement for generic accessibility and that for service based on interaction. A service provides a non-destructive transformation of the information that flows from an original (inaccessible) source of content. A service is an abstraction that collects procedural logic and declarative parametric knowledge and allows assignment of other (external) information resources to this service. This assignment enables controlling of the servicing process by these external information resources.

One example is an external resource that collects user preferences regarding information presentation. It is important to note that the actual location of a software service is not limited to one specific device. Using modern network protocols - especially wireless protocols - distributed service based content provision can be achieved. An example of this design paradigm can be found in Middleware technology, such as Enterprise Java Beans, CORBA and .NET.

A high level of service can be achieved by connecting various software processing services as described above into a network of services. The original content can be transformed *on demand*. Also a high level of scalability is achieved, since the network can be expanded with more and more specific services on the fly.

3.0 Accessible design methodologies

By building on recognised Design For All methodologies, systems *should* be built in such a way that the mainstream solution should be easily adaptable and extensible to add functionality for *niche* markets. As a result of the comprehensive lack of understanding of this concept at the fundamental design level, and strict deadlines to complete software projects, most accessible solutions become piggy-backed onto an ill-suited system as an afterthought. The accessible solution is then itself ill-conceived and unlikely to meet the needs of the end user. This often raises the question (though rarely explicitly) of whether the specialised needs of the niche market merit the effort involved in providing a solution.

The advantage of using objects for accessible design lies primarily in their re-usability and adaptability. If accessibility is considered at the birth of the enterprise, and the concepts of modern software design are utilised to their full extent, the accessible solution is ultimately better designed and other niche markets can be more easily addressed. This rudimentary concept is rarely considered.

Accessibility can be considered to be an interplay of various processing domains. These processing domains relate to processes in production and perception of musical material. Through clear separation of the technological descriptions of music storage and music manipulations procedures, the ideal ‘mix’ of the interplay between the processing domains – logical, gestural, visual and analytical - can be described. Where additional specialised requirements are needed, an linking mechanism of external formats and models is provided. This not only ensures future expandability of the framework and representation capabilities. It also ensures expandability to other musical styles, that require different music notation schemes. Because of the logical domain description that is not effected by these transformations and because of the non-destructive nature of the accessibility transformations, availability of music for groups of end-users is ensured.

One note of caution should be sounded here. In the MPEG(7) arena an explicit distinction is made in the direction of engaging an information flow for consumption. One direction is the *pull* strategy and the other is the *push* strategy. The pull strategy is defined by the initiative for the information exchange by the end-user. It is the end-user who decides what to see, hear or interact with. In the push scenario, information is pushed or proposed in the direction of the end-user in a proactive manner, preferably because of preferences an end-user has defined regarding this kind of application behaviour. The selection of a *pull* scenario relies heavily on the assumption that an end-user knows what he/she wants to retrieve and is able to navigate though the presented content freely and without barriers. This is the domain in which print and vision impaired end-users are set back. Most of the content provided is *not* freely explorable by this category of end-users.

In order to provide accessible music in MPEG environments, a first step would be the creation of a Braille music decoder so that the following could be made possible:

- Greater standardisation through use of the following elements of Braille music production: Braille music printing preferences; Braille music output based on

the International Braille Music Manual; Braille representation preferences at a document level
- Interpretation of the MPEG framework to Braille Music
- Protocol for communicating the musical information to the relevant output device (Braille embosser etc.)

A Braille music decoder could be built around a Braille interpretation component which would be native to several Braille music modules. An MPEG to Braille music decoder would make use of this component and return the required result based on the user's preferences. The various output media required by users of such a system would each have a similar module specific to the output requirements for that medium.

This could take place if all modification of the input components takes place with output specific settings for classes and objects. This allows specific separation to take place between the generic MPEG framework and the various decoders required to meet the needs of users of alternative music notation formats. This can be seen as extensibility, where the modifications are specific to the application logic. This extensibility becomes important to encourage re-use between the components of such a system. The Document settings module may require re-use of the interpretation module's classes and objects.

4.0 Building within MPEG environments

In a sense, the MPEG initiatives exist as families. All the family members depend upon one another and have evolutionary relationships: hence they have an evolutionary base for future development. The various family members operate at different abstraction levels with some communication between these abstraction levels.

The process of contriving a procedure to interface the various processing levels should be based on *use*. The difficulty lies in achieving a level of description of the user requirements that allows re-description in technological terms. This re-description ideally leads to specifications and ultimately implementations. These implementations 'prove' the viability of the concept: it is the proof of the hypothesis. The process of standardisation that runs in parallel with this ensures extraction of higher level descriptions and these are aggregated down to the earlier family members. Using this built-in feature to provide 'slots' for common and specialised accessibility requirements would create what we refer to as *accessibility from scratch* [8]. Once embedded in the family tradition of the MPEG initiative, accessibility might become a commonly available feature instead of a workaround necessity. It is anticipated that the forthcoming European accessible information processing initiative EUAIN [9] will consider this approach.

Building within MPEG environments requires some travelling through acronyms. The user model interacts with and uses DID (Digital Item Declaration) and DII (Digital Item Identifier). DII can be considered to be pointers to locations and contexts in the media structure. DID describes what the DII is pointing to, so in programming terms DII is a pointer and DID is a memory location, albeit a memory location with its own intrinsic structure. DIP (Digital Item Processing) can be used to describe and specify the dynamics required to meaningfully access a DID. DIP 'points' towards the most appropriate consumption behaviour. Ideally this 'pointer' is based on a robust user representation model and enables incorporation of consumption pointers for common and specialised accessibility requirements. DIA (Digital Item Adaptation) is the foreseen infrastructure to achieve this. DRM (Digital Rights Management) is the 'faucet' that ultimately decides (based on the description of the permissions using REL (Rights Expression Language)) if a user is permitted to gain access to the DID.

The earlier MPEG family members focus on the continuous domain, one example being MP4 streams. Higher level family members, such as MPEG21, are symbol based: MPEG21 is a framework of interacting objects. Providing mapping mechanisms that associate continuous behaviour with appropriate discrete objects of meaning is what every human being has to learn during their life. The difficulty here lies in the individual nature of this task. The exact strategy applied to achieve *useful* mapping between the continuous and the discrete domain depends on the

requirements and expectations of the *end-user*. Additionally it should be noted that potentially *any* human being is an end-user. This includes editors, administrators, publishers and gamers. Since MPEG21 and similar initiatives are the work of human beings, descriptions and implementations of such a framework will encounter this same fundamental challenge.

The representation of the interplay between the various user groups should always remain *accessible*. If all relevant entities in a representation system remain accessible, creating meaningful mappings is a matter of connecting the appropriate entities. For this reason, accessibility *from scratch* is of fundamental importance.

6 References

[1] http://projects.fnb.nl

[2] Crombie, D., Lenoir, R., and McKenzie, N., (2003) Bridging the Gap: Music and Communication, Proceedings Rencontres Internationales des Technologies pour la Musique (Resonances 2003), IRCAM, Paris

[3] Krolick, B (1996), "New International Manual of Braille Music Notation", World Blind Union, FNB, Amsterdam

[4] Crombie, D, Dijkstra, S, Lindsay, N and Schut, E (2002), "Spoken Music: enhancing access to music for the print disabled", Lecture Notes in Computer Science, Vol 2398. Springer-Verlag, Berlin Heidelberg New York

[5] Crombie, D, Lenoir, R and McKenzie, N., (2004), "Making Music Accessible: An Introduction to the Special Thematic Session", Lecture Notes in Computer Science, Vol (*forthcoming)*. Springer-Verlag, Berlin Heidelberg New York

[6] Crombie, D., Lenoir, R,, McKenzie, N., (2003) "Producing Accessible Multimedia Music" in Proceedings 3rd International Conference on Web Delivering of Music, IEEE

[7] http://www.daisy.org

[8] Crombie, D., Lenoir, R., and McKenzie, N., (2004) "Accessibility from scratch: How an openfocus contributes to inclusive design", Lecture Notes in Computer Science, Vol (*forthcoming)*. Springer-Verlag, Berlin Heidelberg New York

[9] http://www.euain.org

[10] Crombie, D., Lenoir, R,, McKenzie, N, (2003) "Eye of the Beholder: Re-defining Accessibility" in Proceedings Electronic Arts & Visual Imaging (EVA2003), Institute of Archaeology, London, UK

[11] Van Ossenbruggen, J., Eliëns, A., Rutledge, L., (1998)"The Role of XML in Open Hypermedia Systems" 4th Workshop on Open Hypermedia Systems, Hypertext

[12] Stephanidis, C. (2001) 'User Interfaces for All: New perspectives into Human-Computer Interaction' in C. Stephanidis (Ed), User Interfaces for All – Concepts, Methods, and Tools (pp 3-17). Mahwah, NJ: Lawrence Erlbaum Associates

Contribution of A.F.I. - the Italian Association of Phonographic Producers - to the "Study on Community initiative on the cross-border collective management of copyright" of the 7th July 2005

Isabella Longo europeanbranch@afi.mi.it
Franco Bixio francobixio@afi.mi.it

Abstract

AFI, the Italian Association of Phonographic Producers (www.afi.mi.it), has a clear interest in the terms and contents of the Commission Staff working document of 7th July 2005 as is entitled to act as a ***collecting society of related rights.***
AFI collects and distributes related rights to over 170 SMEs of phonographic producers and as a member of Confindustria (the National Confederation of Industrial Employers of Italy) it defends and promotes the collective interests of the independent music sector.

1. Introduction

On the overall, with regard to any Commission proposal with respect to the cross-border collective management, AFI maintains that firstly it is essential that the **fundamental principles and rules on related rights** are respected.

From a careful analysis of the Commission's text, the identification of the owner of the related right is still linked to the "record label" parameter which does not correspond to the laws in force and is not suitable to the digital environment.

In addition, it's worth noticing that unlike previous Commission Communication (COM 2004 261) in the text's title the related rights are not mentioned even if the issue of their collective management is taken in exam.

Therefore, a EU legislative initiative on the object will not obtain the expected results unless the above indicated issue is duly clarified.

2. The *inadequacy* of the protection of related rights

Originally the record industry was fully integrated from original production (investment, talent scouting, etc) to the distribution of finished works. Therefore related rights were naturally collected and distributed to the record companies on a record label basis. During the course of time the major record companies have developed the distribution side of their business by getting licences from third parties or from the original phonographic producers of works.

Despite this significant change the collecting societies have continued to distribute the related rights directly to the record labels without taking into consideration the rights of the original producer also in case they have ***not*** been explicitly ceded to the record companies. That was also due to the fact that laws regulating related rights were easily misinterpret able.

Today the **WIPO Performances and Phonograms Treaty (WPPT 1996)**, implemented in the EU Member States by means of the Directive 2001/29 CE, grants related rights to "**the producer of phonogram**", defined as, "*the person, or the legal entity, who or which takes the initiative and has the responsibility for the first fixation of the sounds of a performance or other sounds, or the representations of sounds.*" (art. 2) **It's worth mentioning that in Europe today this role is often held by the thousands of SMEs of the independent producers**.

This new definition of the "Producer of Phonogram" clearly states to whom related rights must be granted and it has a considerable meaning especially as it take into account the way related rights could be exploited in the digital environment.

The great opportunity to exploit a single "track" in the digital market on a worldwide basis added to the loss of importance of the "physical recording " have made it clear that the related rights require more consideration and that owner of the related right is the person that invests in promoting "creativity" rather than the one who " fabricates the original recording" (previous definition ex art 78 Italian copyright Law 633/1941 before the Directive 29/01 implementation).

Even so, the system for the identification of the owner of the related rights still follows "record label" parameters - a system that continues to favour the major record companies. As already underlined even in the Commission's text the "record label" are often mentioned as the owners of the related rights.

It's worth mentioning that in Europe, IFPI collecting societies manage and distribute rights to their record member thus having consequences also with respect to the *Simulcasting license* (below analysed).

This "distortion" derives from the clear interest that major record companies have in keeping this "status quo" unchanged and from the lack of rules that grant a clear identification of the owner of related rights.

Even before the WIPO Treaty, both in national and international laws the reference to the "label record" has not been reported as a parameter to identify the owner of the related rights always granted to the producer of phonogram

As already reported in the contribution submitted on 18th of June 2004 in the context of the consultation launched the 21st April 2004, AFI intends to draw the Commission's attention on the following issues

- The collective management of related rights in the digital context must guarantee the respect of the principle of plurality and must represent all rights owners. To this end it is necessary to adopt rules aiming at clearly identifying the right owner for each single music file. Collecting societies must represent, defend and distribute to the producers of phonogram as defined by the WIPO Treaty and by the law in force. .
- IFPI do not represent all collecting societies. In other countries (such as in Italy and France) there are other collecting societies representing and managing related rights of independent producers

3. IFPI Simulcasting license

The Community approval of the IFPI Simulcasting License is based on the fact that this multi territorial license would enhance the use of digital content and would facilitate the royalties' distribution iter.

Following the above consideration it has to be underlined that IFPI initiative allows only the IFPI collecting societies to gather and distribute royalties to their record labels. Therefore it is obvious that the "customer allocation" is not relevant as any licensing profit can always be controlled, at European level, by the same group of major recording companies and among them distributed.

We believe that any simulcasting license should operate in order to allow all European collecting societies to authorize the use of each single music file and should respect the plurality of the rights owners.

4. Collective management and the wide Community licensing of related rights

Following the above analysis AFI intends to analyse the negative impact that the adoption of option 3 would cause to the rights owners, to the CRMs (not only the medium size ones) and as well to the economy of new member states, having also serious consequences with respect to niche repertoire and to the promotion of new talents

OPTION 3 *Give to right-holders the possibility to authorise a collecting society of their own choice to manage their rights across the entire EU*

We firmly believe that promoting the competition among CRMs making the intellectual property rights (including related rights) objects of negotiation is not acceptable. We are quite puzzled also with respect to the use of the language used in the text relating to the participation to the "royalty cake" as if the subject of licensing intellectual property rights has to be treated only with regards to its economical and commercial power. Option 3 would generate market distortion, concentration of power in few hands favouring only successful and important authors, composers, publishers, producers and artists and damaging a certain category of rights owners, the less "attractive" repertoire and the less powerful CRMs. It is clear that less experienced and wealthy CRMs representing a repertoire not exploitable on a global scale (for music style and for the language) would easily disappear in favour of few and big collecting societies capable to manage few and "big" rights owners.

The CRMs' aim would lead to attract and obtain repertoire exploitable on a global scale. Therefore niche and local repertoire would be still managed by less powerful CRMs not having the same negotiating clout vis-à-vis big and powerful commercial users.

As a consequence this situation will "impoverish" less famous and known right owners and enrich few international "creators". Option 3 could have an effect on the respect of plurality and of cultural diversity.

5. Comments to Option 3 aspects

Hereunder we have analysed option 3 impact with respect to each single aspect examined in the text.

5.1 Legal Certainty (point 4.1.of the text)

The adoption of option 3 will not necessarily grant more legal certainty than now. User will need to obtain licenses from more CRMs and get rights clearance from all CRMs whose members' repertoire is requested

5.2 Transparency/Governance (point 4.2.of the text)

This aspect is linked to the aspect of the relation with right owners below reported.

5.3 Culture/Creativity (point 4.3. of the text)

Option 3 has "*the potential to increase the overall amount of revenues created by copyright licensing in the online environment and thus enlarge the pie"*. This will not lead to support the promotion of all European creativity and of new talents if resources are engaged to support and promote creativity of few important "creators".

5.4 Trade flows (point 4.4 of the text)

Competition among CRMs to obtain direct membership of rights holders having *"attractive repertoire"* again will favour only important rights holders. As written in the text *"above all, fewer more efficient societies will distribute more to their members"*

5.5. Innovation and growth (point 4.5. of the text)

Option 3 "*would also stimulate the roll-out of new online services because it will facilitate management of rights by concentrating the licensing process to a few transactions as opposed to potentially 25 licensing transactions in all Community territories" ….. "the European repertoire will be split among a small number of CRMs"*

In our perspective this means that few powerful CMRs will issue license to few "*powerful*" commercial user interested in licensing globally exploitable repertoire

This " European repertoire" will not cover repertoire representing all cultural and musical traditions and specificities of each European nation but mainly the most *attractive* repertoire which means the one *marketable* on global scale (both for music style and language used).

5. 6. Competition (point 4.6.of the text)

The statement that option 3, "*by giving right-holders the possibility to freely choose and move among CRMs, would create the competitive discipline that forces CRMs to compete among themselves for right-holders and negotiate advantageous royalties on their behalf"* is not acceptable for the already mentioned reasons.

Again we want to underline that the commercial negotiation of intellectual property right is not acceptable.

This will allow each CRM to freely estimate the economical value of an intellectual property work (only with respect to their exploitation and not with respect to their cultural value) with possible discriminatory consequences.

In addition how and who will grant that a right holder of category B (having less famous and exploitable repertoire) would have the same economical consideration of right holder of category A (having famous and exploitable repertoire) ?

5.7. Vertical integration of the media (point 4.7 of the text)

Option 3 would "*also lead to the emergence of limited amount of (three or four) powerful CRMs for online licensing who effectively defend right-holders interest vis-àvis powerful commercial users at a pan-European leve*l".

Again Commission underlines that "t*his gives the collective rights manager – especially the society that has accumulated an attractive repertoire – a strong position in order to increase the royalty flows for its members." …This is because a powerful collective rights manager representing a significant repertoire will be in a strong position to negotiate royalties on behalf of its members and thus ensure that right-holders participate in the increased royalty cake"*.

We have already expressed our concern about this aspect .What will happen to the repertoire having great artistic value created by non international authors, producers and artists ? What will happen to the less big and powerful CRMs of the new Member States whose activity, in spite of their efforts, could not compete with most powerful and experienced EU CRMs?

In this section Commission also states that option 3 "*would foster integration among CRMs*" thanks to the pan - European scope of their role.

CRMs should then offer same economical consideration to their members regardless of their repertoire and nationality. Even if this is granted (who

and how ?) the pan European role of each CRM would not easily solve problems related to all other CRMs activities undertaken at national level.

CRMs' role goes beyond the licensing and rights managing: the fight against piracy, the promotion of the represented works, the monitoring of the correct use of the licensed repertoire and the mandate to legally represent their member in law-suit are all activities based on each nation's specific requirements and different laws.

Indeed nowadays a law suit against a user requests a complex and long iter. With option 3 what will happen if a commercial user of a country A infringes copyright with respect to a repertoire managed by a CRM of a country B on behalf of a right holder of a country C ?

5.8. Employment (point 4.8 of the text)

New employment opportunities could favour those (richest) countries in which the *three/four* CRMs would operate and not the countries whose CRMs are out of the competition .

5.9 Consumers/prices (point 4.9 of the text)

Option 3 " *would allow for premium content to be priced higher because it gives the collective rights manager who has attracted such content a very strong bargaining position vis-à-vis commercial users"*

In addition to the above consideration, it seems that only *premium content* could benefit of a more economical consideration

5.10 Impact outside the EU (point 4.10 of the text)

The hypothesis is that for example any African or American right holder could have their rights managed in Europe by becoming member of a CRM of their choice.

5.11 Impact on specific groups (point 4.11 of the text)

5.11.1 Very large CRMs (pont 4.11.1 of the text)

Option 3 "will, of course, have the most significant impact on CRMs" ,which means on the largest one, as the statement is written under this section

5.11.2 Large or medium size CRMs (point 4.11.2 of the text) Commission recognizes the important role undertaken by these CRMs at national level and with regard to local repertoire. However it is also recognised that they will hardly compete on the basis of the repertoire represented. Therefore "*smaller CRMS which do not attract the membership for the provision of on-line exploitation might find new roles in providing services on behalf of the CRMs to which a right-holder has entrusted his online rights. These CRMs could act as agents in relation to each of the service elements that comprise the collective management of copyright.*

Our question is what kind of new service or role are these CRMs expected to provide.

However Commission underlines that smaller CRMs "*on account of their efficiency, can attract right-holders from other jurisdictions."*

The Universal/SABAM case is reported as a positive example. We would like to underline that it is well known by all music industry that Universal's choice of SABAM for licensing of mechanical rights was manly based on the favourable rates that SABAM offers to users. In short, Universal pays mechanical right at a most convenient price while right holders (in this case authors, composers and publishers) have less advantageous royalties.

This is an example on the impact that negotiation of rates with users could have on rights of right holders.

5.11.3 Right holders (point 4.11.3 of the text) The elimination of administrative costs inherent to the reciprocal agreements will increase the CRMs' resources and therefore will allow *CRMs to transfer a considerable amount of the royalties to their members* However CRMs has to invest in order to grant efficient and competitive services. The entity of these investments and the right holders economical benefit is not foreseeable in advance.

Again in this section the Commission underlines that option 3 "*would be especially attractive to authors of musical whose work is exploited on a Community-wide scale*. For *these authors there is little incentive to choose the local CRM…" "Option 3, by allowing* ***international*** *authors to opt for direct membership in a CRM of their choice, would increase the amount of revenue received"*

As already stated, option 3 will be the best solution for "international star".

5.11.4 Online content provider (pont 4.11.4 of the text) Again in this section Commission underlines that "*Option 3 would increase competition on the level of the right-holders and will lead to the emergence of powerful CRMs* who *effectively defend right holders*

interest vis-à-vis powerful commercial users at a pan-European level. "

In the future, few powerful CRMs will license the rights of few international right holders toward few and powerful commercial users.

Nowadays in the on line environment powerful commercial users are already more interested in major companies' repertoire rather than in the one produced by small and independent producers. It is worth mentioning what happened in Italy (and in other countries such as UK) with respect to the **I -Tune case**.

I-Tune, after attracting and obtaining major's companies repertoire, didn't find interesting to obtain the independent repertoire eluding or in same cases refusing any contact or deal to this end even through independent producers' association (as AIM, the British Association of Independent Music, and AFI in Italy) .

6. Conclusion

As already stated (AFI position on Commission Communication COM/261 of the 14th April 2004) before proposing a legislative intervention, the Commission should take more into account the different situations in different Member States for the management of related rights. Especially where the management and collection of related rights involves more than one collecting society each of them having a different legal status, different rules, different efficiency and, above all, a different approach to the management of the rights.

- It is necessary to recognize that the owners of related rights have ***not*** entrusted their right collection only to a single agent (IFPI) in all the EU Member States. In some Member States more than one collecting society of related rights independent form IFPI is currently active Therefore IFPI do not represent all right holders

- Pan European licensing of related rights must be based upon the common acceptance and respect of related rights principles and rules

Nowadays, in absence of a proper respect of the plurality of right holders and without the respect of the law in force (i.e. identification of the right holder) even option 2 will mainly favour major record companies. It is easy to understand that, through their reciprocal agreement among their local members, they will maintain the control and management on a "label record" basis of the on line repertoire.

The conditions for exercising intellectual property rights vary depending on the rights, the repertoires involved and the right holders. No agreement can therefore be established as a valid model for all rights and sectors.

As technologies and business models are evolving no one solution should be set in advance. In addition to the above indication we believe that any solution related to the European collective licensing should be developed by the market, on a voluntary basis and on the basis of contractual freedom.

Workshop on
Virtual Goods Technical, Economic and Legal Aspects of Business Models for Virtual Goods

Organised by

Virtual Goods

Coordinated by

R diger Grimm, Multimedia Applications, Technical University, Germany

J r gen N tzel, Multimedia Applications, Technical University, Germany

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

The Future of a Service-Focused Recorded Music Industry

Stephen Graham Saunders
School of Business and Economics
Monash University
Stephen.saunders@buseco.monash.edu

Abstract

The music industry is currently undergoing a fundamental shift from distributing music through traditional retail stores to a number of different digital distribution channels such as internet downloads and peer-to-peer (P2P) digital file-sharing networks. The purpose of this paper is to broaden the understanding about the fundamental changes that are taking place in the recorded music industry and the marketing strategies needed to cope with this change. This paper argues that the current strategy adopted by the recorded music industry to legally shutdown some of these distribution channels, such as the unlicensed internet downloading and P2P networks, is ineffective and will only alienate music consumers. To support this argument, the paper explains that the music industry is basically shifting from a product (good) focused industry that had control over the distribution of recorded music, to a service-focused industry that has very limited control over the distribution of digitally recorded music. The paper will examine the inherent difference between goods and services, concluding that the future success of the recorded music industry is as a multi-dimensional service-focused industry where the customer controls the distribution channels.

1. Introduction

The music industry is currently undergoing a fundamental shift from distributing music through traditional retail stores to a number of different digital distribution channels such as internet downloads and peer-to-peer (P2P) digital file-sharing networks. While these new distribution channels have led to a renewed customer interest and resurgence in music, the recorded music industry (particularly the major record label groups) has financial suffered, since most of the internet downloading and P2P file-sharing is unlicensed. The long-term future success of the recorded music industry will rely on its ability to transform itself into an industry that can face the challenges that these new digital distribution channels pose.

The purpose of this paper is to broaden the understanding about the fundamental changes that are taking place in the recorded music industry and the marketing strategies needed to cope with this change. This paper will argue that the current strategy adopted by the recorded music industry to legally shutdown some of these distribution channels, such as the unlicensed internet downloading and P2P networks, is ineffective and will only alienate music consumers. To support this argument, the paper will explain that the music industry is fundamentally shifting from a product (good) focused industry that had control over the distribution of recorded music, to a service-focused industry that has very limited control over the distribution of digitally recorded music. The paper will examine the inherent difference between goods and services, and the approach that is needed by the recorded music industry to successfully operate in a service-focused environment. The paper concludes by providing a unique set of challenges that the recorded music industry must face if they are going to survive the transition from a goods-focused marketing strategy to a service-focused marketing strategy.

2. Background

Before music was encoded in digital form and stored on computer hard-drives, MP3 digital players and other digital devices, the distribution channel for recorded music was through traditional (bricks and mortar) or mail-order retailers. By distributing music through traditional retail stores, the music recording industry had control over the distribution of the product (the product being the CD, LP or tape). Over the last decade, with the introduction of digital technology and then the internet, the number of distribution channels for recorded music has increased dramatically. Furthermore, in recent times, with the advances in digital music compression algorithms, and increasing bandwidth, the distribution channels for

recorded music have fundamentally shifted to internet downloads. With music now being distributed digitally over the internet, the music recording industry has lost considerable control over the distribution of the product. In fact the product (i.e. pre-recorded CDs) is ceasing to exist, rather being replaced by a service provided over the internet. Varadarajan & Yadav refer to this shift as "product digitisation" [1].

The following list offers a brief description of some of these new "product digitised" distribution channels for recorded music:

1. *Traditional Retailers that Offer In-Store Downloads.* These retailers offer digital music in-store that can be downloaded onto a portable digital device (e.g. Apple iPod, Creative Nomad, Sony MD, Rio, Samsung MP3 digital player) or directly burnt onto a CD. A number of major retailers such as Virgin Megastores, Tesco (UK) and Walmart (US) are now offering in-store downloads. Other retailers that are normally not associated with distributing recorded music (e.g. Starbucks coffeeshops) are also now offering in-store downloads.
2. *Licensed Internet Downloads. Licensed internet downloads can be directly downloaded (or streamed) onto a personal computer, portable digital device or burnt onto a CD. Some of the major licensed internet download service providers include Apple iTunes, Microsoft MSN Music, Napster, Sony Connect, Rhapsody, and Virgin Digital.*
3. *Unlicensed Internet Downloads. Numerous unlicensed internet download services are currently available. These include unlicensed websites, File Transfer Protocol (FTP), linked websites, and Peer-To-Peer (P2P) networks. Some of the major P2P networks include KaZaA, eDonkey, Gnutella, DirectConnect and BitTorrent. While most P2P networks are unlicensed, there have been some attempts to offer licensed P2P networks (e.g. Snocap, Qtrax and Peer Impact).*
4. *Mobile Phone Downloads. Mobile phone operators such as Vodafone and Orange offer digital music that can be downloaded directly onto a mobile phone.*
5. *Internet Downloads Directly From the Artist. Some artists are distributing their recorded music directly to consumers through their own websites (e.g. David Bowie) so bypassing record labels. This is commonly referred to as "disintermediation" from artist to customer.*

Some of the new internet distribution channels such as the unlicensed P2P networks have had a profoundly negative effect on recorded music sales for both major record labels and smaller independent record labels. Forrester Research alleges that 36 million of the 40 million individuals downloading music in Europe were not paying for the downloaded music [2]. This unlicensed downloading costs the recorded music industry millions of dollars each year, threatening its very livelihood. The International Federation of the Phonographic Industry (IFPI) estimates that global music industry sales declined by 22% over a five year period to 2003, a reduction of over US$6 billion in revenue. The reaction to the unlicensed downloading by recorded music industry has been to attempt to legally shutdown the P2P networks. According to the IFPI Digital Music Report (2005) in 2004 the music industry launched 7000 legal actions for piracy in North America and Europe [3]. In response, the customers who use P2P networks describe the recorded music industry as "greedy monopolists" and celebrate their demise as a form of Schumpeterian "creative destruction" [4].

3. Distinguishing Features of Goods and Services

To understand why this tension exists, it is important to reveal that the recorded music industry still operates in a product-focused environment while the P2P networks operate in a service-focused environment. According to Pine II & Gilmore the most basic and fundamental difference between a good and a service is that a good is tangible and a service is intangible [5]. This difference is significant, as CDs purchased from retail stores could be classified as a good while digital music downloaded from the internet or P2P networks could be classified as a service. While this bipolar classification is rather crude -on a goods and services continuum- CDs could be classified as "tangibles dominant" while digital music downloaded from the internet and P2P networks could be classified as "intangibles dominant" (see Figure 1). Varadarajan & Yadav further propound that digital products are distinct from other "intangibles dominant" products in that product distribution can potentially take place exclusively through the internet [1].

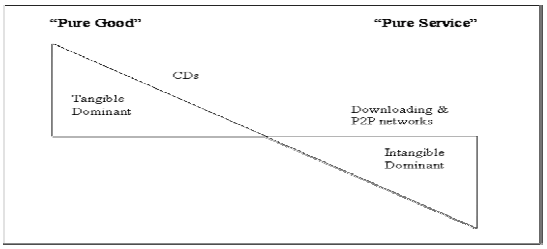


Figure 1. The goods and services continuum

Pine II & Gilmore further highlight that a good is "inventoried after production" while a service is "delivered on demand" [5]. CDs are kept as inventory in retail stores while digital music downloaded from the internet is downloaded on demand.

Another distinction, as advocated by Kotler, Brown, Adam & Armstrong, is that a service is capable of "synchronous conversion, delivery and consumption" [6]. New technologies are allowing digital music to be streamed via the internet onto PCs and existing home stereo systems. This allows for "synchronous conversion, delivery and consumption" through licensed subscription services at standard monthly subscription fees or through unlicensed webcasting.

As music is functioning more and more in a service-focused environment, it is vital that the recorded music industry move from a single product-focused (CD) marketing strategy to a fully integrated service-focused marketing strategy (that includes P2P networks). In order to achieve this, the industry needs an appreciation of the primary nature of a service.

4. The Three Levels of Service in the Music Industry

An analysis by Kotler et al. distinguishes between three different levels of a product or service offering, namely: the core service level, the secondary service level and the augmented service level.

4.1. The Core Service Level

The first level is defined in terms of the core benefits the product or service offers the consumer. In the case of the recorded music industry, the core service is the "experience of music" [7]. In fact, authors such as Pine II & Gilmore suggest that modern customers are "buyers of experiences" and that all industries are really just "providers of experiences" [5].

4.2. The Secondary Service Level

The second level is known as the actual product or service. With CDs, the actual product was well defined with the record company having complete control over the pricing, branding, packaging, and quality levels. The control over these factors was largely because the CD was tangible in nature.

With internet downloads the actual product or service is not well defined with the record company having very little or no control over the product features, pricing, branding, packaging, and quality levels. The reason the product or service is not well defined is that there are many different distribution channels that are evolving such as website downloads, webcasting, podcasting, P2P networks, streaming audio, mobile downloads, on-line video gaming. Within these distribution channels the product features are determined and controlled by the customer, so blurring the distinction between customer and artist. For example the customer can determine which individual songs to download, rather than been forced to buy an entire album. Individual songs can also be remixed, sampled, or manipulated to suit the customers taste or genre. For example, toolkits with all of the necessary software to create new remixes and samples are available on-line. On the BitTorrent P2P network, Jay-Z's (a famous rap artist) Black Album is available in nine different variations with over 1200 clip art images, and a couple of hundred megabytes of samples and breaks. Artists such as DJ Danger Mouse remixed the vocals from Jay-Z's Black Album and the Beatles' White Album and called his creation The Grey Album.

4.2.1. Pricing. The pricing of internet music is a highly contentious issue. Licensed music download companies have adopted a pay-per-track pricing policy that is marginally cheaper than the cost of a CD purchased at a retail store. However, as most internet music downloads are unlicensed, this accounts for a very small percentage of music downloads. Fisher III argues that because of the ease of acquiring unlicensed music downloads, a completely new compensation system for the record industry needs to be developed. One suggestion put forward by Fisher III is to develop a way of tracking digital copies of songs on P2P networks, and then, in theory, implementing a payment system through taxation. Kusek & Leonhard are sceptical about the potential success of any payment system for music downloads, as P2P networks are impossible to monitor or track. They rather recommend that the music downloads should be priced

through bundling various media types, mobile and internet subscriptions, multi-access deals and value-added services. Kusek & Leonhard predict that the future revenue streams of the music industry will increasingly come from the value-added services such as artist management, publishing, touring, and merchandising rather than through recorded music [7].

4.2.2. Branding. According to Palmer the "purpose of branding is to identify products as belonging to a particular organization and to enable differentiation of its products from those of its competitors" [8]. In the recorded music industry, the artist or band has always been the most important brand, however, before the internet the record label was the lever that built and developed the brand. Without the record labels financial backing, it was very difficult to build a strong artist brand. With the internet though, artists are able to build and manage their brand by building direct relationship with the public (with very little available finance). By doing so, the artist is able to create a direct distribution channel from artist to consumer.

The less important brand in the recorded music industry has historically been the record label itself. The record label often branded itself according to a particular genre of music or style (e.g. Def Jam or Motown). More sophisticated consumers would purchase music from a particular label, trusting that the music would offer a similar "musical experience" and quality to their other artists. Today, the record label brand value is being eroded by the internet, as internet download services and P2P networks are becoming brands in their own right. In other words the service itself (KaZaA, eDonkey, Gnutella, DirectConnect and BitTorrent) is the brand, not the record label.

4.2.3. Packaging. Packaging in the traditional sense was very important to the recorded music industry prior to the internet. In the 1970's in particular, LPs came out with elaborate artwork which added to the appeal of buying a licensed copy of the LP. Often the artwork, lyrics and pictures on the album was the only connection with the artist, outside of live performances. With the introduction of the CD in the 1980's, the packaging became less appealing due to the size of the disc. The internet further eroded the appeal of the album cover as customers could connect to the artist via the internet, downloading artwork, pictures or any other material directly from the artist.

4.2.4. Service Quality. The quality dimension of purchasing CDs as opposed to music downloads is quite different. Palmer contends that in goods marketing "quality can be understood as the level of performance of the product". In services marketing, "quality is the *perceived* level of performance of the service". As services are intangible, with service quality based on perceptions, measuring service quality by the customer is often more difficult than measuring goods quality. In addition, the evaluation processes for music downloads is more abstract, more random, and heavily based on symbology rather than on concrete decision variables [9]. For this reason, Zeithaml, Parasuraman & Malhotra have developed an electronic service quality (e-SQ) measure that includes scale items for information availability and content, ease of use, privacy and security, interactivity and entertainment [10].

4.2.5. Word-of-Mouth. Word-of-mouth and P2P networks are vital sources of consumer information. While the power of word-of-mouth per se is nothing new to the recorded music industry, P2P networks have become the new voice for word-of-mouth. The purchase of recorded music is driven by the power of word-of-mouth and networks of consumers which, for a period of time, become immensely loyal to a certain genre of music, band or artist. Bands or artists can go from virtual obscurity to worldwide fame in a matter of weeks, owing to the power of word-of-mouth. Before internet downloads the retail store was the central node (or common meeting point) for consumers to network and to discover the latest band or artist. With the internet, there is no central node (or common meeting point) for consumers to network. Rather the networking or word-of-mouth takes place through the decentralised P2P networks. Unlicensed P2P networks such as KaZaA (with sophisticated P2P software) have managed to lead the way in creating decentralised networks of consumers in a virtual space. Not only can consumers download recorded music for "free" but they can engage in discussion groups and fan clubs with likeminded consumers from around the world.

P2P networks are, as a result, able to provide independent credibility to the review of artists, allow customers to share experiences in an honest and open medium, and create a "buzz" around new artists or genres [11]. Salzman, Matathia & O'Reilly affirm that "buzz" enables the customer to "experience a brand rather than simply use it". As the core service of the music industry is the "experience of music", it is questionable that the music industry is intent on closing down the P2P networks that create the "buzz" around new artists or genres [12].

4.3. The Augmented Service Level

The third level is known as the augmented product or service. The augmented product or service is the additional services and benefits built around the core and actual product or service. Before the internet, the mass marketing approach adopted by the major record labels concentrated on selling volume and often neglected building direct relationships between artists and customers. Record labels offered very few additional services and benefits other than limited edition CDs, live recordings and artist promotional tours. With the internet this all changed as the artist and the customer were able to build direct relationships with each other without the need of a record label. A survey by Pew Internet & American Life Project on USA musicians, showed that 87% of musicians directly promote, advertise or display their music on-line and 83% provide free samples or previews of their music on the internet [13]. The Pew Internet & American Life Project survey also showed that 84% of USA musicians use the internet or email to keep in direct contact with fans. Through fan clubs, merchandise, concert tickets, limited releases, real-time live recordings, chat rooms, and other appealing additional services and benefits, artists are able to build direct relationships with customers. This has created a fundamental shift from a mass product marketing model to a direct service relationship marketing model.

Through building direct relationships comes loyalty from the customer. In the Pew Internet & American Life Project survey, 72% of musicians claimed that the internet has helped them make more money from their music. Reichheld underscores that businesses that understand the basis of true loyalty, recognise the internet as a powerful tool for strengthening relationships [14]. However he cautions that the internet can only dramatically deepen relationships if there is a high level of trust.

5. Private Good to Public Service

The internet and broadband technology has created an environment where the sharing of music through P2P networks has allowed music to be downloaded for "free", regardless of the legal action and lack of trust shown by the recorded music industry. There are two reasons why the legal action taken by the recorded music industry is completely ineffective.

Firstly, Robert Metcalfe, founder of 3Com Corporation, highlighted that the usefulness, or utility, of a network equals the square of the number of users, a function now known as Metcalfe's Law. This means that the more customers who use P2P networks, the more valuable it becomes, and the more new customers it will attract, increasing both its utility and the speed of its adoption by still more users. The number of P2P networks is increasing exponentially on a daily basis. KaZaA, a P2P network owned by Sharman Networks has currently over 60 million global customers [15]. Even if KaZaA is legally shutdown, KaZaA's customer would just migrate to one of the hundreds of other P2P networks.

Secondly, Lessig argues further that the internet has changed the economic characteristic of recorded music from a "private good" to a "public service" [16]. The Oxford Dictionary of Economics defines a public good as a good or service that is "open to use by all members of society". A public service has two characteristics: it is non-rival and nonexclusive. Non rival meaning that the increasing cost of providing the service to an additional customer is zero. Non-exclusive meaning that a customer cannot be excluded from consuming the service [17]. This is certainly the case of music downloads on the P2P networks, where songs can be downloaded millions of times with no additional cost and nobody is excluded from downloading the music from the P2P network.

With the power of Metcalfe's Law and the nature of a "public service", it is and will be impossible to prevent the unlicensed P2P networks from distributing and sharing music for "free".

6. Conclusion

If internet downloads and P2P networks are indeed the preferred distribution channel of the future, it would make strategic sense for record companies and other industry players to accept that P2P networks can become another legitimate distribution channel, even if there is no direct payment for the music. By shifting from a product-focused strategy to service-focused strategy, the recorded music industry will be able to face the challenges that these new digital distribution channels pose.

Rather than engaging in costly legal actions that attempt to shutdown P2P networks, record companies could instead spend more time and resources in developing service partnerships with these P2P networks. Legal actions against P2P networks only alienate the P2P consumer, so strengthening their loyalty in promoting this distribution channel. Only

through service partnerships with the software developers and consumers of major P2P networks, can record companies gain some control over these distribution channels. With the record company controlling the band or artist, surely the record companies could come up with a number of creative ideas and incentives to ensure that these P2P networks would be prepared to legitimise their services in return for better contact/connection to the band or artist?

Moreover, Van Raaij & Poiesz propose that successful marketing in the future will

- integrate products and services into cross-domain packages,
- build long-term relationships between suppliers and customers; and
- use information technology to interact between suppliers and customers, and create customisation of products and services to individual characteristics and desires [18].

Internet downloading and P2P networks are able to capture all the aspects proposed by Van Raaij & Poiesz. As Kusek & Leonhard note: the future success of the music industry will rely on giving customers a "completely integrated and cross-marketed mix of recorded music, live shows, merchandising, tickets, artist access, mobile music, video games, television, radio, film, software and other publishing and information products" [7].

In conclusion, the future success of the recorded music industry is as a multi-dimensional service-focused industry where the customer controls the distribution channels. It is not a single product-focused industry where the recorded music industry controls the distribution channels.

References

[1] Varadarajan, P.R. and M.S. Yadav, "Marketing Strategy and the Internet: An Organising Framework", *Journal of the Academy of Marketing Science*, 2002, Vol. 30, No 4, pp. 296-312.

[2] Forrester Research, "Europeans Love Online Music - As Long As It's Free", August 2004, http://www.forrester.com (assessed 16 February 2005).

[3] IFPI Digital Music Report 2005, http://www.ifpi.org (assessed 18 June 2005).

[4] Fisher III, W.W., *Promises to Keep: Technology, Law, and the Future of Entertainment*, Stanford University Press, Cambridge Massachusetts, 2004.

[5] Pine II, B.J. and J.H. Gilmore, *The Experience Economy*, Harvard Business School Press, Boston Massachusetts, 1999.

[6] Kotler, P., Brown, L., Adam, S. and G. Armstrong, *Marketing*, 5th Edition, Pearson Education, French's Forest NSW, 2001.

[7] Kusek, D. and G. Leonhard, *The Future of Music: Manifesto for the Digital Music Revolution*, Berklee Press, Boston Massachusetts, 2005.

[8] Palmer, A., *Principles of Services Marketing*, 3rd Edition, McGraw-Hill, Singapore, 2001.

[9] Zeithaml, V., "How Consumers Evaluation Processes Differ Between Goods and Services". In Donnelly, J.H. and George, W.R. (eds). *Marketing of Services*, American Marketing Association, Chicago, 1981.

[10] Zeithaml, V., Parasuraman, A. and A Malhotra,. "Service Quality Delivery Through Websites: A Critical Review of Extant Knowledge", *Journal of the Academy of Marketing Science*, 2002, Vol. 30, No. 4, pp. 362-375.

[11] Silverman, G., *The Secrets of Word-of-Mouth Marketing*, American Management Association Books, New York, 2001.

[12] Salzman, M., Matathia, I. and A. O'Reilly, *Buzz: Harness the Power of Influence and Create Demand*, John Wiley and Sons, New Jersey, 2003.

[13] Pew Internet and American Life Project 2004. "Artists, Musicians and the Internet", Pew Internet and American Life Project, Washington D.C., http://www.pewinternet.org Assessed 18 June 2005).

[14] Reichheld, F.F. *Loyalty Rules: How Today's Leaders Build Lasting Relationships*, Harvard Business School Press, Boston Massachusetts, 2001.

[15] Sharman Networks, http://www.sharmannetworks.com (assessed 17 June 2005).

[16] Lessig, L., *Free Culture*, The Penguin Press, New York, 2004.

[17] Mansfield, E., *Microeconomics*. 8th Edition. W.W. Norton and Company, New York, 1994.

[18] Van Raaij, W.F. and T. Poiesz, "Rethinking the Value Concept in Marketing". In Kitchen, P.J. (ed). *The Future of Marketing: Critical 21st-Century Perspectives*. Palgrave MacMillan, Hampshire, 2003.

An economic analysis of music download platforms*

Andreas Will
Technische Universität Ilmenau
andreas.will (at) tu-ilmenau.de

Abstract

Music download platforms are analyzed with respect to economic criteria such as utility and transaction costs. Starting from customer needs, supplier objectives, and product characteristics, the analyses focus on the explanation of the reluctant growth of online music distribution. Some recommendations for improved business models are given.

1. Introduction

Online music distribution could be expected to be a front-runner of the development of markets for virtual goods: The music market has a total of a multi billion Euro volume, the consumers are comparatively young and are thus characterized by a high affinity to new technologies, and among the suppliers we find global media groups with solid financial background and investment power. All this combined with the perfect fit of music to a fully digitalized online distribution, seems to indicate quite strong reasons for a successful growth of this market. Nevertheless we observe a considerable reluctant demand on online music markets compared to the total market volume for music. Moreover, only a small portion of all downloads are legal ones; in Germany for example, 8.5 Mio downloads out of a total of 475 Mio downloads were legal (1.8%) [2].

This paper is aiming at an economic explanation of the reluctant development of online music markets. Therefore we use an analysis and comparison of some of the current online music distribution platforms and their respective suppliers. An analysis of the reluctant development can start from three points relating to each other: (1) the supplier side, (2) the product side, and (3) the consumer side. Thus, the remainder of this paper is organized as follows. After a brief overview of the method used for our study (section 2), section 3 presents the results from consumer, supplier and product side. Section 4 discusses the results and gives recommendations for improved business models to overcome the shortcomings of reluctant demand on online music markets.

2. Method

Within a project investigating the state of the art of current digital rights management (DRM) application systems concerning privacy and usability [1], several analyses were performed which can help to a better understanding of the problems mentioned above.

By means of a detailed inspection, a functional analysis determined the consumer utility of several online music platforms. By a transaction cost analysis the gross utility is put in contrast to the cost of performing the transaction. Functional analysis and transaction analysis correspond to the measurement of effectiveness and efficiency within the scope of Software Engineering's usability testing.

Within a second step supplier strategies were investigated to find out the contribution of online music distribution to strategic portfolio of the respective suppliers. By further analyzing the business models, these results were refined on a more detailed level. Especially, we learned about the revenue model and obtained some interesting hints concerning the product side.

The analyses included four online music platforms, currently available on the German market, i.e. Apple

* This study was sponsored by the German Federal Ministry of Education and Research under the program "*Innovationspotenziale der Informationstechnik"* 2005 [1].

itunes, T-Online Musicload, Sony Connect, and Bevision / Potato System.

3. Results

3.1. The consumer side

In general, consumers are interested in a broad and long-term supply of a maximum number of music titles, older ones as well as new comers. Given the supply, consumers decide on a transaction basis. A single transaction on an online music market will take place, if the consumer is better off after the transaction compared to his respective opportunities (e.g. doing nothing, buying offline, or downloading illegally). Consumers are better off by maximizing their net utility, i.e. the gross utility minus the price minus the transaction costs. Thus, their reluctance might be explained by a combination of three reasons: (1) the gross utility of the product is too little, e.g. they miss the CD booklet, (2) the price is too high, e.g. compared to the expected sanction of illegal download, and (3) the transaction costs are too high, e.g. due to poor usability of the download platform.

We will concentrate on aspects where - compared to the opportunities - special advantages of online distribution could be expected. We look at the information and search stage of the online transaction; high quality search at low cost seems to be one central advantage of online vs. offline distribution models. Further, we look at community building, having in mind the sophisticated recommender and rating functionality of successful online communities such as Ebay or amazon.com. Another important point is personalization, a feature not available in the offline world. Finally, we look at (the restrictions to) the use of the downloaded content, compared to the free use of illegally downloaded content.

Even if we see some innovative aspects (e.g. user specific hit lists, music news, or the "artist of the week"), none of the suppliers offers *information* and *search* functions which are to a large extent satisfactory. Combined search, fault tolerance, and pre-listening are standards which are offered by three of the four systems. No system offers complete additional information such as publication dates, belonging to an album, discography of the artist, downloadable single/album charts, or editorial recommendations and reviews.

Additionally, help functions, including FAQs, E-Mail and phone hotline support, feedback and contact forms, are insufficient and in some cases not easy to find.

Community building is marginal compared to platforms like Ebay or amazon.com. Two platforms offer little functionality to build and publish user lists, and to comment on artists, songs, or albums. We do not find recommendations like "who bought this … also bought" or "who searched for … expressed interest in".

In no case, search and buying behavior influences the offers presented to a registered user. Moreover, no platform allows to *personalize* the offers presented to the user.

Concerning the use of the content, one system dispenses with DRM; free use and forwarding of the content is possible. Among the DRM based systems, only one comes with transparent and extensive user rights, the others are restrictive and not transparent.

All in all we find a considerable lack of customer orientation.

3.2. The supplier side

Suppliers consider offering music online in the context of other media markets: offline music markets, TV and radio markets, markets for live concerts and so on. Strategic planning will lead to cross media concepts aiming at maximizing the supplier's market value. Consequently, an analysis of the supplier side has to take in account the firms' strategy and the resulting business models [3]. Here, we concentrate on business models, especially on the respective revenue sources.

Our analysis of the supplier side showed that four types of revenue sources were more or less apparent on the online music market:

1. Content, i.e. music files, is sold at low transaction costs. Suppliers try to build a large customer base and are aiming at keeping customers for repeated purchases.
2. Content is combined with additional products and services such as meta information, fan products, or concert tickets.
3. Content is used to sell complementary goods such as iTunes for iPod.
4. Banner ads etc. - placed on the content platform - are sold.

All business models we observed concentrate on type 1 revenues: the traditional model of selling CDs or DVDs is transferred from the offline world into an open Internet environment. PotatoSystem gives an example of an extension to type 2 revenues. Apple iPod and iTunes is an important example for type 3 revenues. Finally, additional type 4 revenues are generated by all suppliers. But all in all, type 1 revenues remain the core of all business models in use. Suppliers try to defend this revenue source by legal and moral arguments. We do not see innovative business models; we do not see suppliers who 'proactively create novel kinds of distributing music' [4].

3.3. The product side

Finally, the analysis of the *product side* asks whether suppliers sell the right product to consumers, with respect to consumer preferences, consumer opportunities, and characteristics of the product. Here, we concentrate on the aspect of pricing.

We observed that selling (music) data files remains the core business model of download platforms (type 1 revenues, cf. 3.2). Consequently, the *product* is a stream of bits, transporting the same information as it can be found on a CD. The suppliers go to the market and try to exchange data files for money within single transactions. By realizing this model, suppliers seem to ignore a common principle of setting prices: the orientation by customers' willingness to pay.

From an economic perspective, buying a data file which is alternatively available at no cost on an illegal platform can be explained as follows: The customer values a prospective punishment and expected costs of (illegally) downloading defective or infected files higher than the price of the legally purchased file. The customer is willing to pay for the avoidance of expected transaction costs when downloading illegally. He is not willing to pay for the usage of the data.

Obviously, for many customers the expected value of fines is very low, and the protection against infected files is available easily at low cost. This gives evidence of the problems of a successful market development.

4. Conclusions

A better customer orientation is a necessary prerequisite to overcome the shortcomings sketched above. Customer orientation builds trust, it creates value to customers and lowers transaction costs.

Customer orientation consists of a bundle of interdependent topics, we discuss the most important ones in the context of music download platforms.

- Building trust: Principles of *privacy* (cf. [5] in this volume) should not only be seen in their function as constraints of a legal business; their improvement should rather be an important corporate objective. A high privacy standard is a competitive advantage on markets for virtual goods. In particular, this holds true compared to illegal platforms. Obviously, communicating a high privacy standard is a good opportunity to differentiate from competitors at a slight cost.

Moreover, trust is built by a good *reputation* of a supplier. A supplier's brand is an important measure to signalize reputation. Reputation building via branding is especially difficult in the music market. Primarily, the bands and artists are brands, signaling style and quality of the product. But bands and artists do not effectively signalize user-friendly download platforms – and telcos or IT firms do not signalize high quality music. In spite of these difficulties, suppliers from other industries entered the market by establishing new brands such as 'itunes' or 'Musicload'. Our analysis could not reveal, why this way was chosen (co-operations with well-known brands or extension of an own brand, e.g. Apple, to the new market would be other opportunities). Finally, it remains unclear why the strong retail brands dominating the offline market are not transferred to online music distribution.

- Creating value for consumers: Additional value for consumers is the second contribution to customer orientation. In the context of virtual goods, *individualization* of the product and service offerings is of high importance to create additional value. Like in other fields of electronic commerce, user profiles could be used to offer individually tailored product bundles. Moreover, individualization could be used in a different way from offline markets, e.g. by creating individual albums on request. Thus, individualization would open up novel revenue sources to suppliers by skimming off the differing willingness to pay among the consumers. Considering the prevalent model of selling data files on a single transaction basis, the industry is at present far away from establishing individualization strategies. Instead of differentiating from offline markets by novel

individualization models, we observe suppliers cautiously avoiding supposed cannibalization of offline distribution channels.

- Reducing transaction costs: Finally, low transaction costs help to build customer orientation. As long as online distribution reflects the offline world by only selling music on a single transaction basis, transaction costs has to be evaluated in comparison to the traditional retail model of selling CDs and DVDs. Thus, the *usability* of the download platform has to be designed in a way that obvious advantages of buying online (saving in time and distance) are not neutralized by a complicated operation, or by difficulties due to restricted digital rights: A consumer who has to move to a different country or has to buy new hardware, is hardly willing to invest time and money in legal or technical issues, but wants to use his music data as easy as taking out a CD from the removal van and putting it in a new CD player.

These points indicate some of the advantages we can expect when customer oriented business models are established. It also became clear that some of the measures could be realized by the suppliers on their own, e.g. promoting high privacy standards; others however will probably not take effect within the prevalent constellation of firms, e.g. individualization.

Finally, we will discuss a further option for music download platforms. As we have seen, up to now all platforms sell music files. Even if their core business model is selling devices (the Apple case), on the platform itself a transaction based revenue model is in the centre. At present, suppliers have to ask themselves how cheap a music file has to be offered in order to drag users out of illegality and to make them paying customers. This 'pricing policy' implies probably negative margins. Instead, suppliers could ask if customers are willing to pay for services 'around the file'. Obviously, the currently used additional services, complementary goods, and banner ads could serve as such revenue sources. More general, a positive willingness to pay exists if additional utility – exceeding the utility of a simple usage of the music file - is generated.

Additional utility could stem from individual advice, e.g. considering preferences, mood, and occasion. The business transaction would not be a purchase of a music file, but the service of advising the user – for example combined with the right of a single use. Different pricing options such as subscription would contribute to a transparent business model. The revelation of personal data would bring along the advantage of individual service offerings.

Another example for additional utility could be an internet based radio station offering the service of purchasing the music immediately when it is on air. Here, the willingness to pay is for saving in search time and for convenience. Without even knowing title or artist (a necessity with conventional search), rapid transactions would easily be possible.

Out of the discussion of our findings we obtained some recommendations for novel business models for online music distribution (cf. [6] for further examples). The key to improvement seems to be a switch from music files as the virtual good traded on the market - to virtual services, oriented by customer preferences and willingness to pay. Such a switch might help to overcome the reluctant demand on online music markets and make them a serious alternative to both, common retailers and illegal platforms.

5. References

[1] R. Grimm, S. Puchta, M. Müller, J. Bizer, J. Möller, A. Will, A. Müller, S. Jazdzejewski, *Privacy4DRM – Datenschutzverträgliches und nutzungsfreundliches Digital Rights Management*, Study for the German Federal Ministry of Education and Research (BMBF), Ilmenau and Kiel, 2005.

[2] IFPI (ed.): *Brennerstudie 2005*, http://www.ifpi.de/wirtschaft/brennerstudie2005.pdf.

[3] W. W. Fisher III, *Promises to keep – technology, law, and the future of entertainment*, Stanford University Press, Stanford, 2004.

[4] P. Tschmuck, *Kreativität und Innovation in der Musikindustrie*, Studienverlag, Innsbruck, 2003.

[5] R. Grimm: *Privacy for Digital Rights Management Products and their Business Cases*, Contribution to Axmedis 2005, Virtual Goods Workshop.

[6] H. Varian: *Copying and Copyright*, http://www.sims.berkeley.edu/~hal/Papers/2004/copying-and-copyright.pdf, Berkeley 2004

The Value of Metadata in the Digital Music Industry

Markus Cremer
Gracenote, Inc.
mcremer@gracenote.com

Peter DiMaria
Gracenote, Inc.
pdimaria@gracenote.com

Abstract

An enormous amount of multimedia data is available to end users not only through the internet, but also on portable devices such as mp3 players. With technology manufacturers, this has kindled awareness of the need for alternative means to search and navigate large content libraries efficiently. Many different ergonomic and clever user interfaces have been designed and put to market so far to facilitate easier accessibility. However, all of these interfaces require that additional informative data attached to the content in order for them to function. This additional data, commonly referred to as metadata, is essential for the quality of the navigation experience for the consumer. This paper focuses on the existing methods to acquire this descriptive data, and on their relevance for current multimedia applications. The current relevance of such technologies is supported by analytical data, and in the conclusion an outlook is given on future developments.

1. Introduction

Recent years have seen the advent of numerous powerful applications to help users enjoy the experience of multimedia content anywhere at an given moment. The wide range of available devices and software components for the home PC bears testimony to this development. While the first generations of portable audio players only supported a very limited number of songs equivalent to barely an hour playing time [1], current devices are incorporating memory capacities that would have been considered impossibly high just ten years ago. These devices not only allow users to access an archive of music equivalent to a playing time in the range of days, they also contain multiple codecs to facilitate playback of the consumer's preferred music formats [2, 3, 4]. Portable music players are becoming more and more versatile, now capable of displaying not only archived photos on their color displays, but also video clips. Additionally, other devices not originally intended for multimedia playback are expanding into this area as well. Cell phones and PDAs with extended multimedia capabilities offer content and services to the consumer on the go [5, 6, 7]. The significant new characteristic of these devices is the expanded connectivity available through wireless LAN and GSM data channels. Thus, the user is no longer confined to synchronizing the content on the device with her or his home computer, but is able to search for music over the internet, and to receive streaming or downloaded content form a remote server. Though the bandwidth offered by wireless telephone networks is still very limited, larger bandwidths are at the brink of breaching the market in answer to this demand [8,9].

The large amount of multimedia content available via multiple services, channels and devices require an efficient means for the consumer to search and navigate both online and local content. On the other end of the chain are the music service providers that ambitiously compete in providing the consumer with the content to her or his liking by establishing elaborate user profiles [10, 11].

All of these technologies require descriptive data or metadata that is used to access a certain piece of music or a video clip. This data is most often created manually, though more and more focus is directed towards automated means to find similar items. The acquisition and archiving of metadata, be it the most basic factual data or more elaborate descriptions that facilitate enriched presentation of multimedia content is the focus of this publication. As one of the first businesses concentrating on services and applications built around metadata, Gracenote has benefited and expanded from the increasing connectivity and multimedia capability of the products and trends mentioned above.

In the following chapters, several possible methods to collect metadata for audio and video shall be explored.

Current technologies to associate metadata with content will be presented, by which users can re-organize their collections and search the internet for music fitting their individual taste. Thereafter, a number of applications and products will be enumerated where special detail will be directed on the use and delivery of metadata to the consumer. In the subsequent chapter statistical data will be presented that demonstrates the use of Gracenote's services and applications. Special attention will be directed towards relatively new types of multimedia products that are on the brink of replacing the traditional home and car stereo equipment.

The publication will conclude with a brief summary and an outlook on the immediate future of devices with multimedia capabilities.

2. Metadata Services and Technologies

The history of Gracenote follows the cliché of many successful start-up companies. The original idea of automatically sorting and cataloguing a significantly sized private CD collection led to the development of an internet-based CD recognition service. The underlying principle is the identification of a music CD using the Table of Contents (TOC) to compute a look-up index from the playing durations and order of the individual audio items. This key is then used to query a central database over the internet and retrieve matching metadata to this CD, namely the performing artist, album title and song names. If the appropriate CD data has not yet been entered into the database, the software application asks the end user to enter the data for future retrieval [12]. Though an extension to the original CD Redbook Standard [13] supports the embedding of metadata on the CD itself, only a few manufacturers make use of these data fields, and few devices in the market display this data. The service rose to immense popularity almost overnight and it soon became apparent that the hosting and maintenance of the database exceeded the financial and temporal effort adequate for a hobby. This demand, on the other hand, yielded promise of a business service that provides metadata to end users, and led to the decision to found Gracenote, Inc., a company dedicated to the archiving and delivery of first audio, and later multimedia metadata. Since then the service known as MusicID (formerly “CDDB”) is not only available to devices connected over the internet, but has also successfully been integrated in embedded environments to perform look-ups in a local database. This is particularly useful in mobile multimedia equipment such as the newest generation of car entertainment systems.

Another tangent of Gracenote's growing business portfolio is the adaptation of the CD recognition principle to DVDs. Here, the structure of the multimedia files residing on the DVD is analyzed and an appropriate look-up key is constructed to facilitate the association of metadata. The application field for such a service differs from the CDDB service, since only a limited number of consumers are currently compiling large databases of movies on their home computer systems. Instead, automotive applications are expected to be a key market, where in-car DVD entertainment systems are becoming increasingly popular. In most deployments, the DVD content displayed for viewing by the rear seat passengers is controlled from the front seat entertainment console. The requires efficient navigation of the DVD changer discs as well as the individual content chapters within each DVDs Since the location of the main movie feature on the DVD, and other information such as the movie title are not available as machine-readable data on the medium itself, a database containing this information is required to implement the above mentioned features.

The multimedia experience can be further improved for the consumer with the utilization of a new generation of algorithms designed to retrieve information from multimedia content directly. A basic form of such a technology is the identification of a multimedia item independent of its current form or storage media. While the look-up index to MusicID is retrieved from the CD TOC, so-called fingerprinting solutions analyze the music signal itself and generate a numeric representation of the audio. Since multimedia content in general, and audio in particular can be altered significantly (from a signal systematic point of view), this index retrieval method has to be robust against common changes of the original recording. Hence these fingerprinting technologies are also often referred to as "robust hashing". As a side note, fingerprinting is often confused with watermarking, another technology entirely that involves embedding additional information in an audio signal with the goal of imperceptibility. The thus hidden signal can be retrieved by an appropriate decoder even after manipulation of the signal, such as mp3-encoding, and is commonly used to “brand” a particular multimedia item. Though this technology can also be used for the purposes of metadata association, fingerprinting is the far preferable method for indexing because it does not require the embedding of additional information but

instead derives all necessary data from the content itself.

Fingerprinting can be deployed in a variety of ways depending on its primary application. While in a file based environment only a short excerpt of an audio item, e.g. the first 15 seconds, needs to be analyzed and stored, other scenarios target streaming content and have to be able to identify an item at any position into the audio. For the later case, data for the whole audio item is needed in the reference database. Applications range from identification of a song aired over radio or TV to the identification of content using a mobile phone service. Here, the end user dials a service number, and points the handset to the audio source. The audio is transmitted to the service provider through regular communication channels (e.g. GSM network) and analyzed on the server site. Since the audio quality is harshly degraded by environmental noises, the speech codec and the transmission channel, the original signal arrives strongly altered and distorted at the server site. This calls for ultra-robust fingerprinting technologies that in themselves demand a comparably large amount of fingerprint data per time segment. The technology used by Gracenote has been developed by Philips Research and enables successful recognition of the audio item after as short a time segment of four seconds [14].

The successful association of metadata to multimedia content opens numerous doors for potential applications. By delivering rich descriptive metadata to the end user, intelligent automatic playlisting can be achieved. The user can either sort content by directly accessing this data, or use the data related to one particular item to search for similar items. The latter method is commonly referred to as "query by example".

Not only are these methods useful for navigation through personal archives, but they apply also to online delivery of content and multimedia stores [15, 16]. Here, content can be browsed using stylistic descriptions such as rich genre, region and era metadata. Since download portals and online retail stores usually have their own proprietary way of archiving and accessing their content, methods have to be provided for semi-automatically matching this content to Gracenote's metadata archive. This service, offered under the product name “Link” comprises automated matching according to textual as well as available numeric metadata such as e.g. UPC and ISRC [17]. Additional manual matching is often required for a complete mapping of the archive. As databases are constantly updated (e.g. with new releases), this is a continuing process.

3. Descriptive Data Acquisition

The great benefit of Gracenote's end-user-based data acquisition system is its scalability and global coverage. Additionally, since end-users enter the metadata, it is most likely that the same phrasings and spellings will be used when searching for this content. The most common forms of such spellings and misspellings provide valuable data for optimizing commercial search and retrieval applications and services. Also, in times of increasing market globalization, the textual representation of artists and titles in countries using non-western character sets is required for enabling international servicing.

However, for professional applications such as download portals, the traditional method for collecting metadata by consumer entries is not sufficient alone to provide the required level of accuracy and consistency in the database. Elaborate filtering and text processing mechanisms that follow a set of statistically determined vote of majority rules and other heuristics have been designed and implemented to filter out spelling errors, data that was entered in the wrong data fields, and additional irrelevant or inappropriate data such as links to music download portals and malicious misinformation.

This carefully calibrated automatic distillation of essential metadata still is not always satisfactory for services that require the utmost consistency of data. The cataloguing order of words, for example, calls for a significant manual editorial effort. For illustration of this predicament a few examples are given:

- The artist’s first and then last name, or the last name followed by a comma and the first name,
- The frequently encountered omission of part of the name e.g. the article 'The' in the name 'The Beatles',
- Rules for capitalization and special treatment of non-standard names such as AC/DC and U2.

Another method for data acquisition involves partnering with other music industry organizations such as record labels and music publishers to submit metadata for their own content directly to Gracenote. A combination of software and services, referred to as the Content Management System, facilitates the submission of rich metadata along with the submitter’s proprietary identifier to the Gracenote Media Database. This way, a reference link to the submitting business

can be easily facilitated that allows end users to purchase or browse for more content offered by the partner.

4. Application Scenarios

As detailed above, the originally targeted platform for using the CD look-up service was the connected PC. Consumers wanted to comfortably select tracks on their music CDs for listening, and index their content while encoding files from the CD media to their storage archive.

Since the music metadata consumes considerably less storage capacity than the actual associated multimedia content, it is feasible in more recent generations of home and car stereo devices to perform the recognition look-up and metadata delivery in a completely embedded environment. Compression techniques are well-known for textual content which allows even more metadata to be stored in such a device.

A number of audio devices are available that feature internet connectivity, and thus offer additional content services such as internet radio streaming [18]. Here, also, fingerprinting technologies can be applied in order to identify audio content by analyzing its waveform. Thus, audio files transferred to the device using other media such as USB flash drives or data CDs can also be appropriately labeled and made available for intelligent browsing and playlisting.

Taking identification even one step further, fingerprint look-ups can also be conducted on the device itself. This requires somewhat more computational and memory resources. Under the assumption that only complete audio items or songs will be looked up, and the audio quality will be acceptably high (e.g. mp3@128kbps) the hardware constraints can be feasibly overcome.

With current technologies, 500MB of hard drive memory can contain a potential reference database of approximately 7 million fingerprints, each corresponding to an audio item. Tests have been conducted using an Intel Strongarm™ processor clocked at 206MHz and with 64MB of RAM, and a 2GB PC Card hard drive. When performing 10,000 queries, the results yielded a recognition rate higher than 99% and an average look-up time of approximately 2 seconds per item.

This setup resembles a typical combination of components used for car radios and home stereo equipment.

In the mobile communication market, impulse shopping is a rapidly growing phenomenon. Consumers are increasingly comfortable with purchasing items at the punch of a button using their cell phones. To enable comfortable browsing for music the query-by-example method easily compensates for the limited user interface. For example, if a consumer hears a new song playing that they like in a nightclub, they can use the phone to identify it. The query result can also be linked to immediate mobile commerce opportunities such as ring tones, wallpapers, concert tickets, and music downloads.

User look-up statistics can enable service providers to create user profiles that can be used to generate, personalized recommendations of similar items - following the common concept of indirect customer recommendations ('customers that have bought this also have shopped for these items'). Music clip previews of recommended content can be also be inserted in consumers' playlists to spur curiosity and enable deeper exploration into the world of music.

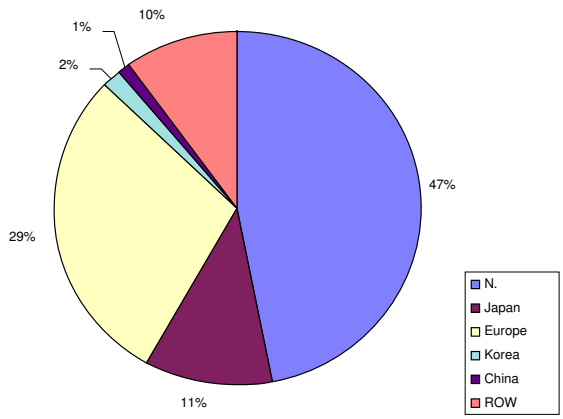


Figure 1: Unique albums per region

These examples give an idea of the manifold applications that are benefiting from audio identification and associated services.

5. Facts and Figures

In the years since the creation of the Gracenote CD recognition service, its popularity has grown immensely. Though end-user identifying data is not stored, other information is used to gather statistics which characterize the aggregate use of the service. At this time, 97 million unique end-user computers are using the CD identification per year via roughly 4,000

different applications. These users are looking up approximately 6 million CDs every day, amounting to over 2 billion look-ups over the lifetime of the service.

The underlying database currently recognizes about 44 million tracks on 3.5 million albums by 780,000 artists. To achieve this coverage, the service is available in over 213 countries and 80 languages - which in itself poses a significant challenge. Figure 1 shows a regional breakdown of the global coverage of unique CD releases.

On the submission side, about 2,500 new CDs are submitted to the database every day, with a significant number contributed by the 250 labels, artists and content owners that enter their data into the database using the Content Management System.

Since the fingerprinting technology has been more recently deployed, this database is not yet as large as the CD identification database. However, fingerprint data is currently available for over 6 million tracks, with an average of 50,000 fingerprints added daily.

It is, however, still feasible to create database subsets for memory constrained environments encountered in stand-alone consumer devices. A large number of these tracks and CDs are looked up on a less frequent basis. In particular, 250,000 albums account for approximately 90% of all album look-ups aggregated over the last 5 years. The important and determining factor, though, is the selection of the 'right' tracks and albums for inclusion in the database subsets. This in turn requires thorough knowledge of region-specific look-up statistics that can only be derived from a comprehensive and therefore significantly larger database. Furthermore, in order to assure near-perfect lookup results for any individual's unique music collection anywhere in the world, the size of the database must be well over an order of magnitude larger.

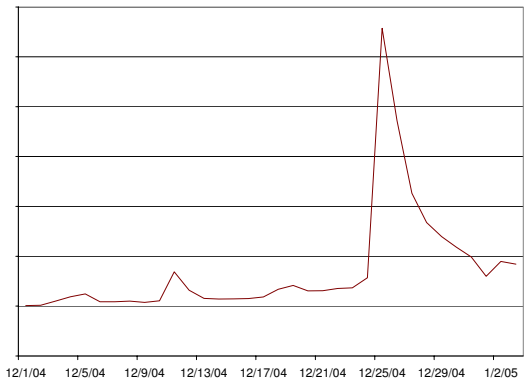


Figure 2: Increase of queries at new product launch

Interesting to observe is also the introduction of a new mass application into the market. With the recent deployment of one large consumer application during the Christmas season, the amount of look-ups increased over night almost by a factor of five. In figure 2 the number of queries over the observed period is displayed.

Another notable fact is that the market for audio books is gaining. For example, the recent release of J.K. Rowling's "Harry Potter and the Half-Blood Prince" [19] was the third most often looked up audio media item globally only one week after its release date, and two weeks later by far the most queried for.

Also, as a result of the use of promotional pre-release CDs within the industry, it is quite common that CDs are queried for identification weeks prior to the official "street" release date.

To enable additional services such as music library auto-organization, intelligent playlisting, and recommendations, more data than basic factual artist, title, and album information is necessary. For example, over 1,500 different micro-genre categories have been defined by Gracenote's music expert team and assigned to the music content. This data together with the original release date of an album, regional data and other descriptive information fuels an algorithm for automatic playlist creation that yields very satisfactory results.

6. Summary and Outlook

It has been shown that the value of metadata and identification of audio, especially music data, is significant enough on a global basis to support a viable business. And with the increase in data capacities of mobile devices such as car radios and portable music players efficient browsing and playlist creation methods will become increasingly important. A basic set of metadata such as artist, title and album help consumers sort and access music in their audio environment. However, it is a deeper, richer level of metadata that powers smart applications like playlist creation or finding similar songs to a query item. As devices become more and more connected, the currently utopian idea of exchanging audio across multiple networks and devices and hence availability of an immense number of songs from multiple sources at almost any time and any place will become more and more reality.

Another trend can be observed with the increasing computational power and memory size of small mobile devices, which leads to more complex analysis

algorithms that can be locally performed, and the accessibility of larger databases for rich media content such as e.g. CD cover art.

However, to enable a good user experience, appropriate software still has to be written and the proper data and has to be provided. A particular predicament is the design of user interfaces on such devices, where the consumer has to locate desired content by navigating numerous levels of menus. A constant demand will be required for smart interfaces that help the user to efficiently, effectively and enjoyably access the media that they are seeking at any moment. One encouraging direction for user interface evolution is the ability to control and interact with the multimedia system in a hands-free manner by using a speech interface. The individual technologies exist to day to create such an experience, but are yet sparingly used in devices.

Finally, people are experiencing music more and more on the go, Gracenote enables this on-the-go experience via its embedded metadata playlisting engine that allows users to generate novel automatic playlists anytime / anywhere without access to CDs, a PC, or even an Internet connection. Finally, in the future, where portable devices will be increasingly often connected through wireless data channels [8,9]. metadata services enable a common user interface and library functions for content that is available both online and offline. For example, if access to music subscription services and local files are combined in a single device, this should happen in a manner transparent to the end user, following his or her elementary desire: 'I want to hear music right now' - whatever content is available should be easily accessible almost instantly at any time.

7. References

[1] “Diamond Multimedia Announces Rio PMP300 Portable MP3 Music Player”, Press Release, Sept. 14th, 1998, http://news.harmony-central.com/Newp/1998/Rio-PMP300.html

[2] Apple, iPod product description, 2005, www.apple.com/ipod/

[3] Creative Labs, Nomad product description, 2005, www.nomadworld.com/products/Jukebox/

[4] RCA, “Thomson Expands RCA LYRA mp3 Product Line with Newly-Designed mp3 Jukebox and Lightweight Flash Products”, Press Release, Jan. 7th, 2004, www.rca.com/content/viewdetail/1,2811,EI700567-CI258,00.html

[5] Siemens, SL45 mobile phone product description, 2005, www.my-siemens.com/sl45

[6] Samsung, “SAMSUNG's SGH-i300 Offers Incredible Storage as the World's First Music Smartphone with 3GB HDD“, Press Release, Mar 11th, 2005, www.samsung.com/PressCenter/PressRelease/PressRelease.asp?seq=20050311_0000102071

[7] Microsoft, Windows Mobile product description, 2005, www.microsoft.com/windowsmobile/default.mspx

[8] GSM 09.18, "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Serving GPRS Support Node (SGSN) - Mobile Switching Center (MSC): Gs interface Layer 3 specification"

[9] IEC, Online Education, “Universal Mobile Telecommunications System (UMTS), Protocols and Protocol Testing: 1. Evolution of Mobile Communications”, 2005, www.iec.org/online/tutorials/umts/topic01.html

[10] Amazon.com, www.amazon.com

[11] Yahoo!, “Yahoo! Premiers Yahoo! Music Unlimited”, Press Release, May 10th, 2005, http://docs.yahoo.com/docs/pr/release1237.html

[12] International patent no. WO9847080A3, “Network delivery of interactive entertainment complementing audio recording”

[13] Red Book Standard IEC 60908 (1999-02), “Audio recording - Compact disc digital audio system”

[14] Pedro Cano, Eloi Batlle, Ton Kalker, Jaap Haitsma, “A review of algorithms for audio fingerprinting”, *Proceedings of 2002 IEEE Workshop on Multimedia Signal Processing*, Virgin Islands, USA

[15] Apple, iTunes product description, 2005, www.apple.com/itunes/

[16] 24/7 Downloads, media download shop product page, 2005, www.247downloads.com/

[17] ISO 3901:2001, "Information and documentation - International Standard Recording Code (ISRC)”

[18] Philips, Streamium product description, 2005, www.streamium.com/whatisstreamium/

[19] J.K. Rowling, Jim Dale (Reader), “Harry Potter and the Half-Blood Prince” Audio CD edition (July 16, 2005), *Listening Library (Audio)*; ISBN: 0307283658

Privacy for Digital Rights Management Products and their Business Cases*

Rüdiger Grimm
University Koblenz and Fraunhofer Institute for Digital Media Technology
grimm (at) idmt.fraunhofer.de

Abstract

DRM systems and their shops are analyzed with respect to the privacy of their customers. The analysis follows a structure of privacy principles in accordance with the European Directive for Privacy Protection. As an example, Apple's iTunes is analyzed in detail. From the analysis, recommendations for a better practice are derived.

1. Introduction: the problem with DRM

Before the efficient digitization of intellectual property such as music, films, games, images, and text, their content was strictly bound to physical media, from which they could not easily be copied and sent without loss of quality. The traditional business model of intellectual goods was the business model of their physical media. With the digitization, latest with the highly efficient compression method of MP3 (Brandenburg/Stoll 1992, and MP3 1992/94), this business model was corrupted. Customers don't pay for CDs, DVDs, or books, if they get the related content in (almost) original quality for (almost) free in the Internet.

Digital rights management compensates for the loss of physical binding, in that their mechanisms block the duplication of digital products on the end-user devices. More sophisticated, digital rights enforcement mechanisms execute specified user rights, coming as part of the content code, within the end-user devices. Digital rights management and enforcement aim to restore the classical business models in that they make sure that users can consume only those products they have paid for. Moreover, users can consume the products only in the very specific way, they have paid for. The history and functioning of DRM systems are discussed by many authors, see for example the excellent books of Rose et al. 2002, and Becker et al. 2003.

There are many DRM products out in the market, for example the Fairplay DRM kernel in Apple's iTunes, Microsofts Windows Media Rights Manager (WMRM) in many shop systems, for example in Musicload of T-Online, the OpenMG in Sony Connect-Europe, to mention only the music market. For electronic documents, Adobe PDF has integrated DRM functions to support the E-Books format. There are many more systems in the market, they are all mutually incompatible, difficult to use, and – this is the topic of this paper – they are intransparent with respect to their handling of personal data. After all, they are badly accepted by the market, compared to the market potential. The Musicload Factsheet of 05.04.2005 (www.musicload.de) with reference to the 2004 Digital Music Report of the International Federation of the Phonographic Industry (IFPI), estimates 200 million legal music-downloads world-wide, while they estimate that this covers only 20% of all downloads: that is, 80% or 800 million downloads were illegal. Moreover, the numbers of legal downloads have remained constant form 2003 to 2004. The download market of digital music is both far beyond its potential, and it is in a stagnation phase.

In our study (Bizer et al. 2005) we claim, that this is due to the fact, that DRM systems are uncomfortable to use, they don't meet the real needs of the users, and they undermine the trust of users, in that they either misuse personal data of their customers or, at least, handle personal data with little care. In this paper I will explicate especially the privacy problem of existing DRM systems.

For this purpose, this paper explains how existing DRM systems and their shops can be analyzed with respect to the privacy of their customers. The analysis follows a structure of privacy principles in accordance with the European Directive 95/46/EC on Privacy. (www.cdt.org/privacy/eudirective/). As an example, Apple's iTunes is analyzed in detail. From the

* This study was sponsored by the German Federal Ministry of Education and Research under the program "*Innovationspotenziale der Informationstechnik*" 2005

analysis, recommendations for a better practice are derived.

An economic analysis of DRM systems is discussed by Will (2005) in his contribution to the Axmedis conference 2005.

2. The privacy problem in existing DRM systems

There are three fundamentally different approaches to the protection of rights on digital items. Approach number 1 is "strong DRM" which enforces rights: users *cannot* act illegally. Copyright mechanisms are the most simple form of "strong DRM". Approach number 2 would not prohibit users to act illegally by technical means, but would personalize products in order to identify the origin of products in illegal environments. An example for this "trace model" is the LWDRM technology (Grimm/Aichroth 2004) by which users earn fair usage of digital items if they sign them: users do *not dare to* act illegally. Approach number 3 does not use technical rules enforcement mechanisms, but would encourage users to act legally by incentives: users do *not want to* act illegally. An example for this "incentive model" is the "Potato system" which encourages users into a provision model (www.potatosystem.de).

Obviously, the trace model and the incentive model must somehow use personal data in order to identify users either for prosecution or in order to realize the incentives. Both approaches must work carefully on a personal data protection model. The third approach, however, the "strong DRM", which prohibits users to act illegally by technical means, should be strictly product oriented, with no reference to the person who legally owns the product. Because the technical mechanisms enforce well behavior, there is no need to either prosecute or reward the user.

However, most shop systems which use DRM, do not trust the built-in mechanisms of DRM to enforce the usage rules in the end-user devices. Therefore they use the trace method as a second line of defense. They collect data to identify users, not only for business purposes, but also to link products to their buyers in order to identify the origin of products in illegal environments. Most often, the usage of personal data within DRM protected products is intransparent to the customers.

3. The analysis structure of the privacy principles

It is helpful to use a privacy model in order to analyze DRM systems with respect to their usage of personal data. The European Directive 95/46/EC on privacy provides such a model, which is widely accepted, not only within Europe, but also, for example, in the USA by means of the Save Harbor Principles (US DoC 2005). The directive defines "personal data" as "any information relating to an identified or identifiable natural person". The identified or identifiable person is called the "data subject" (Art. 2). The directive regulates the storage, processing, and usage of personal data by explicating the following principles

- Quality (Art. 6): the data must be lawful, fair, adequate, relevant;
- Legitimation (Art. 7): personal data must be bound to the purpose of the service, they may be used only by consent of the data subject or by a legal obligation;
- Purpose binding (Art. 7): the personal data must be necessary for the purpose, e.g. a contractual cooperation or the administration of a service, etc.;
- Transparence (Art. 10-12): the right of access by the data subject;
- User control: beyond transparence, the right of access, esp. the right of rectification (Art. 10-12), the right to object (Art. 14);
- Confidentiality (Art. 16): the organization must ensure the confidentiality of the personal data;
- Correctness and security (Art 17 on security, and the right to rectify the data, in Art. 10-12): the organization must protect the personal data against loss, distortion, and correctness with respect to the content;
- Supervision (Art. 18-19, and 28-30): regulations on a supervisory authority;
- Remedies, liability, and sanctions (Art. 22-24); regulations on the sanctions in case the service provider does not comply with the principles.

There is also an obligation to ensure that personal data may be transferred to a third country outside of the EU only if the "third country in question ensures an adequate level of protection" (Art. 25). The US Safe Harbor is an agreement which service providers may join freely in the USA in order to guarantee such an "adequate level of protection". The principles of the European Directive 95/46/EC on privacy are not exactly, but somewhat closely mapped on the seven Safe Harbor principles: notice, choice, onward transfer, access, security, integrity, and enforcement. (US DoC 2005)

Of course, these principles may be checked against any shop that offers digital music. Web shops offer privacy statements on their Web pages, for example http://www.apple.com/de/legal/privacy/statements.

However, it is hard to know, where exactly, at which

point of communication between buyer and seller, and – most difficult – for which purpose, data are collected, stored and used by a shop provider, either directly, or by means of its DRM system.

In order to find out, which data are collected and stored at which point of the communication and at which place, each Web shop and its incorporated DRM system can be analyzed according to the following categories (Bizer et al. 2005, 2.4-2.5):

- Data flow *before* concluding a deal: while preparing a purchase, during user registration, and when selecting a product for purchase (placing a product into the shopping cart);
- Data flow *at* conclusion of a deal: at end of selection (closing the shopping cart), for payment of the products, and for delivery of the products;
- Data flow by checking the right to use a product: at first initialization of a player; at repeated usage; for rights update;
- Data flow through service functions: for example improvement of service, direct marketing, and security functions such as encryption;
- Data flow through hidden interfaces and by linkage of different functions: cookies, pixel tags (web bugs), combining customer data with clickstream data such as IP addresses or encryption keys.

There is an additional category of personal data, which is orthogonal to the other five categories, which we call "general data traces":

- General data generated by the user consciously by filling forms;
- General data collected by the shop from the communication data;
- General data encoded within the product itself.

4. For example, Apple's iTunes

4.1 Overview

Apple's iTunes is a music portal with online stores in 19 countries worldwide. It offers 800 000 titles in 28 genres (by January 2005). Five major and numerous independent labels offer their titles in iTunes stores. Until March 2005 iTunes has counted over 300 Million downloads over all stores, that is ca. 40 Million downloads per month. In addition to music the iTunes stores offer videos, reading books, film trailers and radio streams.

The number of customers is not published. The target customer group is young people who love music, who are used to online surfing, and who are ready to pay for what they get.

The iTunes servers are located in the USA, they support the formats of M4P (MPEG-4 protected) and Codec AAC. The DRM system incorporated is called "Fairplay".

See (iTUnes Musicstore 2005).

4.2 Process model

Apple's iTunes works like this, see figure 1. The user interacts with a shop server (steps 3 and 5), which stores registered content from the content provider (step 1). After installing the iTunes client software (2), the user browses the shop's Web site and selects a piece for purchase (3). The shop first organizes the payment of the chosen product (4) and then delivers the content to the user (5). The user may consume the purchased products or share it with a specified number of devices according to the iTunes usage rules (6).

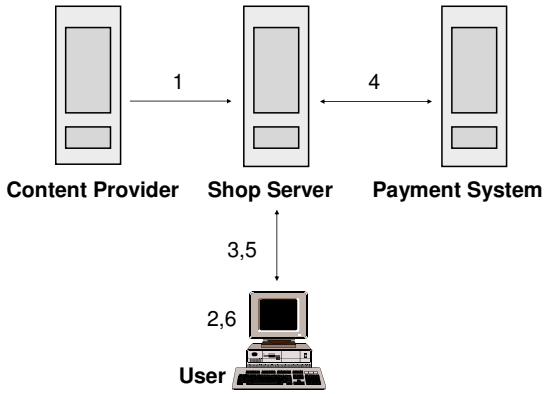


Figure 1: The process model of iTunes

The user may burn the purchased iTunes products on up to 10 CDs, and he may transfer the products to Apple iPods. He may also share the products with not more than six other computer devices which must be registered to the iTunes shop server in advance. Before a product can be transferred to a seventh computer, one of the first six must be de-registered before. iTunes products are encrypted with a symmetric key, which is encrypted by an asymmetric public key of the user device, before the product is delivered. The key pair of the user device is a function of some hardware parameters of the device, such that only a registered device can decrypt the symmetric content key (Bizer et al. 2005, section 3.1).

4.3 Data traces

Following the structure of section 3 above, we found the following personal data collected and stored by iTunes:

General data provided consciously by the customer at registration:

During registration, the customer fills a form which contains these mandatory fields:

- name, address, and telephone number;
- the user's e-mail address, which is used as his "Apple-id", and a related password on the user's choice;
- a secret question plus an answer, as well as the birth date in order to reveal a lost password;
- payment information such as credit card number and validity date.

Optionally, a client can provide his

- postal address for delivery of hard goods;
- fax number;
- mobile telephone number;
- tax number.

General data collected by the shop during communication:

- The operation system used by the client (version, language), the version of the iTunes software used, and the IP address of the client by means of the HTTP protocol;
- Cookies and session-ids. Cookies are expressively used to check how often which special sites of iTunes are visited by clients in order to improve the service; session-ids are used to organize the online session effectively;
- Device-id: for registration of a client device, the iTunes client-software derives a so-called Device-id from the hash of some hardware parameters and sends it to the iTunes shop; the Device-id will be used by the shop to encrypt a symmetric user key with which the delivered products are encrypted in order to bind it to a specific registered client device: this is the kernel of the Fairplay DRM system.

General data encoded within every iTunes product:

This part is rather intransparent to the customers of iTunes. However, it is simple to find these data in a hexdump of an iTunes product:

- Product-id, provided and maintained by the store;
- Apple-id, which is identical with the e-mail address of the customer;
- Meta data about the product such as title, name of composer, genre, year of production, etc.

It may be surprising to many customers, that their e-mail address is part of the product code. So, if they copy and send a product to other devices, their email address will reveal the origin of this file.

Data flow before concluding a deal:

Before completing a purchase, a user must be a registered iTunes customer. For this purpose, a user fills an online registration form. See above the "general data provided consciously by the customer at registration". Moreover, the registration is done over the HTTP protocol, by which more personal data are sent to the shop, like the operation system used by the client (version, language), the version of the iTunes software, and the IP address of the client, see above the first bullet point of the "general data collected by the shop during communication".

Other data are sent when using coupons from other users to fill up an internal iTunes account.

Data flow at conclusion of a deal:

A registered iTunes customer concludes a purchase in that he logs into the system with his Apple-id (which his e-mail address) and his password. The products in his shopping cart are associated with Product-ids by the shop. The shop stores the association of the Apple-id (the user's e-mail address), the list of Product-ids and the download-status. Even if the download status is "completed", this line remains in the store's database.

The shop encodes Product-id, Apple-id, and some meta data about the content within the code of every product before the content is encrypted with a user key and then downloaded by the customer.

Data flow by checking the right to use a product:

A PC must be initialized before it can play an iTunes product. Remember, that the content of iTunes products does not come in clear text, but it is encrypted with a symmetric key. The symmetric key is encrypted itself individually for any registered end-user device. For this purpose, a PC establishes a user account which contains private information to decrypt the symmetric product keys. The private decryption information of the user device is derived from some hardware parameters of the device (Device-id) and from the e-mail address of the user (Apple-id). For registration, the client PC sends its Device-id related Apple-id to the iTunes shop and, in turn, receives the necessary information to decrypt the symmetric keys that decrypt the content of iTunes products.

Another device, which is not registered, would not be able to decrypt the symmetric product keys, because the private information necessary for this step includes the local hardware parameters. Therefore, only registered end-user devices can play iTunes content (unless the DRM system is broken).

Registered devices play iTunes products offline. Therefore, on consumption, there is no more data flow between customer and shop.

Data flow through service functions

When a registered iTunes customer browses through an iTunes shop, key-words for search and visited Product-ids may be associated with IP addresses or Apple-ids.

A special feature is the so-called iMix. An iMix is a personal hit list of products which are assessed as best by the user. An iMix is associated with a self-chosen pseudonym (persona). An iMix together with its pseudonym (but not more) can be published. An iTunes customer can recommend his personal iMix to another user, in that he fills a special form which includes his own e-mail address and the e-mail address of the recipient. The recipient needs not to be an iTunes customer (but might now become one, because he likes the iMix). iMix, pseudonyms and e-mail addresses are associated by the iTunes system. This personal data reveals a certain user behavior.

Another special feature is the "pocket-money account". There is a personal user-account internal to the iTunes system. It contains a kind of "pocket-money", which can be used to purchase iTunes products. The smart idea behind the "pocket-money account" is that a registered iTunes customer can send money to another user who is not necessarily an iTunes customer. But he will certainly become one in order to enjoy the money presented to him. All related personal data (e-mail addresses as Apple-ids, money flow between the accounts, and subsequent purchases) can be linked by the iTunes system in order to learn more about its customers.

Coupons are another way by which customers can send money to other users. Again, the e-mail addresses which are or may become Apple-ids, are used for sending coupons.

Metadata of content can also be communicated in the iTunes system. There is a central metadata service, CDDB run by Gracenote (www.gracenote.com/music). An iTunes customer can access the metadata server of CDDB during his iTunes session. He can send the metadata of his content to the meta data server. And in case he has no metadata or he wishes to update his existing metadata, an iTunes customer can send content to CDDB which would recognize it send metadata back to the user. During metadata communication all associated data like Apple-id and Product-ids are accessible by iTunes.

Data flow through hidden interfaces and by linkage of different functions

At every communication step with iTunes, all clickstream data out of the HTTP protocol, including IP address, language, HTTP-referer ("from which site am I coming") and search key-words are accessible by the iTunes system and may be associated with other personal data stored in the internal data bases of iTunes, especially the Apple-id which identifies the customer by his e-mail address.

The following data set is an example of an HTTP header of a client request to an iTunes store:

```
GET
/WebObjects/MZStore.woa/wa/com.apple.jingl
e.app.store.DirectAction/viewNewReleases
?fcId=14094475&pageType=newReleases&id=100
HTTP/1.1
Referer:
http://ax.phobos.apple.com.edgesuite.net/W
ebObjects/MZStore.woa/wa/storeFront
Accept-Language: de-de, de;q=0.75, en-
us;q=0.50, en;q=0.25
X-Apple-Tz: 3600
User-Agent: iTunes/4.6 (Windows; U;
Microsoft Windows 2000 Professional
Service Pack 4 (Build 2195)) DPI/96
Cookie: countryVerified=1
X-Apple-Validation: 0F56EB06-
7A5FFCA9109C3FC4E2B0CCA304ADC981
Accept-Encoding: gzip, x-aes-cbc
X-Apple-Store-Front: 143443
Host: ax.phobos.apple.com.edgesuite.net
```

More information about personal behavior is uncovered by the association of the Apple-id, iMix (favorites hit list), sent and received coupons and pocket-money. From the frequency and type of communication iTunes may learn a lot about its customers, not only as an anonymous customer, but as a real person in that iTunes knows name, address etc.

E-mail addresses are certainly personal data. Even if they may provide a certain degree of pseudonymity, the iTunes system can map it to the real person by means of its customer record exactly.

A clearly hidden interface is the e-mail address of the customer encoded in every product he has purchased. With this information every reader of the product knows its origin.

5. Conclusion: recommendations for a better practice

To complete the picture, other DRM system must be analyzed by the same structure. This would allow to compare the systems in a fair and transparent way. This has been done in the study (Bizer et al. 2005), which compares iTunes with Microsoft's WMRM in Musicload (T-Online), the OpenMG in Sony Connect-Europe, and the alternative PotatoSystem. As a result, the state-of-the-art of DRM systems can be described as follows:

They all collect more personal data from their customers than necessary to fulfill the purchase service. There are many hidden interfaces, both by encoding personal data within the products, and by linking clickstream data with contractual data.

It must be noted, that the online shops utilize a good part of their knowledge about their customers for service improvement or extra features to the benefit of the users, like the assessment of more or less purchased or visited products. But there is no need to link this information to the real users and their personal behavior like favorite lists or personal relationships to other users.

There are two important parameters which govern business, not only, but also in the electronic world: Trust and reputation. Hidden interfaces and encoded personal data demonstrate that the shops do not trust their customers. The second line of defense is prosecution of customers who use the products illegally by copying them to other users. But by encoding personal data within the products, *all* customers are put under suspicion. If this is done secretly (intransparent to the users), customers lose trust to their shops, and the shop will lose reputation. And this will reduce the business of a shop beyond its market potential. This is the situation as is.

It is recommended to do all personal data processing in a clear transparent mode. Customers, who are ready to pay for what they get, are ready to provide personal data if they know what it is good for. If it is really for the benefit of the customers, they would accept it. If not, these data shouldn't be used.

6. References

[1] Becker, Eberhard; Buhse, Wilms; Günnewig, Dirk; Rump, Nils (2003): Digital Rights Management. Technological, Economic, Legal and Political Aspects. Springer Lecture Notes 2770, Springer, Berlin etc. 2003.

[2] Bizer, J.; Grimm, R. Will, Andreas; Möller, J.; Puchta, S.; Müller, A.; Müller, M. (2005): Privacy4DRM – Datenschutzverträgliches und nutzungsfreundliches Digital Rights Management. Study for the Ministry of Education and Research (BMBF) of the Federal Republic of Germany. July 2005.

[3] Brandenburg, K.; Stoll, G. (1992): The ISO/MPEG-1 Audio Codec: A Generic Standard for Coding of High Quality Digital Audio. Journal of the AES, Oktober 1994, 780-792. First publication at AES-Convention, Vienna 1992.

[4] EU (1995): Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data (www.cdt.org/privacy/eudirective/EU_Directive_.html).

[5] Grimm, R., Aichroth, P.: Privacy Protection for Signed Media Files: A Separation-of-Duty Approach to the Lightweight DRM (LWDRM) System. Proc. ACM Multimedia and Security Workshop 2004, Magdeburg, Germany, September 2004.

[6] iTunes Music Store (last download at 8.7.2005): www.apple.com/de/itunes, www.apple.com/de/itunes/store, www.apple.com/de/legal/privacy/statements.

[7] MP3 (1992/94): MPEG-1: ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 11172-3, Coding of moving pictures and associated audio for digital storage media at up to about 1.5 MBit/s. Part-3: Audio. 1992. And MPEG-2: ISO/IEC JTC1/SC29/WG11 (MPEG), International Standard ISO/IEC 13818-3, Generic coding of moving pictures and associated audio. Part 3: Audio. 1994. („MP3" is „MPEG-1 Audio" + „MPEG-2 Audio low sampling rates").

[8] Musicload (2005): The Musicload Factsheet of 05.04.2005 (www.musicload.de) with reference to the 2004 Digital Music Report of the International Federation of the Phonographic Industry (IFPI).

[9] Rosenblatt, B.; Trippe, B.; Mooney, S. (2002): Digital Rights Management – Business and Technology. M&T Books, Hungry Minds Inc., New York, 2002, 288 pages.

[10] US Department of Commerce (2005): Safe Harbor. http://www.export.gov/safeharbor/ (last update 3/2/05).

[11] Will, A. (2005): An economic analysis of music download platforms. Contribution to Axmedis 2005, Virtual Goods Workshop.

iRights.info: The need for reliable and trustworthy consumer information after copyright revision in Germany

Matthias Spielkamp
iRights.info
ms@iRights.info

Abstract

The rising use of file-sharing networks by users exchanging copyrighted works is one of the reasons for an increased potential for conflict between rights holders and consumers of copyrighted works. The author describes how this has been reflected in recent changes in copyright law in Germany as a result of international conventions. He analyses the current legal situation and comes to the conclusion that the situation is confusing for the majority of users of digital media, and even for many legal professionals. He then gives an example of the campaign the German film industry adopted in order to deal with this confusing legal situation, which can be described as an attempt to criminalize potential customers and spread uncertaint. In closing, the author contrasts this with the approach adopted by the site iRights.info, which aims at informing users in a neutral and balanced fashion about their rights regarding copyrighted works in digital form.

1. Introduction

In spite of the recent US Supreme Court decision in MGM vs. Grokster [1], file sharing is not likely to decrease in the near future.[2] According to a recent study by Cambridge, England, company Cache Logic, up to just over 80 percent of internet backbone traffic in some parts of the World is P2P-traffic – ranging from just over 50 per cent in Europe to just over 80 per cent in Asia.[3] The market research agency Jupiter Research of Darien, CT, interviewed executives of Europe's leading broadband providers and a quarter of those questioned said that more than 75 per cent of their subscribers used file sharing networks at least once a month.[4] Forrester Research of Cambridge, MA, says that around 35 million online consumers in Europe have downloaded music files from file-sharing services, amounting to about a third of all Europeans online.[5] At one point, file-sharing software Kazaa was the most downloaded program on the Internet – the software was downloaded approximately 230 million times until May 23, 2003, claiming top spot from the ICQ communication software [6]; in the week ending February 13, 2005, there were six file-sharing clients among the top 15 downloads at Download.com, one of the most popular software download sites worldwide.[7]

Some studies say file sharing does not influence album sales in any way [8], but others claim that the music industry lost 20 percent in sales in recent years [9]. At least that is the assertion of the music industry says so, at least [10], and the film industry started saying the same about their movie ticket and DVD sales. How this correlates with the fact that since 1999 – the year the DVD first played a major role in the sale and rental of movies – revenue from these sales and rentals (on DVD and video tape) in Germany has more than doubled, remains to be explained by the industry. According to a press release by Bundesverband Audiovisuelle Medien (German Association of audio-visual media), revenue rose from 860 million Euros in 1999 to 1.747 million Euros in 2004. The press release further states: "The market for rentals, heavily shaken by movie piracy, shows a rise in revenue by 1,4 per cent to 306,4 million Euros, and a rise in rental transactions of 2,3 per cent to 116,2 million rentals – the first increase after the slump in recent years."[11]

There is no clear evidence either way, but what is apparent is that rights holders and their associations have stepped up their lobbying efforts to change laws, and these changes lead to a growing confusion among citizens/consumers concerning what they are allowed to do on the Internet. This is one major reason for the conception of iRights.info.

In this paper, I will first present feasible examples of consumer activity and how this activity can be assessed with regards to the specifics of file-sharing regulation under German copyright law (Urheberrecht). I will then present a description and analysis of a campaign launched by a consortium of German film

industry associations, meant to deter people from "pirating" movies. Concluding, I will characterise the intention and aim of iRights.info in this context.

2. File-sharing and the law in Germany

Many uses of file-sharing networks are completely legal. Some people know this, some may take it for granted, but to some people this will sound rather surprising. Reading the newspapers or watching TV, they can certainly get the impression that everything that has to do with file-sharing is so called "illegal piracy". Not so.

Sharing your own works – texts, music, pictures, videos, software, games, animations and so on – is completely legal. Or, to be more specific: It is legal to share works if the person sharing them holds the rights to these works. If a programmer sells the rights to a piece of software he coded to Apple Computer Inc., he does not hold the rights to the program any more – so if Apple says he is not allowed to put this piece of software on a file-sharing network, then he is not allowed to do so. But if someone holds the rights to his or her work, they can do with it whatever they like.

For example, more and more companies put out files to share as well: music for promotional purposes, movie trailers and the like.

In addition to works someone owns, sharing is allowed for works the copyright holder allows to be shared – this sounds obvious, but one has to be aware that the rights holder must specifically assign those rights. This is done quite often, though, i.e. with works under Creative Commons licences [12], the GPL (GNU General Public Licence) [13] and many others.

Then there are works in the public domain (in Germany, there is no such thing as a public domain in the specific US sense. But what is described here is applicable to works that are "gemeinfrei", meaning copyright protection has run out for these works). An example for this is the Project Gutenberg [14], where scholars, students, and activists digitize classical texts from Shakespeare to Schiller and make them available in a searchable data base.

Most uses that are actually practiced on today's filesharing networks are illegal, though. The vast majority of music, films, software, texts are copyrighted and the rights holders prohibit sharing. This is made clear in most cases by copyright notices ("all rights reserved", the copyright-© etc.). It is important to point out that there is no obligation to use a copyright notice – on the contrary, unless the rights holder specifically assigns rights, users have "none" – pertaining to distribution, which is the case looked at here. Of course users have the rights to use the work – read it, listen to it, quote from it. But these rights are different from a right to publish, or to make available, for example on a file-sharing network.

Since the so called "first basket" (first round) of the German copyright revision came into force in September 2003 [15], it is illegal for individuals to make available works in a file-sharing network without holding the the rights to them – which is the majority of works on file-sharing networks today. So most of the actual uploading being done is clearly illegal under German law.

Downloading is a different matter, though. If a user in Germany downloads a song from a file-sharing network, it is seen as a duplication – a copy of the song. If this copy is for private use, it is perfectly legal – like copying a CD or a video tape. This permission is granted by an exception to copyright ("Schrankenregelung"), resembling – but not equalling – the fair use provision in US copyright law. Of course it is not allowed to sell or lend this copy, because then it would be a commercial use, which is prohibited.

Here things get complicated. Copying for private use is only allowed if the original is lawful; if the work from which the copy is made is itself "evidently an unlawful copy", it is prohibited. But how can someone tell whether it is evident that this work found on the file-sharing network was produced unlawfully?

This is a tough question. Imagine you find a copy of the movie "Independence Day" on the file sharing network Kazaa and decide to download it. Is this lawful?

It might well be. It has been shown on TV in Germany. So someone might have recorded the TV broadcast on his PC and converted the recording into a digital file. With this he is making a copy for private use, which is perfectly lawful.

But if he put the file on a file-sharing network, he would clearly be breaking the law because he doesn't have the right to distribute the movie, or to make it available.

But someone downloading the file would not be breaking the law, because it was not evident that the copy that was made available was produced illegally. It was illegal to make it available, but the subsequent copying of the file is legal.

Confusing? It gets better.

Now imagine someone finds a copy of "The Aviator" on a file-sharing network. Is it legal to download

it? As we have seen, it would be, if it were not obvious that the copy found on the network was produced an illegally. But is it obvious that it is a copy produced illegally?

To answer this question, one has to be able to answer the following questions:

Has the movie in question been broadcast on TV?

Answer: Probably not, it just came out in Germany, it is a big production and in cinemas at that moment.

Has it been released for home viewing?

Answer: This is difficult to determine. It is a pretty new movie. But who knows, US movies often come out in the US long before they come out in Europe. (for example the drama "House of Sand and Fog", which was released in the US on December 26, 2003, came to the theatres in Germany on February 17, 2005 – more than a year later. At the time the movie was still showing in German theatres the DVD was already available in the US, where it was released March 30, 2005.[16]) And if the person planning to download the movie lives in a small city with only one cinema, then she is familiar with the situation that movies come out a lot later there than in Berlin, Madrid, or London. So if it came out in the US a year ago already, it might have been released for home viewing in the US a while ago.

Therefore someone could have bought the DVD of the movie, made a private copy of it and put it on the file-sharing network – this way it would be legal to download it.

But what if the DVD is copy-protected?

Because of anti-circumvention legislation, it may be illegal to make a copy, even for private use.

In the US, article 1201 of the Digital Millennium Copyright Act ("Circumvention of copyright protection systems") states in paragraph (a) ("Violations regarding circumvention of technological measures") that "No person shall circumvent a technological measure that effectively controls access to a work protected under this title."[17] In Germany, a comparable rule applies, which is laid down in article 95a UrhG ("Schutz technischer Maßnahmen").[18] Both provisions stem from Article 11 of the WIPO Copyright Treaty, adopted December 20, 1996 by the Diplomatic Conference on Certain Copyright and Neighbouring Rights Questions: "Obligations concerning Technological Measures – Contracting Parties shall provide adequate legal protection and effective legal remedies against the circumvention of effective technological measures that are used by authors in connection with the exercise of their rights under this Treaty or the Berne Convention and that restrict acts, in respect of their works, which are not authorized by the authors concerned or permitted by law."[19]

But for one, all these laws are very complicated to understand and interpret, even for legal professionals[1]. Additionally, how would a downloader know whether "The Aviator" is copy-protected or not? In our sample case, he does not even know whether it has been released on DVD yet. Furthermore, it could have been released on videotape, which means it would be copy-protected in the US by the Macrovision Video Copy Protection system [20] – but in many cases not in Europe.

So after exhaustive and careful deliberation the user decides to download the movie.

By doing this, he committed a crime – at least that is what the rights holders say.

Because "The Aviator" has not been released for home viewing to date, the file on the file-sharing network has to be a copy someone made with his video camera in a cinema, and therefore illegal.

So the user has not only waited for hours for an abysmally bad and grainy copy of "The Aviator" to download onto his PC, he also has the studios demanding damages.

Now he consults a lawyer. The lawyer knows the law and asks him: Was it obvious that the file was illegitimate? The answer is: There is no way to be sure.

Whether something is obvious to someone is a very subjective category – and there are no guidelines, because there have not been any court cases dealing with this question in Germany yet. So until someone who has been indicted by a rights holder has the money and persistence to go to court against the movie or music industry, we won't know the answer. And that could take a couple of years.

In case the German ministry of justice is successful with a proposal brought forward as part of the "second basket" of copyright revision, we will know a little earlier. This proposal calls for a ban on the download of works in case it is obvious that they are made available illegally. What does this mean? The subjective category "offensichtlich" ("obviously" or "evidently") would still be part of the wording, but the law would – in the opinion of the ministry of justice – still be a lot clearer than the current version. Remember: It is not allowed to make available a copy of "Independence Day", even in case the copy itself is

[1] for reference see the Electronic Frontier Foundation's collection of various disputed cases: "Unintended Consequences: Five Years under the DMCA" at http://www.eff.org/IP/DMCA/unintended_consequences.php

a legal private copy. So the rationale of the ministry of justice would have to be that it is "more obvious" (or obvious to a greater number of people) that it is illegal to distribute a copy of "Independence Day" than it is obvious whether the copy itself is a legal (private) copy. If this were the case, then it must be more obvious (or obvious to a greater number of people) that it would be illegal to download it from a file sharing network. Whether this rationale rings true to the parents of a 12-year-old who downloaded a bunch of movies – readily available on the net – to his mom's computer, is at least debatable.

Situations similar to that of the legal uncertainty described in the case of file-sharing abound. Examples include the unresolved legal implications for private customers in Germany, the US and elsewhere, when buying music from Internet music store Allofmp3 in Russia.[2]

3. The film industry's campaign

Because the legal situation is as complicated as it is, iRights.info's creators see the dissemination of information concerning the consequences of acts deemed illegal as a major objective. In doing this, the iRights.info team has to take into account the approaches of other players in this field.

A while ago, a public relations agency, commissioned by a variety of associations of the German movie industry, started a campaign called "Hart, aber gerecht" ("firm but just").[21] They put up posters in cinemas, in trains, above urinals and other public places. They also show short promotional films – infomercials – in cinemas. In one of them [22] the viewer sees two young men in prison, looking insecure and afraid – obviously it is their first time in prison – being led to their cell by a guard. The camera cuts to two old inmates leaning against a handrail, watching them. The first one says to the other: "Mmh. Pirates!" Then he licks his lips and the two young inmates are led past them on their way to the cell. The other old inmate looks after them and says: "Even more pirates." Now the first one says: "Yes, but mine has the firmer ass" ("Ja, aber meiner hat den geileren Arsch"). Then the viewer sees the insert "Hart, aber gerecht" ("Tough, but just"), the picture fades to black and a text is shown, accompanied by a voice-over reading the text, which says: "Since September 2003, pirates go to jail for five years."

This clip is debatable on a lot of levels. Whether it is a good idea to convince people not to engage in file sharing by threatening them with the prospect of being raped in jail is one thing. Readers can judge for themselves. The focus here is on the legal aspects.

The slogan "Since September 2003, pirates go to jail for five years", which appears in all of the campaign's clips and on the posters, can be interpreted as a deliberate inaccuracy. The five year jail term for "pirates" ("Raubkopierer") has not only not changed with the revision of copyright legislation in 2003. It was five years before, it is five years now. It also only applies in case of commercial distribution or sales of illegitimate copies of works ("gewerblich"). This is clearly not the case with the majority of people who engage in file-sharing.

Still the film industry's clip intends to give viewers the impression that they will go to jail for five years if they are caught downloading a song or movie from a file-sharing network.

This example shows that the film indutry has no interest in informing people in an unhurried and dispassioned manner about the possible consequences of their offences. Instead, the big budget campaigns designed by lobbying groups more often than not disseminate dubious and misleading information in order to misinform and scare users.

4. iRights.info's approach

Of course it is important to tell users that, yes, it can have bitter consequences if you break the law and are caught. Damages can be very high, and no one wants to have to go up against a major label in court.

But at the same time, copyright information has to distance itself from the malicious propaganda of campaigns like the one described above.

This conviction was the rationale behind the project proposal made to the German Ministry for Consumer Protection, Food, and Agriculture (BMVEL) by "mikro – Verein zur Pflege von Medienkulturen e.V.", a non-profit organization based in Berlin, dedicated to fostering media culture. The proposal, developed by iRights.info's project lead Dr. Volker Grassmuck of Humboldt University's Hermann von Helmholtz-Zentrum für Kulturtechnik (HZK), was accepted in the summer of 2004; the project itself runs from September 2004 until March 2005.

How can iRights.info achieve the goal outlined above? One way is by employing a style of writing as neutral as possible, giving correct and trustworthy

[2] for an evaluation see: MP3 site settles for $10 million with RIAA, CNET News.com, October 25, 2004, http://news.com.com/MP3+site+settles+for+10+million+with+RIAA/2100-1027_3-5425885.html, and accompanying discussion forum threads

information and withholding opinion. One of the editors of iRights.info, Till Kreutzer, is one of the most renowned copyright experts in Germany. He is frequently invited to the deliberations of the justice ministry and is one of the co-founders of the "Institute for legal questions of free and open source software" ("Institut für Rechtsfragen der Freien und Open Source Software (IFROSS)").[23] He checks and validates each and every one of the texts that appear on the site, so users can be sure to find correct and trustworthy information regarding the legal interpretation of law and court decisions.

Another way to achieve this balance is by pointing out alternatives. In most cases, be it with texts, music, films, or software, there are licenses authors can employ to make their works available to others in a much easier way than before. One well-known example for programmers and software engineers are copyleft licenses like the GNU General Public License (GPL). There are many others in that field[3].

But open content or copyleft licenses do not only work for programmers. Content producers, like artists, musicians and writers, can as well can use licenses to make content more freely available. Creative Commons licences are one example[4].

These alternative licences do not imply that people have to give up their copyright. On the contrary, licenses like the GPL or Creative Commons are founded on copyright, they could not exist without a copyright system. They allow for a differentiated set of permissions: for example, letting other people know that they are allowed to use a picture under certain circumstances like putting it on their webiste for non-commercial use or allowing to make derivatives – remixes and covers – of songs. But the author can, for example, determine: you can use it as it is, but you must not change it. Many producers could use this to promote their works, and many already do. So, if a lot of people use these licenses, there would be a much bigger pool of works to choose from when creating new creative works.

That is the reason why iRights.info showcases examples like this on the site as well – to call attention to alternatives.

iRights.info is structured into three parts. The first section of the site is called "Kopieren" ("copy") and explains what users can do with copyrighted works in their daily life – what are the implications of file sharing, is it allowed to make copies from CDs and DVDs, is it legal to digitally record radio or TV broadcasts and so on. The second is called "Selbermachen" ("do it yourself") and provides advice on how copyright applies when someone creates his or her own work, be it software, text, audio, video and more. The third section gives "Hintergrund" ("background") information on copyright: where does it come from historically, how did the current legislation come about, what does it mean for business and culture, what will happen in the foreseeable future.

iRights.info's approach is three-fold in order to cover a wide range of issues important to the general public targeted by the site. First of all, the editorial staff consists of four people from diverse backgrounds – law, multimedia art, politics, and journalism. This is meant to provide for a diversity of issues being included in the site. Secondly, iRights.info has an advisory board made up of experts on different fields of expertise, ranging from activists to some of the most renowned experts on copyright law in Germany (among them the main commentator of copyright law, Thomas Dreier). They take an active role in advising the editors on what issues to include and whether the topics covered are dealt with in a appropriate and correct fashion.

The field of copyright has an extraordinary wide range, though. Therefore it is important for iRights.info to receive feedback from users on all sorts of levels: does the site cover all the issues important to users? Are these issues explained in a way a legal layperson can understand? Is the information useful in dayly life – using and "consuming" works on the one hand, and creating them on the other hand? So thirdly, to receive as many comments and remarks on these questions, a discussion board is an integral part of the site. Here, users can comment on articles, suggest further topics they would like to see covered, and interact with other users discussing questions they share.

With the combination of these channels of information and communication, iRights.info's creators and editors hope to be able to achieve the set goal: to meet the need for reliable and trustworthy consumer information after – and during an ongoing – copyright revision in Germany.

10. References

[1] Supreme Court of the United States: "Metro-Goldwyn-Mayer Studios Inc. et al. v. Grokster, Ltd., et al.",

[3] for a collection of these different licenses, please consult the IFROSS License center at http://ifross.de/ifross_html/lizenzcenter-en.html (German version at http://ifross.de/ifross_html/lizenzcenter.html)

[4] others can be found at http://en.wikipedia.org/wiki/Open_content or the IFROSS Open content section: http://ifross.de/ifross_html/opencontent.html (in German only)

June 27, 2005, accessed September 2, 2005 http://www.eff.org/IP/P2P/MGM_v_Grokster/04-480.pdf

[2] A. Gilbert, "Ruling won't slow file swapping, experts say", *Cnet news.com*, June 27, 2005, http://news.com.com/Ruling+wont+slow+file+swapping%2C+experts+say/2100-1030_3-5764998.html, accessed September 2, 2005

[3] CacheLogic: The True Picture of Peer-to-Peer Filesharing, July 2004, http://www.cachelogic.com/research/slide12.php, accessed September 2, 2005

[4] Jupitermedia: File sharing is no longer just a problem for record labels – broadband service providers are being swamped with file-sharing traffic, press release, February 2003, http://www.jupitermedia.com/corporate/releases/03.02.10-filesharing.html, accessed September 2, 2005

[5] ClickZ Marketing Solutions: Downloading music files a hit with Europeans, October 2002, http://www.nua.ie/surveys/index.cgi?f=VS&art_id=905358424&rel=true, accessed September 2, 2005

[6] S. Neuhaus, "KaZaA erzielt Download-Rekord", *Heise Online*, May 23, 2003, http://www.heise.de/newsticker/meldung/37062, accessed September 2, 2005

[7] "Most Popular in Windows"-section, *Cnet download.com*, http://www.download.com/3101-2001-0-1.html?tag=pop, accessed September 2, 2005

[8] F. Oberholzer-Gee and K. Strumpf, "The Effect of File Sharing on Record Sales: An Empirical Analysis", http://www.nber.org/~confer/2004/URCs04/felix.pdf, accessed September 2, 2005

[9] R. Rob and J. Waldfogel, "Piracy on the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students", NBER Working Paper No. 10874, November 2004, http://www.nber.org/papers/w10874, accessed September 2, 2005

[10] IFPI: Fact sheet - Facts on File-sharing, August 2004, http://www.ifpi.org/site-content/press/20041007c.html, accessed September 2, 2005

[11] Bundesverband audiovisuelle Medien: DVD-/Videogesamtmarkt konsolidiert im ersten Halbjahr 2005 auf hohem Niveau, http://www.bvv-medien.de/aktuell.html, accessed September 2, 2005

[12] Creative Commons Website, http://creativecommons.org, accessed September 2, 2005

[13] Free Software Fourndation: GNU General Public License, http://www.gnu.org/copyleft/gpl.html, accessed September 2, 2005

[14] Project Gutenberg, http://www.gutenberg.org/, German version at http://gutenberg.spiegel.de, accessed September 2, 2005

[15] Bundesministerium der Justiz: "Neues Urheberrecht ab morgen in Kraft", http://www.kopien-brauchen-originale.de/enid/5j.html, text of law available at http://217.160.60.235/BGBL/bgbl1f/bgbl103s1774.pdf, accessed September 2, 2005

[16] "House of Sand and Fog"-website, http://www.dreamworks.com/dvd_features_hosaf.html, accessed September 2, 2005

[17] U.S. House of Representatives, "Digital Millennium Copyright Act", October 8, 1998, http://www.copyright.gov/legislation/hr2281.pdf, accessed September 2, 2005

[16] Bundesministerium der Justiz, "UrhG § 95a – Schutz technischer Maßnahmen", http://bundesrecht.juris.de/bundesrecht/urhg/__95a.html, accessed September 2, 2005

[18] World Intellectual Property Organisation, WIPO Copyright Treaty, December 20, 1996, http://www.wipo.int/documents/en/diplconf/distrib/94dc.htm, accessed September 2, 2005

[19] Wikipedia: Macrovision, http://en.wikipedia.org/wiki/Macrovision, accessed September 2, 2005

[21] ZKM Zukunft Kino Marketing GmbH: "Hart Aber Gerecht"-website, http://www.hartabergerecht.de, accessed September 2, 2005

[22] available (MPEG, 7,5 MB) at http://www.edelman-newsroom.de/cgi-bin/WebObjects/app.woa/1/wa/Nav/showAsset?oid=711&cid=2&wosid=ltpyFahHs8FUZaxtBme2G0, accessed May 21, 2005

[23] Institut für Rechtsfragen der Freien und Open Source Software, http://ifross.de, accessed September 2, 2005

Justice Psychology Meets Digital Rights Dilemma

Nicola Döring
Ilmenau University of Technology
nicola.doering@tu-ilmenau.de

Franziska Fellenberg
Ilmenau University of Technology
franziska.fellenberg@tu-ilmenau.de

Abstract

The Digital Rights Dilemma addresses the conflict of interest between providers and consumers of virtual goods. The former are mainly interested in ensuring their intellectual property rights and copyrights, and try to enforce usage rules with the help of Digital Rights Management (DRM) Systems. The later are interested in unconstrained access to virtual goods and often make illegal copies (Digital Piracy) and refuse DRM measures. This paper argues that, based on theories and empirical findings of both Justice and Moral Psychology, digital piracy could be reduced and DRM acceptance increased if the consumers' subjective experiences of morality and justice would be taken into consideration. Implications for further research, design and marketing of DRM are discussed.

1. Virtual Goods and Digital Rights

Unlike physical goods, v*irtual goods,* such as digital music and videos, electronic books, computer games or software, can be reproduced and distributed requiring almost no time, effort or money.

1.1. The Digital Rights Dilemma

Unconstrained access to virtual goods, including unlimited reproduction and distribution, is in the customer's interest. For the copyright holders and providers of virtual goods, however, it is clearly preferable to restrict access to digital goods in order to ensure their eventual purchase by the respective users. After all, an important motivation for the production and maintenance of high-quality informational goods is the possibility to generate profit. This seminal conflict of interest between rights-holders and customers regarding access to copyrighted virtual goods is denominated the *Digital Rights Dilemma.*

In this conflict of interest, both sides represent extreme positions: many users demand a completely free flow of information ("information anarchism"), whereas a number of rights-holders tend toward a highly restrictive policy of technical access ("information feudalism"). In many commercial sectors, the trading of virtual goods is both legally and technically restricted (e.g. the prevention of illegal copies through mechanisms of copy protection devices) and/or controlled (e.g. the detection of illegal copies and thus possible prosecution) through the application of DRM or *Digital Rights Management*. Consumers, on the other hand, go to great lengths—both individually and collectively—to circumvent respective measures and to illegally appropriate virtual goods (*Digital Piracy* [1]).

1.2. Psychological Aspects of Digital Rights Management

This paper looks at DRM and digital piracy from a psychological perspective. Concretely, we focus on the experiences and behavior of consumers. Taking into account a number of psychological aspects, we have chosen to concentrate on *fairness and moral perceptions of consumers*:

- To what extent do consumers of virtual goods describe the DRM strategies and the prices of virtual goods to be fair or unfair?
- To what extent do consumers consider the illegal elusion of DRM measures and the unpaid appropriation of virtual goods as morally justifiable?

Alongside other components, we can assume that assessments of justice help to determine the acceptance level of DRM's. An analysis of a fairness assessment can practically serve to:

- Modify DRM strategies with respect to those characteristics which are considered to be especially unfair.
- Communicate DRM strategies in such a manner so that they have a positive influence on perceptions of fairness and moral assessments.

A basic supposition of psychology consists in the fact that human actions derive from a variety of motivations, of which self-interest plays a key role. On the other hand, there also exist numerous examples of peo-

ple who act out of altruism, meaning in the interest of others. Developmental psychologist Lerner [2] was the first psychologist to defend the notion that the *production of justice must be seen as an independent motive of human beings*. For example, a person feels guilty if she acts contrary to her sense of justice [3]. As a result, privileges are not always experienced positively, but rather can very well produce negative feelings if one determines that those priviliges are unjust [4]. A great deal of studies has empirically proven that a person's actions (also) derive from the production of justice, without exclusively seeking individual gain in the process.

According to this line of thought, copyright protection and fair pricing of virtual goods should be a socially accepted concern which can nevertheless enter into conflict with the motives of self-interest. Accordingly, a poll surveying n=126 internet users in Germany demonstrated that 74 percent of those questioned recognized that illegal software copies represent a financial detriment to rights-holders [5]. At the same time, 75 percent admitted to the fact that they themselves were in possession of illegal copies. Those polled overwhelmingly (78%) demanded that the use of illegal copies for private use not be punishable. They found it to be evidently unjust to have to pay for digital goods, and saw their illegal behavior as morally justifiable, meaning they had no sense of guiltiness.

It is not the responsibility of psychologists to *normatively* assert how to solve the Digital Rights Dilemma in a just manner, and/or how to act morally in this context. In lieu thereof, psychologists proceed descriptively and reconstruct the subjective sense of justice and morals, and the behavioral consequences of those involved.

2. Justice Psychology and Digital Rights

Justice Psychology primarily deals with the question as to what human beings experience as just and unjust, and how they attempt to generate justice. The Digital Rights Dilemma constitutes a new and underresearched topic in Justice Psychology.

2.1. Concepts

The *concept of justice* addresses the process of ensuring the legitimate rights and entitlements of persons or groups: a) in decision-making and distributive processesses (*procedural justice*); and/or b) in terms of the outcomes of theses processes (*outcome justice*).

Individuals perceive unjustice, if:

- their legitimate entitlements are violated,
- it is possible to attribute cause or responsibility for the violation to a certain actor, and
- the actor does not or can not sufficiently justify his or her actions.

Both the justice motive and the sensitivity for unjustice are relatively stable individual traits which are distinctive according to each and every individual. Furthermore, justice perceptions are not only determined by traits but also by situative factors. The justification of pretensions are not only derived from the specific characteristics and behavior of each person individually (e.g. loyal customers expect to receive discounts), but are also based on social rules and norms and the comparison of similar situations.

In addition to the terms *just* and *justice,* the concepts of *fair* and *fairness* are often times applied. *Just* contains a stronger normative component and thus describes a sense of justice which is based on *principles*. The term *fair*, on the other hand, deals more strongly with the aspect of social exchange, as a result placing much more emphasis on *interpersonal relationships*. Within justice psychology, "fair" and "just" are frequently applied as synonyms.

Justice perceptions are relevant in different social contexts: within personal relationships (*interpersonal justice*, e.g. fair distribution of housework between spouses), within formal organizations (*organizational justice*, e.g. fair treatment of employees, just wages) and within society in general (*social justice*, e.g. a just tax system). From the consumer's point of view, the Digital Rights Dilemma is embedded in the customer-seller-relationship.

2.2. Criteria and Methods

Consumers can assess the fairness of DRM and prices of virtual goods based on criteria of procedural and outcome justice:

- *Procedural Justice*: The individual criteria on which a just decision-making process is based are multiple, and include such *normative principles* as freedom from discrimination. Additionally, an important criterion is the possibility that the affected persons have the ability to influence the process (*participatory decision making*). Alone the possibility to express an opinion in the course of a decision-making process helps the parties to accept the future results, even when the end-result may be unfavorable [6]. Another criterion of procedural justice is access to all relevant information and justification of the decision-making process (*informational justice*). Even if individuals are not able to actively participate, they will still evaluate the decision-making process as more fair after having received exhaustive information: for example, if wage cuts were well explained this would mean a significant decrease in the levels of theft and

layoffs as it would when the applied cutbacks were to be explained in a succinct sentence which only refers to the financial necessities of the business [7]. Perceptions of *procedural injustice* are related to processes that violate social norms and laws, that do not permit the participation of all relevant parties, and that are intransparent to observers. The procedural justice of legal and technical DRM decision is regarded as limited because they are dominated by content providers while at the same time consumer organizations only play a marginal rule.

- *Outcome Justice*: The concept of outcome justice suggests that fairness occurs when all parties involved in an exchange process share the same ratio of profits to investments. That means that individuals expect rewards proportional to their investments. At the same time—as part of the perception of a fair outcome—they also take into consideration the entitlements and rewards of the other exchange parties (*dual entitlement*). Buyers and sellers believe that they are respectively entitled to a reference price and reference profit [8]. If either party does not receive its entitlement, the relationship will be perceived as unfair. For its part, price fairness psychologically addresses the balancing of both entitlements. Specifically, dual entitlement suggests that perceived unfairness results when a reference price is increased to such an extent that a company increases its profit. An increased price is perceived as fair when it maintains the company's existing level of profit. Also the theory suggests that people reflect on the respective seller's motives and reputation when they judge the fairness of pricing policies [9]. Until now, the perceptions of different consumer groups regarding the subjective fairness of prices and profits of different virtual goods have not yet empirically investigated.

The *research designs and methods of justice psychology* are as diverse as its topics. In general, justice perceptions are measured through standardized questionnaires: The respondents are asked to rate the fairness of a certain situation, outcome or decision-making process on five-point to seven-point rating scales. Additionally, cognitive, emotional and behavior reactions regarding justice (e.g. satisfaction, happiness, trust) or injustice (e.g. anger, depression, revenge) are operationalized through appropriate questionnaire or interview items. On the other hand, manifest behavioral patterns are primarily researched in experimental studies, whereby a group of participants receives shared rewards and needs to distribute them in a fair manner between the single group members. All these methods are applicable to DRM scenarios.

2.3. Just DRM and Pricing

The concepts outlined above provide us with a basis from which to understand issues of just DRM and just pricing of virtual goods.

2.3.1. Customer-Seller-Relationship. The willingness of consumers to accept DRM and to pay for virtual goods depends on the relationship to the seller. The content provider must be accepted by the consumer as an exchange partner with her or his own legitimate entitlements. Virtual piracy is largely related to ignoring the right-holders as human beings and as real victims, perceiving them instead as, "nonhuman, remote, oblivious entities" [10].

In the context of digital goods, a central hypothesis consists in the fact that the lack of physical or visual presence of sellers, other customers, etc.—including the lack of physical contact with these people—all help to explain why the individual places his or her own perception in the foreground. As a result, one's own interests are over-valued, while at the same time undervaluing the interests of the others. One possible countermeasure could be that of a vivid representation of the rights-holder on the online-shopping platform (e.g. giving the right-holders a face through the presentation of employees' photographs).

Customers would more easily accept the providers' entitlements if the providers themselves would acknowledge the consumers interests (e.g. by open information politics), and would show prosocial engagement (like charity activities).

2.3.2. Rational Justification of DRM. Contrary to the sense of procedural justice, if customers are to be forced to use DRM, then this decision should be well justified in order to reduce the perception of unfairness. Frequently, the use of DRM's are *legally* justified. Since copyright laws are both complex and rarely understood by consumers, it is a very weak line of argumentation. Moreover, it conceals a criminalization of the user.

Seen from a perspective of justice psychology, it would be better to employ *economic* arguments and rationally justify the use of DRM by highlighting the negative consequences of virtual piracy. These should be made more public and transparent [1] for the different goods (e.g. music, e-books) and parties (e.g. rights-holders, editorial companies and other intermediaries), respectively. By providing examples of the financial consequences for the different branches and virtual goods, the actual damages provoked by digital piracy could be made more visible. Moreover, by employing such a rational basis, it would then be clear if and to what extent the possession of illegal private copies

does or does not represent a minor offense. Subsequently, the complications caused for the customer by DRM systems may then seem to be more fair when the justification of the measures is made comprehensible.

2.3.3. Improving DRM Characteristics. In order for a customer to classify the DRM measures as fair, her/his justified interests of security, interoperability, and usability must be met. More often than not, this is not the case.

The more DRM's are unsecure and time-consuming, many customers will come to the conclusion that it is an act of unjustice to have to accept disadvantages which line the pockets of the provider.

Thus, as seen from a justice psychology perspective, the existing DRM's will have to be improved or alternative models employed [11] in order to increase customer acceptance.

2.3.4. Explaining DRM Characteristics. As long as problems of security and interoperability exist and customers are not included into the decision-making and designing process, at the very least an proactive informational policy is a basic requirement: the limitations and problems of the DRM's should be prominently displayed on the websites of online shops and be made generally understandable for all (informational fairness).

2.3.5. Compensating for DRM Burden. As a result of the complexity and risks of the DRM's, attractive rewards should be made available for the customers. This could be in the form of discounts or value-added services.

2.3.6. Rational Justification of Pricing. Since virtual goods are not tangible items and practically free to reproduce, their value is often times subjectively seen as minimal. Accordingly, it would be seen as unfair to have to pay high prices for bits and bytes. Empirical studies demonstrate that customers generally have no idea as to the costs of production and distribution [10]. Therefore, further investigations should be carried out as to how costs of copyrights, of production and of marketing of virtual goods can be made more transparent for many people. To further explore consumers' notions of fair pricing of virtual goods comparative studies including similar material goods could be of interest. To illustrate the empirical approach a selection of exemplary questionnaire items is presented:

Sample Item 1 on Buying versus Renting Virtual Goods:

Please use the 7-point rating scale to indicate your justice evaluation of the pricing of the following music services:

- Select from more than one million songs. Purchase and download the music you want for 0,99 EUR per song. Keep it forever.
- Select from more than one million songs. Purchase and download all the music you want for 9,99 EUR per month. Keep it as long as you are a member.

7-point rating scale to measure justice perceptions:
1 = not at all fair
2 = unfair
3 = rather unfair
4 = middle rate
5 = pretty fair
6 = fair
7 = very fair

Sample Item 2 on Fair Pricing of Virtual Goods:
What do you think?

- How much of the price of a music album *on CD* usually *goes* to the artist? ____Percent
- How much of the price of a music album *on CD should* go to the artist? ____Percent
- How much of the price of an *Online* album usually *goes* to the artist? ____Percent
- How much of the price of an *online* album *should* go to the artist? ____Percent

Sample Item 3 on Fair Pricing of Virtual Goods:
Please insert the pricing scheme you feel is most fair. Think of a recently published music album:

- The album *on CD* (new) for ____EUR
- The album on CD (used) for ____EUR
- The album *Online* for ____EUR

By realistically and transparently presenting each and every service offered by the rights-holders and service providers, the sense of regarding the price formation of apparently free virtual goods can be increased.

Nevertheless, seen in terms of prices, it is not just the customer's consciousness about the investments of the service provider which is incorporated into the dual-entitlement principle. The reference price of the consumer should also be taken into account, which is usually derived from prices of alternative products: for example, the official download-price for a twenty-page online article from a scientific magazine would be considered as fair when it more or less coincides with the cost and time-consumption of the consumer related to a respective paper copy.

Apparently, DRM acceptance is closely related to fair pricing. Until now, there has been no public rational discourse regarding adequate pricing for virtual goods. As can in the following example of the online-music store iTunes, fairness is a promise: "In all, the iTunes Music Store offers music that's fair to you, fair to artists and easy to enjoy." Whether or not this self-appraisal coincides with the fairness assessment of potential customers should be examined empirically.

3. Moral Psychology

Whereas philosophy and theology focus on the issue of how human beings should act, moral psychology is more interested in the question how people react in different situations, and how they morally justify their behavior or the behavior of others. Moral refers to the question of right or wrong. It is a concept more general than justice: justice is considered to be morally positive and injustice as morally negative.

3.1. Moral Dilemmas

The classical method of moral psychology consists in presenting people with stories of moral dilemmas. A moral dilemma consists in a problem in which there is no clear right or wrong resolution: The more or less legitimate interests of different parties stand against each other. The subjects are asked to name and justify a course of action which in their opinion is morally correct. On the basis of this justification, the subject's moral beliefs are infered. One such example is the so called Heinz dilemma [12].

The Heinz Dilemma

In Europe, a woman was near death from a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging 10 times what the drug cost him to make. He paid $400 for the Radium and charged $4,000 for a small dose of the drug. The sick woman's husband, Heinz, went to everyone he knew and tried every legal means to borrow the money, but he could only get together about $2,000, which is half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said: " No, I discovered the drug and I`m going to make money from it." So having tried every legal means, Heinz gets desperate and considers breaking into the man's store to steal the drug for his wife.

- *Should Heinz steal the drug?*
- *Why or why not?*

The method of the moral dilemma situation was developed by the moral researcher Lawrence Kohlberg. Kohlberg [13] investigated the development of moral judgment and conceived a three stage model of moral development which even today exercises an important influence in the field:

1) *Preconventional Level / Personal Interest Schema.* Here, moral decisions are based on one's own personal interests and necessities as well as the approval of people who are close to us.
2) *Conventional Level / Maintaining Norms Schema*. One orientates him- or herself on the laws and rules of larger systems, such as the state, religious communities, etc.
3) *Postconventional Level / Postconventional Schema.* At this level, norms and laws are understood to form part of a societal contract, and the people orientate themselves around attempts to find the best possible solution for all those involved.

Whereas Kohlberg assumed that there was a progressive stage model, Neo-Kohlbergians describe different equivalent schemas of argumentation which determine the moral judgment of people depending on each particular situation [14]. The preferred method of the Neo-Kohlbergian-approach is the Defining Issues Test (DIT) [15]. Starting from a moral dilemma situation, the subject is given multiple-choice items. They are then asked to assess these statements according to their importance for the decision that was made. On the basis of the assessments, the subject's moral schema is derived.

3.2. Moral Framing of DRM

On the basis of the until now more limited studies (primarily carried out with student samples), we find that the consumers of virtual goods are almost exclusively motivated by their own interests. Within the context of the *Personal Interest Schema* or the *Preconventional Level of Morality*, the illegal appropriation of digital goods is not understood to be a moral dilemma at all, but instead as non-detrimental normality, because the other side—embodied in people who have just but converging interests—is blinded out as such [16]. Meanwhile, the orientation on the personal interest schema is strengthened when the same stance overwhelmingly exists within the social environment itself, be it among students at the university [10] or in peer-to-peer networks [17]. Within the framework of the *norm-schema*, rarely do students judge digital piracy to be amoral in relation to copyrights rights and right holders. In part, postconventional arguments are employed, whereby the free flow of information is presented as a contribution to the common welfare of everybody. Nevertheless, it is likely just a line of argumentation used to rationalize the defense of one's own interests [10].

Concretizing the Digital Rights Dilemma in terms of moral dilemmas in order to explore the moral understanding of consumers and promote a public moral discussion offers us very promising research and application opportunities for the future.

Sure enough, it will be difficult to operate with existential and dramatic scenes such as in the Heinz dilemma. Dilemma situations relating to digital rights could variate employing the following aspects:

- the reason for illegal appropriation of virtual goods (high school student not fully aware of copyright issues; university student on low budget in need for an expensive software tool to finish her dissertation thesis),
- type of visible right owner (global player or small software firm),
- price of virtual good etc.

4. Conclusion

At least among certain populations, digital piracy is extremely normalized. If the claims of the rights-holders and service providers are considered to be fully inadequate, respective DRM measures appear to be unfair from the very beginning. Subsequently, a fundamental psychological intervention consists in first making the Digital Rights Dilemma visible in the eyes of the consumers as a moral dilemma in the first place.

Systematic studies of consumers' justice perceptions of single attributes of DRM's or price schemes for virtual goods are lacking. Nevertheless, several practical proposals can be derived from previous investigations. These not only deal with the technical and economic design of DRM's, but primarily address the relevant communication processes between both sellers and consumers of virtual goods as well as the broad public.

The concepts, findings, and methods of justice and moral psychology could prove to be very fruitful for addressing the Digital Rights Dilemma. In the future, the theoretically deduced proposals for the design and marketing of DRM measures still need to be tested and further refined. The social psychological perspective of justice and morality should always be seen in the context of other factors influencing DRM acceptance.

5. References

[1] J. Gantz and J. Rochester, Pirates of the Digital Millenium. Financial Times Prentice Hall, 2004.

[2] M.J., Lerner, "The justice motive. Some hypotheses as to its origins and forms", *Journal of Personality, 45*, Blackwell Publishing, Farmington, 1977, pp. 1-32.

[3] L. Montada, "Justice, Equity, and Fairness in Human Relations", In: T. Millon & M.J. Lerner, Eds., *Handbook of Psychology*, Vol. 5 , Wiley, New Jersey, 2003, pp. 537-568.

[4] T.R. Tyler, R.J. Boeckman, H.J. Smith, and Y.A. Hugo. *Social Justice in a diverse society*, Westview Press, Boulder, CO, 1997.

[5] K. Böhle, About the mind-set of software pirates. *Editorial of INDICARE Monitor, 1, 8,* 2005.

[6] J. Greenberg and R., Folger, "Procedural Justice, participation, and the fair process effect in groups and organizations." In P. Paulus, Ed., Basic group process, Springer, New York, 1983, pp. 235-266.

[7] J. Greenberg, "Stealing in the Name of Justice: Informational and Interpersonal Moderators of Theft Reactions to Underpayment Inequity", *Organizational Behavior and Human Decision Processes, 54,* Elsevier Science, San Diego, 1993, pp. 81-103.

[8] J.L. Cox, "Can different prices be fair?", *Journal of Product and Brand Management, 10, 5*, Emerald, Bradford, 2001, pp. 264-275.

[9] S. Spiekermann, "Individual Price Discriminaton in E-Commerce – An impossibility?" 2005 Online Document: http://www.wiwi.hu-berlin.de/~sspiek/Pricing_IEEE.pdf.

[10] S. Hinduja, „Trends and Patterns among Software Pirates", *Ethics and Information Technology, 5*, Kluwer Academic Publisher, Netherlands, 2003, pp. 49-61.

[11] R. Grimm and J. Nützel, Potato System and Signed Media Format – an Alternative Approach to Online Music Business. *Proceedings of the Third International Conference on WEB Delivery of Music* 2003, pp. 23-26. IEEE Computer Society

[12] A. Colby and L. Kohlberg, *The Measurement of Moral Judgment.* Vol. 2: Standard Issue Scoring Manual. Cambridge University Press, Cambridge, 1987b, reprinted 1990.

[13] L. Kohlberg, "Moral stages and moralization: The cognitive developmental approach", In T. Lickona, Ed., *Moral Development and Behavior,* Holt, Rinehart and Winston, New York, 1976, pp. 31-53.

[14] J. Rest, D. Narvaez, M. Bebeau, and S. Thoma,. A Neo-Kohlbergian Approach: The DIT and Schema Theory. *Educational Psychological Review, 11, 4*,1999, pp. 291-324.

[15] J. Rest, D., Narvaez, M. Bebeau, and S. Thoma, *Exploring Moral Judgement: A Technical Manual for the Defining Issue Test*. Center for the Study of Ethical Development, 1999, Order, Leeds, UK.

[16] R.B. Kini, H.V. Ramakrishna and B.S. Vijayaraman, "Shaping of Moral Intensity Regarding Software Piracy: A Comparison Between Thailand and U.S. Students", *Journal of Business Ethics, Volume 49, 1*, Springer, New York, 2004, pp. 91-104.

[17] J.S. Svensson and F. Bannister, “Pirates, sharks and moral crusaders: Social control in peer–to–peer networks “, *First Monday, 9, 6*, University of Illinois, 2004, Online Document: http://firstmonday.org/issues/issue9_6/svensson/

Automatic Image Theft Detection in eBay by Digital Watermarking

Martin Steinebach, Ellen Kremer, Lucilla Croce Ferri
Fraunhofer Integrated Publication and Information Systems Institute (IPSI), Dolivostr. 15, D-64293 Darmstadt
{Martin.Steinebach}, {Ellen.Kremer}, {Lucilla.Croce-Ferri}@ipsi.fraunhofer.de

Abstract

Digital images are used in the Internet for a broad range of applications. One well known example for image usage are product photographs in eBay auctions. In recent times the misuse of these images is often discussed and reported. Either images are re-used from third party auctions or they are copied from other web sites like for example online catalogues. As eBay offers a high number of auctions, image theft often stays unnoticed. We introduce an automated method for image theft detection using digital watermarking and an eBay online interface. Images are scanned based on product description filters, downloaded and scanned for embedded watermarks.

1. Motivation

Online auctions are very popular today. eBay is the best known representative for this business type. Many people use digital images to show potential bidders how the products they offer look like. Figure 1 provides an example.

Various reasons may lead a seller to a situation where he would like to provide an image, but has no access to one. Examples are lack of a digital camera, difficulties in providing professional images or a fraud where the seller actually does not own the product he is offering. In all cases sellers tend to use images already existing on the Internet, perpetrating a misrepresentation fraud. They can originate from other eBay auctions or from online shop catalogues where the products are sold. Either the image is copied by the seller to a web storage to which he refers or he simply uses the URL of the original copy of the image and places it into his offer. We call a seller using copied images "pirate seller" for the rest of this article. Both methods are copyright violations against the original creators of the digital images, leading to complains like this:

"someone named as […] stole my pictures and now he is selling the same items as mine. what will i do for him to stop copying my picture?"[1]

But it can be assumed that many copyright violations pass unnoticed by the original owners of the images due to the vast number of auctions hosted at eBay. Still owners of web shops using high quality image material to advertise for their products are looking for an efficient way to identify copyright violations. An automated system would be necessary to be able to scan the huge amount of images in an acceptable amount of time.

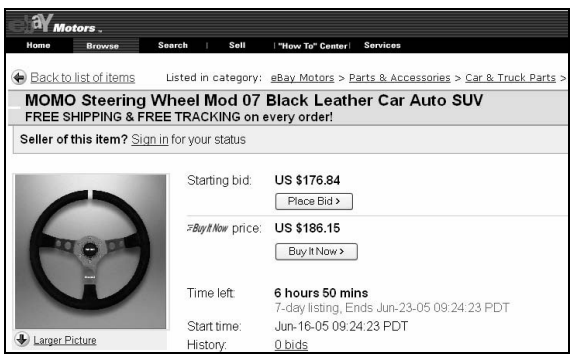


Figure 1. Example eBay auction with image

[1] http://forums.ebay.ph/thread.jspa?threadID=300000585&tstart=0&mod=1117000770127

When an image theft is noticed, various reactions are possible, depending on the method of theft.

When the pirate seller uses an URL pointing to the original image, known countermeasures include (a) changing the image or (b) changing the image location. In (a) the image could be changed into a sign indicating a copyright violation. Of course, the original owner of the referred image must change the URL of his image and his references to display the correct image. Method (b) is simpler and only leads to a broken link in the auction of the pirate seller.

But if a pirate seller copies the image to a web space the original owner has no access to, a third party must be called to handle the problem. In the case of eBay, a web form (see figure 2) for complains about pirate sellers is available. This leads into stopping the auctions of the pirate seller. Of course this method can also be chosen when the pirate seller is using only an URL to the original image[2].

As one can see, an owner of an image is not helpless against copyright violations of his images.

In this article, we address the part of automatically detecting image copyright violations in eBay using digital watermarking algorithms. In section 2 we introduce digital watermarking. In section 3 we introduce a concept using an eBay scanner and a watermarking detector. In section 4 we describe our implemented prototype with respect of reliability and performance.

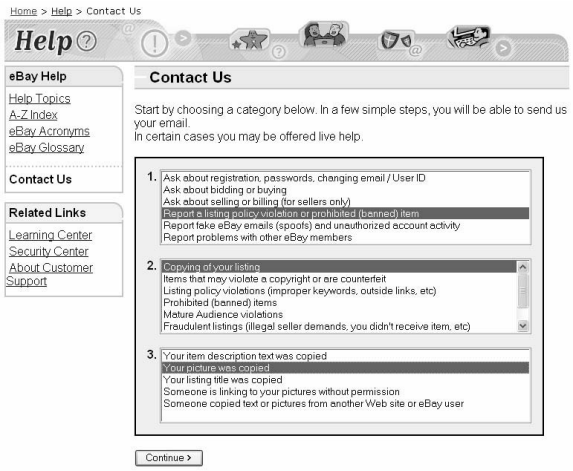


Figure 2. eBay image piracy notification form[3]

[2] http://cgi3.ebay.com/ws/eBayISAPI.dll?ViewUserPage&userid=net_enforcers_inc

[3] http://pages.ebay.com/help/contact_us/_base/index.html

2. Digital image watermarking

Digital watermarking ([1], [2]) invisibly embeds information into a cover with the help of a secret key. This information refers to the cover, and provides additional information about it accessible only by those who own the watermarking algorithm and the secret key. The most common application of this technology is copyright protection or customer tracing. Both can be seen as a simple stand-alone alternative to complex and more restrictive digital rights management (DRM) or easily inserted into existing security concepts.

One of the most important advantages of watermarking is the fact that content can escape from DRM environments based only on cryptography and access control. This would render it unprotected. One example for this is when the content is transmitted via an analogue channel. A well-designed digital watermark, on the other hand, stays in the content even after printing and scanning or manipulations like strong JPEG compression. So when watermarked content is found in an illegal environment, copyright claims can be proven or original customers can be identified.

For this reason, the challenge in DRM is mainly to keep material in a protected environment, while in watermarking one needs to find watermarked content which is used illegally. This makes efficient search strategies an important aspect of digital watermarking.

2.1. Image watermarking state of the art

A large number of watermarking algorithms has been proposed in the recent years. Most of them deal explicitly with still images and share common approaches.

Some basic watermarking requirements can be identified independently from the various applications. These requirements are related mainly to the perceptual transparency after the watermarking embedding, to the watermarking capacity, i.e. the quantity of information that can be embedded into the data, the security of the watermarking technique and the watermarking robustness against common processing techniques or intentional manipulations of the data. Even if transparency, capacity and robustness are trade-off parameters, for images the robustness represents the most challenging parameter. The geometrical transformations of images caused by printing and scanning are very challenging processes for watermarks search mechanisms. They can cause

serious robustness problems to many watermarking algorithms not explicitly designed to survive them. General methods to achieve high robustness against these transformations are based on resynchronization techniques, such as the usage of the original image for non-blind detection methods, registration patterns or extraction of characteristic feature points found in the original image for blind ones.

A local exhaustive search mechanism can be necessary in combination of the previous mentioned methods. Other possibities are invariant watermarks, which remain unchanged under the considered geometrical transformation and autocorrelation techniques for period watermarks. [3]

Geometrical transformations are not the only manipulations that can corrupt the embedded watermarks. Also lossy compression, noise and luminance changes can modify the images in such a way that the watermarks cannot be detected anymore or only partially. The approached used in this case to achieve the required robustness are based mainly on the redundant embedding into perceptually significant parts of the image. Spread spectrum techniques [4] are used for this purpose in the spatial and frequency domain.

2.2. Search strategies

Different commercial applications and services based on web-crawling already exist, that search the Internet for watermarked data. Perhaps the most known one is the service offered by Digimark[4], but also other companies are providing commercial watermarking and searching services, promising secure online distribution of images with usage tracking[5]. Not only images need to be protected. Audio mp3 data are particularly vulnerable and new business models based on their individual fingerprinting are adopted by various publishing companies [5].

A generic prototype for crawling peer to peer services for watermarked data was introduced in [6]. It is based on the Gnutella architecture.

3. Concept

In this section we describe the general idea of our watermarking approach for eBay.

The section discusses the different stages and the necessary components. The actual implementation of these is described in the next section.

Our approach is to watermark images used by original owners and search for misuse of these marked images in eBay using keywords to identify fitting auctions. A complete process including fraud detection would feature the following steps:

1. **Original image is watermarked by its owner.** Only the watermarked copy is used in the public. The unmarked copy is stored in a secure place or deleted. The embedded message identifies the original owner. The secret key used in the embedding process is stored.

2. **Marked image is put in the public.** This can either be an eBay auction, an online catalogue or even a printed catalogue if the watermark is robust against printing and scanning, Thereby the image is made accessible for potential pirates.

3. **Image is included in scanner list.** The original owner informs the online eBay scanner that it should search for misuse of this image. He provides the secret key which is necessary to retrieve the watermark and a list of keywords which describe the product to be seen on the image.

4. **Scanning starts.** The scanner is now continuously looking for auctions with fitting keywords. If such an auction is found and it features an image, it is downloaded and the secret key is used to check if the original owners watermark is present. If the original owner uses his image in an eBay auction, this process will also find his own auction showing the successful operation of the scanner.

5. **Pirate seller places auction using stolen image.** Now the pirate seller copies the marked image into his eBay auction. He describes the product he wants to sell and uses some of the keywords the original owner submitted to the scanner.

6. **Scanner detects misuse.** The auction of the pirate seller is found by the scanner using the keywords. The image is downloaded and the watermark of the original owner is detected.

7. **Alert.** The scanner now informs the original owner about the misuse, sending a copy of the auction including the image to him. The original owner can now ask eBay to shut down the

4 http://www.digimarc.com/
5 http://www.bluespike.com/giovanni.html, http://www.alphatecltd.com/watermarking/eikonamark/eikonamark.html

auction. As an alternative, the scanner could also inform eBay automatically about the misuse.

3.1. Example

A seller wants to start an auction of a steering wheel. He takes a photograph of the product to sell and embeds a watermark consisting of his name and using his secret key into it. The marked copy is then used in the eBay auction illustrating the product to be sold. Figure 3 shows the usage of the image.

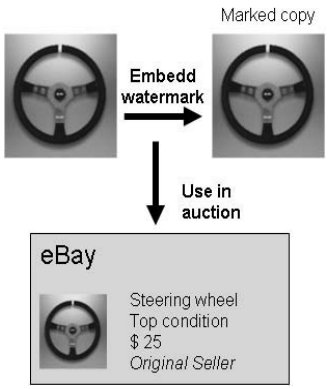


Figure 3. Image of steering wheel is marked and used in an auction

Now a pirate seller also wants to sell a similar steering wheel. Instead of taking an own photograph, he is looking for other auctions selling steering wheels in eBay. He chooses the image of the original seller and copies the image to his own web space. Then he uses the marked copy in his own auction. As seen in figure 4, he offers a "top steering wheel" which is "used, as new".

But the original seller is aware of potential image theft and therefore sends his secret key and a list of keywords to an eBay scanner. The scanner now checks all images in auctions of the transportation domain featuring the keywords "car wheel" and "steering wheel" as shown in figure 5. Over time, it finds three images. One is from the auction of the original seller and a watermark is detected in it. The auction is then listed in a report sent from the scanner to the original seller. The next image is the stolen image of the pirate seller where the watermark is also found in. The occurrence of the image is also noted in the report. The third image is a different image of a steering wheel. No watermark of the original seller is detected here. The scanning is repeated for a given amount of time specified by the original owner. Images or auctions already checked are stored in a data base to prevent repeated downloading and detection.

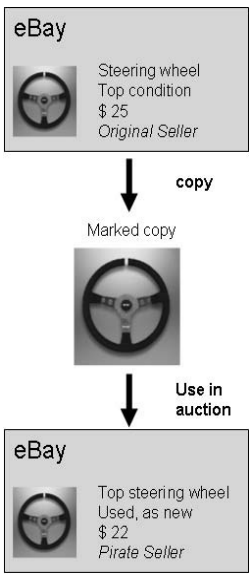


Figure 4. The image is copied and misused by a pirate seller

The original seller regularly receives the report of the scanner. If any other than his own auctions use his images, he can either react on his own or address eBay to stop the pirate seller auction. In either way, he does not need to scan for auctions misusing his images on his own. Any auction in the transportation domain selling steering wheels or car wheels will be scanned for his image automatically for a period defined by him.

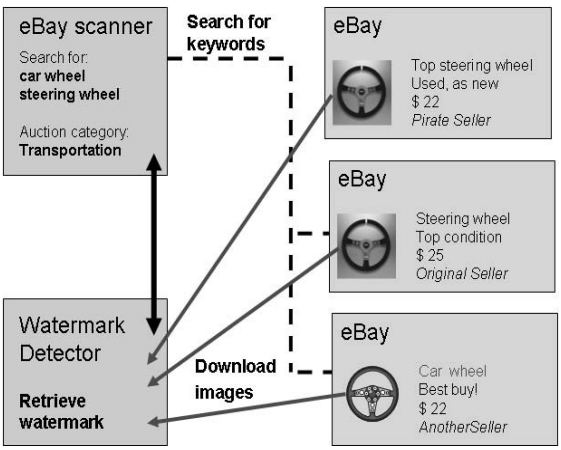


Figure 5. The eBay scanner is searching for fitting keywords and finds the auction of the pirate seller

4. Implementation and test results

In this section we describe our prototypic implementation based on the concept introduced in the previous section. We also provide test results and identify possible bottlenecks in a commercial application.

4.1. System design

Our scanner uses the eBay product search feature to search image files. It accesses the eBay platform via Internet, utilizing an API provided by eBay. This API enables to access eBay via web services (Figure 6).

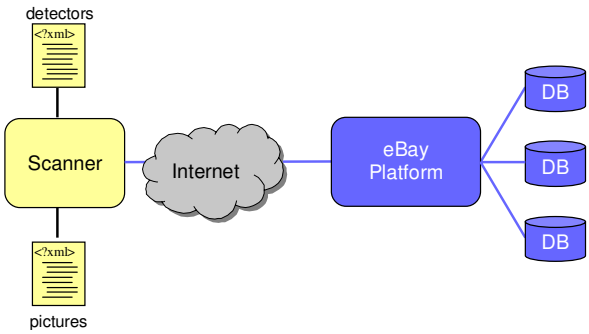


Figure 6. eBay Image Scanner communicates via internet with the eBay platform

The information whether an already scanned image is watermarked or not, is stored in an xml file in order to avoid multiple scanning of images.

Following steps are performed by the system to process a search query:

1. Submit keywords given by the user to the product search offered by eBay.
2. Identify the product offers containing an image.
3. Download the images from this selected product list and calculate their hash value
4. Check if the hash value already exists in the xml file. In this case, check if the image was watermarked.
5. Otherwise, check if the image is watermarked and store its hash value.
6. Display the results of these steps to the user as shown in Figure 7.

The result of the search query contains the URLs of the image and the offer, together with the offer and seller IDs. Furthermore the country, where the eBay offer originates, is also displayed (Figure 7).

4.2. Implemented functionalities

The graphical interface for the scanner (Figure 8) offers the possibility to use advanced search functionalities for the keywords in the offer title. The user can specify or exclude selected words or all the listed words as criteria for the search. Also portions of sentences can be used and the search can be performed not only in the title but also in the complete offer text.

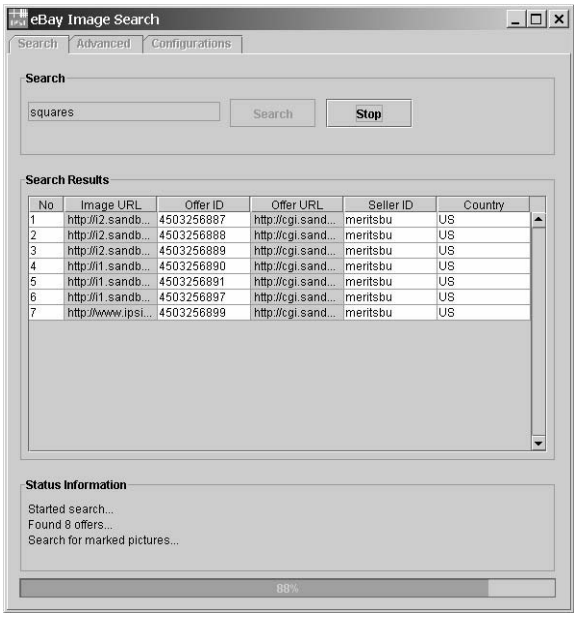


Figure 7. Example search in the eBay image scanner

Furthermore the user can decide to search only selected eBay categories, the eBay stores or offers from a specific date.

The eBay Image Scanner allows using different watermarking detection algorithms. They are shown in the "configuration" page. The user can decide which one is used for the current search.

The current advanced prototype can be enhanced in the future particularly in relation to the configuration and documentation facilities and its user friendliness.

A possible extension of the searching filters is considered and a more structurable display of the results. Other new possible functionalities are related to an automatic notification of the results to the user or directly to the eBay supporters. In this last case, mechanisms should be added to definitively ensure the fraudulence of the founded watermarked images.

Another important issue is the amount of information stored in the result list. At the moment,

the user receives only a list of possible illegal usages of a specified image. For forensic applications, it would be necessary to collect all information about the suspected offer and a screenshot that can be used as fraud evidence.

To enhance user friendliness, user profiles can be supported, containing the different configuration settings.

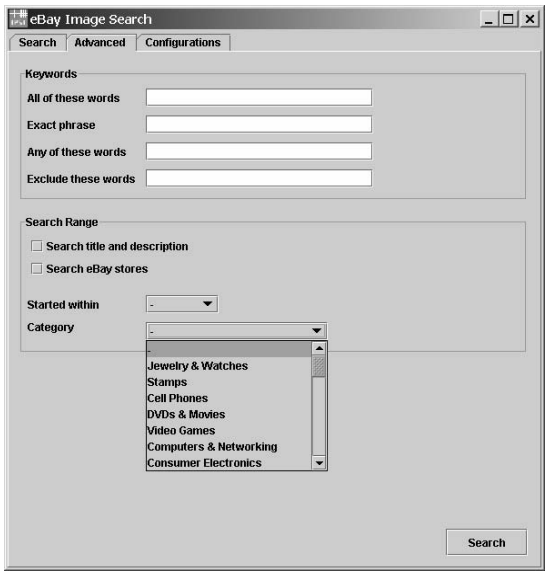


Figure 8. Advances search options of the eBay image scanner

To enhance user friendliness, user profiles can be supported, containing the different configuration settings.

4.3. Performance

Two types of performances are important for our eBay Image Scanner. The first one is related to the false and positive detection errors of the watermarking algorithm and the second one to the time performance of the whole system.

Compared to other existing commercial crawling tools, for example the Digimarc image tracking service[6], the Image Scanner is not limited to a specific watermarking algorithm. Different methods can be linked to the Image Scanner thought a well defined interface specification.

[6] http://www.digimarc.com/products/imagebridge/MarcSpider/default.asp

Following requirements have to be satisfied by the watermarking methods, to optimize the system performance:

- Blind detection, i.e. the detection algorithm does not utilize the original image to extract the watermark. This is, of course, a mandatory requirement.
- High robustness against scaling and cropping, since pirate images can be a slightly different version of the legal ones.
- High robustness against luminance and compression changes for the same reason.
- Low complexity, at least for the detection process, in order to reduce the detection time. This will play a fundamental role in minimizing the processing time of the whole scanning process.

4.4. Possible attacks

If the watermarking algorithm is available to everyone, attacks trying to overwrite or destroy the watermark are possible [7]. The robustness and security of the watermarking algorithm have to be proved critically, before the method is registered by the scanner.

It has to be pointed out, that the knowledge necessary to perform complicated attacks are normally behind the possibilities of most of normal eBay users.

Also secure protocols for the transmission of the secret key needed during the detection process have to be utilized. If a pirate has access to the secret key, he could try to generate his own watermark with the purpose to replace the original one.

It could be also necessary to make the scanner anonymous and to mask its IP address, to avoid misleading it. A professional attacker could monitor the searching activities of the scanner, temporarily substitutes the illegal used images with some other images thereby hide his illegal usage of the images.

Possible technical interferences of the scanner with the auction functionalities are not an issue, since the scanner acts like a normal user, looking for a specific product. Still an explicit cooperation with eBay would be desirable, may in the form of an additional security service offered by eBay.

5. Discussion and future work

In this section we briefly describe the advantages of our approach for different potential groups of users and the planned extensions to the prototype, in order to enhance its usability.

5.1. Potential users

The main benefits of our automatic eBay Image Scanner are related to the possibility of monitoring copyright infringements for images. This is an open issue in particular for professional online catalogues, online image archives or professional eBay sellers, whose images are often stolen and somewhere else illegally used. With the eBay Image Scanner, the catalogue operators would be able to track these abuses, having a mean to demonstrate their legal position. Since the watermarking algorithm has to be designed to be robust against geometrical attacks, images need to be watermarked only once. The legal user can publish them in different contexts, in different Web pages with different formats, without the necessity to mark all the different versions of the images.

The deterrent effect would produce also an advantage for eBay itself, since it would reduce the number of reclamations that they have to process, as answer to stolen victims' protests. But, of course, also users and new potential sellers would appreciate the enhanced security against image misrepresentations. This could contribute to increase the transactions' volume and to open new online markets for high valued images and photographs.

Another application for the eBay Image Scanner is the online brand monitoring. Services based on the search of illegal usages of marked images in eBay could be offered to assist marketing and branding professionals to protect to their brands.

5.2. Extension

The current advanced prototype can be enhanced in the future particularly in relation to the configuration and documentation facilities and its user friendliness.

A possible extension of the searching filters is considered and a more structured display of the results. Other new possible functionalities are related to an automatic notification of the results to the user or directly to the eBay supporters. In this last case, mechanisms should be added to definitively ensure the fraudulence of the founded watermarked images.

Another important point is about the protocols of the results. At the moment, the user receives only a list of possible illegal usages of a specified image. For forensic applications, it would be necessary to collect all information about the suspected offer and a screenshot that can be used as fraud evidence.

About enhanced user friendliness, user profiles can be supported, containing the different configuration settings.

6. Summary and conclusion

We proposed an automated method for image theft detection based on digital watermarking and providing an eBay online interface. Images are scanned using product description filters, downloaded and scanned for embedded watermarks. The listed results can be analysed by the legal image copyright holder, who decides the countermeasures against the pirates. It is important to point out that, in order to ensure the optimal image scanner performance, all online existing images have to be watermarked before their publication, otherwise the attacker could utilize image versions which are not protected and these would not be found by the scanner.

Another strategy used to discourage image theft is the embedding of a visible watermark, such as a company logo, or the website URL, or any other copyright text into the images.[7]

The most important advantage of our approach is that the high quality of the published images is not damaged by the visible mark. Furthermore, in most cases, the visible mark could also be easily cropped out, while the invisible watermark can be detected by the scanner also after cropping transformations.

Not only the sellers of eBay actions would benefit from the image scanner services, but also eBay itself should have interest in supporting and offering them to its users.

Acknowledgements

The work described in this paper has been supported in part by the European Commission through the IST Programme under Contract IST-2002-507932 ECRYPT. The information in this document reflects only the author's views, is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.

[7] http://www.vendio.com/my/ihost/promo_wm.html

References

[1] I. Cox, M. Miller, and J. Bloom, “Digital Watermarking”, San Francisco, CA: Morgan Kaufmann, ISBN 1-55860-714-5, (2002)

[2] J. Dittmann, "Digitale Wasserzeichen" Springer Verlag, Berlin/Heidelberg, ISBN 3-540-66661-3, (2000)

[3] P. Bas, JM. Chassery and B. Macq, “Geometrically Invariant Watermarking Using Feature Points”, IEEE Trans. Image Processing, Vol. 11, Nr. 9, pp. 1014-1028, (2002)

[4] Cox, I., Kilian, J., Leighton, T., Shamoon, T., "Secure Spread Spectrum Watermarking for Multimedia", IEEE Transactions on Image Processing, Vol.6, Nr.12, pp.1673-1687,(1997)

[5] Steinebach, Dittmann: "Design Principles for Active Audio and Video Fingerprinting, Multimedia Security", Chapter V, Idea Group Publishing, Chun-Shien Lu (Hrsg.), ISBN 1-59140-275-1,pp 157-172, 2004

[6] Steinebach, Dittmann, Lang: "Konzepte zur Vermeidung oder Verfolgung von Urheberrechtsverletzungen in Netzwerken auf der Basis digitaler Wasserzeichen", Competence in Content, Ralf Schmidt (Hrsg.), Tagungsband 25. Online-Tagung der DGI, S. 113 – 125, ISBN 3-925474-58-x, 2003

[7] Johnson, Duric, Jajodia, “Information Hiding: Stenography and Watermarking Attacks and Countermeasures”, Kluwer Academic Publishers, ISBN: 0-7923-7204-2, (2001)

Multi-Level Markets for Virtual Goods

Andreas U. Schmidt
Fraunhofer–Institute for Secure Information Technology SIT
Dolivostraße 15, 64293 Darmstadt, Germany
Andreas.U.Schmidt@sit.fraunhofer.de

Abstract

As an alternative to rigid DRM measures, multi-level or networked marketing of virtual goods has raised some interest. We report on a theoretical study of those markets which was hitherto lacking, and devise a generic, kinematic model for the monetary flow in them. Building on it, the incentives buyers receive through resales revenues and the competition of goods are examined. Some practical implications, in particular for the efficacy of multi-level markets for countering free-rider phenomena, are outlined.

1. Introduction

Information goods share the attributes of transferability and non-rivalry with public goods, and additionally are durable, i.e., show no wear out by usage or time [1]. Like with a private good, however, original creation can be costly, whereas reproduction and redistribution are cheap. This is the more true for *virtual goods* [2], i.e., information goods in intangible, digital form, which are distributed through electronic networks. Free-rider phenomena plague their creators and distributors, a problem which is conventionally approached using copy protection measures and/or digital rights management (DRM) systems. This practise has aroused public controversy and an ongoing discussion about the various fundamental [3] and economic [4] issues arising from it. The general legitimacy of DRM measures which tend to disrupt consumers' expectations on their individual usage of the good [6], seems doubtful in light of empirical findings on the effect of illegal file-sharing on record sales [7], which seems negligible. As an alternative to the protection of virtual goods by DRM, so called incentive management (IM) systems have recently emerged. They promise to yield a fair remuneration to the originator of the good, who may be identical with its creator or not, without necessitating copy protection or disruption of users' expectations on "fair" and "personal" uses. One of the first such systems, and one which is already in practical use is the so called Potato System [8, 9]. It is based on super distribution of the virtual good from buyer to buyer, whereby each buyer obtains, along with the good itself, the right to redistribute it on commission. Upon resale, she will obtain a share of the purchase price as an additional incentive. The rationale behind this kind of scheme, called here **multi-level IM (MLIM) systems**, is as obvious as appealing. Rather than to discourage illegal distribution of the good by more or less unpopular measures, the aim is to make legal distribution more attractive than "piracy". Concurrently, the scheme purports to attribute a fair remuneration to the party from which the good originated, for instance the creator of a work of which the virtual good is an embodiment.

The present report contributes a building block to the presently lacking study of MLIM in the framework of theoretical economy. Section 2 introduces a simple model for the monetary flux in a general multi-level market and derives the most basic results pertaining to it. The model is complemented by a dynamical model for the competition of two goods in such a market in Section 3. The results the competition model yields with respect to the free-rider problem and the competition between two goods are treated in 4. Section 5 offers a qualitative discussion of the issues raised in the preceding theoretical ones. It is argued that MLIM can be a fair scheme despite its similarity to illicit schemes. The free-rider problem presents itself as a genuine issue of information economy. We then offer some thoughts on the potential of MLIM to influence markets through determining the incentive via dynamical forward pricing, and outline the potential problem of market inhomogeneities. Section 6 concludes by noting some directions for further work. The full version of this abridged report, containing in particular proofs of all propositions, is found in [10], cf. also the related paper [11].

2. Monetary Flux Model

The model we devise is continuous and kinematic, i.e., all quantities are variables with continuous range and the model describes the monetary flux between the market play-

ers. Other relevant quantities, such as the expected resales revenue, are to be derived from the kinematics. About the market players no special assumptions are made. The model is thus neutral with respect to the detailed structure of the monopolist firm marketing the good and the consumer. Agents are solely discriminated by the time t at which they enter the market, i.e., buy the good form another agent. Consequently, buying the good occurs only once for per agent, while resale can happen to arbitrarily often, subsequently. The market in turn is assumed to be homogeneous, i.e., all agents have equal probability for mutual trade.

We let the number $n(t)$ of agents in the market at time t be an unspecified function with continuous, non-negative, finite or infinite range. The resales price at time t is denoted by $\pi(t)$. This fundamental price function π, as well as the market dynamics, is left completely unspecified and can be generated by any underlying mechanism without affecting any general results derived from the model. The expected (average) monetary incentive v_i for an agent entering the market at time t is given by

$$v_i(t) = v_r(t) - \pi(t), \tag{1}$$

i.e., the expected revenue v_r from resales to later market entrants, diminished by the price at which the good was bought. To calculate v_r, note that the influx of agents into the market is given by $\dot{n}(t') = \mathrm{d}n(t')/\mathrm{d}t'$ at any later time $t' > t$, and if the agent was alone then one could integrate $\pi(t')\dot{n}(t')$ over an interval to obtain the resale revenue accumulated in its duration. But since there is competition in the reseller market, and all $n(t')$ agents have equal probability to strike a deal with the newcomers, the integrand must be divided by $n(t')$. Thus

$$v_r(t) = \int_t^\infty \frac{\pi(t')}{n(t')}\dot{n}(t')\,\mathrm{d}t'. \tag{2}$$

Reparametrisation by the monotonously increasing number of agents $n(t)$, makes the independence of the market dynamics manifest and yields

$$v_r(n) = \int_n^{n_\infty} \frac{\pi(n')}{n'}\,\mathrm{d}n', \tag{3}$$

in which the market size n_∞ may be finite or infinite. However, it makes sense to specialise to finite markets, see [10], and we assume $n_\infty < \infty$. Then, a nonsingular reparametrisation can be applied, replacing n with the market saturation $s = n/n_\infty$, $0 \le s \le 1$. The integral operator $K\colon \pi \longmapsto v_i$, mapping price to incentive, is a Volterra operator of the second kind, given by

$$(K\pi)(s) \stackrel{\text{def}}{=} v_i(s) = \int_s^1 \frac{\pi(s')}{s'}\,\mathrm{d}s' - \pi(s). \tag{4}$$

As this operator describes a closed market, one would expect it to satisfy a conservation law. Here, this law takes the form of a game-theoretical zero-sum condition.

Proposition 2.1 (Zero-Sum Condition). *For bounded* π

$$\int_0^1 v_i(s)\,\mathrm{d}s = 0. \tag{5}$$

This condition expresses that wins and losses in incentive compensate each other. One important feature of the model is that the incentive is *scale-free*, i.e., does not depend on n_∞. For regular enough π, the inverse of K is obtained as a solution of the inhomogeneous equation $K\pi = v_i$. The derivatives of π, v_i, are denoted by $\dot{\pi}$, $\dot{v}_i$, respectively.

Proposition 2.2. *K maps $\mathscr{V} \stackrel{\text{def}}{=} C^1([0,1])$ bijectively onto*

$$\mathscr{W} \stackrel{\text{def}}{=} \Big\{ v_i \in C^1((0,1]) \;\Big|\; \int_0^1 v_i = 0,\ v_i = o(\tfrac{1}{s}),\ \textit{and}\ \dot{v}_i = O(\tfrac{1}{s})\ (s \to 0) \Big\}. \tag{6}$$

The inverse of $K\colon \mathscr{V} \longrightarrow \mathscr{W}$ is

$$(\check{K}v_i)(s) \stackrel{\text{def}}{=} -\frac{1}{s}\int_0^s \sigma\dot{v}_i(\sigma)\,\mathrm{d}\sigma. \tag{7}$$

Although nothing in principle prevents a forward monetary flow from earlier market entrants to later ones by negative prices $\pi < 0$, the more conventional case is that of positive resale prices. The necessary and sufficient condition for positive prices reads as follows.

Proposition 2.3. *Let $\pi \in C^1([0,1])$. Then, π is positive if and only if*

$$\frac{1}{s}\int_0^s v_i(\sigma)\,\mathrm{d}\sigma > v_i(s) \quad \textit{for all } s. \tag{8}$$

This result has a rather direct interpretation. It says that the monetary flow is always directed backwards if and only if the expected incentive at a certain time is smaller than the average expected incentive before that time.

The basic model can easily amended by further features. In particular it is desirable to take transaction costs and a commission into account. In the resale process, the buyer as well as the seller can incur transaction costs. We assume them to be constant. While the buyer's transaction cost $\beta \ge 0$ directly adds to the price $\pi(s)$ and can therefore be absorbed in it, the seller's transaction cost $\sigma \ge 0$ modifies the integrand for the calculation of v_r from $\pi(s)/s$ to $(\pi(s) - \sigma)/s$. Upon integration, this yields a negative contribution in the incentive of the form

$$v_i(s) = \int_s^1 \frac{\pi(s')}{s'}\,\mathrm{d}s' + \sigma\ln s - \pi(s). \tag{9}$$

If there is an entity, called the **collector**, which collects part of the resales revenue, e.g., to remunerate the creator of the good, and pays only part of it as a commission to resellers, the market turns into an open system. The **commission factor** $0 \le \gamma \le 1$ diminishes the revenue of a single

resale from π to $\gamma\pi$, and the modified operator K_γ yielding the incentive $v_{i,\gamma}$ becomes

$$(K_\gamma \pi)(s) = \int_s^1 \frac{\gamma(s')\pi(s')}{s'}\,\mathrm{d}s' - \pi(s). \qquad (10)$$

Its inverse for differentiable π can still be calculated and reduces, for constant commission, to

$$(\check{K}_{\gamma=\text{const.}} v_{i,\gamma})(s) = -\frac{1}{s^\gamma}\int_0^s \dot{v}_{i,\gamma}(\sigma)\sigma^\gamma\,\mathrm{d}\sigma. \qquad (11)$$

The market with commission no longer satisfies the zero-sum condition but rather an analogue balanced with the collector's share. Further details are found in [10, 11].

A continuous model is an idealisation of a realistic market where buyers enter one by one, i.e., the market size evolves in discrete steps. This entails artifacts, most notably the logarithmic singularity for $v_i(s)$ as $s \searrow 0$ when $\pi(0) > 0$, see Figure 1 a). Therefore one needs to examine the discrepancy between the incentive obtained from the continuous model and the one calculated by discrete summation somewhat more closely. For a constant price $\pi(s) = \pi$, the discrete model can be solved directly. Agents are labelled with $k = 1, \ldots, n_\infty$, by the order of market entrance, and this yields for the expected incentive $\overline{v}_i$ of the discrete case

$$\overline{v}_i = \pi\left(\sum_{k'=k+1}^{n_\infty} \frac{1}{k'-1} - 1\right) = \pi\left(\Psi(n_\infty) - \Psi(k) - 1\right), \quad (12)$$

where the Digamma function $\Psi(z) = \Gamma'(z)/\Gamma(z)$ is the logarithmic derivative of the Gamma function. In the general case, we have to look at the difference between $v_i(s)$ and the discrete incentive $\overline{v}_i(s \cdot n_\infty)$ at the corresponding point.

Proposition 2.4. *For bounded, non-negative π holds*

$$|v_i(s) - \overline{v}_i(sn_\infty)| \le \frac{\pi_{\max}}{2}\left[\frac{1+s}{sn_\infty} + O\left(\frac{1}{6}\frac{1+s^2}{(sn_\infty)^2}\right)\right], \quad (13)$$

with $\pi_{\max} \stackrel{\text{def}}{=} \max_{s\in[0,1]} \pi(s)$, and in which the term of order $(sn_\infty)^2$ is strictly dominated by the previous one.

Te error behaviour of the continuous model is rather benign in that it decays with the inverse of the market size at any finite saturation $s > 0$. For fixed $k = sn_\infty$ on the other hand, a constant error bounded by $c_k\pi_{\max}$ for some $c_k > 0$, will always remain.

It is to be expected that markets based on super distribution, in particular MLIM systems for virtual goods, are related to network externalities. They can be endogenously produced in those markets as well as influence the market's dynamical growth. Network effects are understood in the literature as the benefit that accrues to a user of a good or a service because he or she is one of the many who use it. Simple functional forms of network effects for special types of networks, e.g., telecommunication networks, such as Sarnoff's, Metcalfe's, and Reed's law, are often taken as heuristics to explain the dynamics of the growth of networks of the respective type. The most prominent phenomena traced back in this way to network effects are a "slow startup", the existence of a "critical mass", and strong growth after this mass has been reached. Models for network externalities and their effects on prices and utility are numerous, see [12] and references therein, where also possible functional forms of network externalities are discussed.

Network utility can spatially be understood as the *aggregate* value, summed over all members of the network, or as the *individual* value enjoyed by single members. In models depending on a dynamical parameter, each case is in turn subdivided on the temporal axis into the *dynamic* utility given as a function of the saturation s, as a relative variable, and the *kinematic* utility, which is the scaling behaviour of the utility with the market size n_∞. The only kinematic aggregate utility arising is that obtained by the replication of the good and redistribution of it, a contribution which is always of order $O(n_\infty)$, like in broadcast networks. The incentive contributes to aggregate utilities only in a dynamic way, since it is given by

$$n_\infty \cdot \int_0^s v_i(s')\,\mathrm{d}s', \qquad (14)$$

which approaches zero for $s \to 1$, respectively is of the order $O(-n_\infty)$, more precisely $-n_\infty \int_0^1 (\gamma(\sigma) - 1)\pi(\sigma)\mathrm{d}\sigma$ if a commission is in effect.

The only contribution to the dynamic, individual utility is v_i, since the kinematic, individual utility, i.e., the scaling behaviour of v_i with n_∞, is $O(1)$ precisely if π is $O(1)$ $(n_\infty \to \infty)$, i.e., if the price stays bounded. This *scale-freeness*, which becomes manifest in the continuous limit $n_\infty \to \infty$, is an essential property of the model presented. It is not an artifact of the continuous idealisation, since the error bound (13) shows that it is stable for nonzero s. However, for small, fixed $k = sn_\infty$, and if $\pi(0) > 0$, a scaling of the kinematic, individual utility of order $O(\ln n_\infty)$ appears (meaning that in pyramid schemes profiteers gains scale logarithmically with the number of participants). In conclusion, the incentive is the only network externality affecting the agents, except for a logarithmic, kinematic effect on early buyers. This was to be expected since the market described has no special structural properties.

Figure 1 a) shows the most basic example of resales revenues and incentives resulting from a constant price. It exhibits the logarithmic singularity present in the continuous model, and which will always emerge if $\pi(0)$ is positive. The singularity is avoided if $\pi(0) = 0$ as in b) and c). Additionally, in c) the incentive is forced to zero as $s \to 1$ by letting π approach zero, and also shows a case where v_i is not always monotonic decreasing and π is still positive. The effect of a commission factor is exhibited in Figure 1 d).

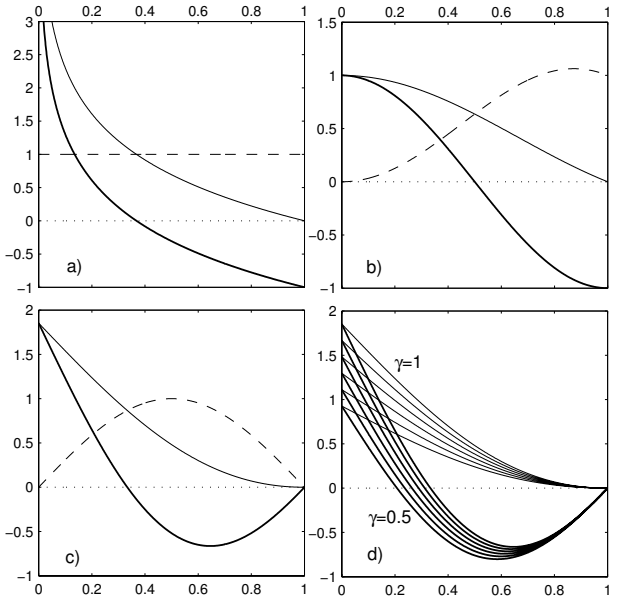


Figure 1. Examples for prices π (dashed), expected resales revenues v_r (thin solid), and incentives v_i (thick solid).

3. Competition Model

To devise a dynamical model for the competition of two goods, say A and B, in a multi-level market described by the model above, an utility-theoretic approach is suitable. Let $s^\bullet$ ($\bullet = A$ or B) denote the partial market sizes, or **market shares** for good A, and B, respectively. As all other variables introduced below, they are considered as dependent variables $s^\bullet = s^\bullet(s)$ satisfying $s^A + s^B = s$. This account manifestly treats A and B as substitute goods, i.e., agents decide exclusively for either one or the other.

To describe the decision probability $\rho^\bullet = \rho^\bullet(s)$ for buying A or B, respectively, at saturation s, at least three factors need to be taken into account. The first is the distribution of the genuine, individual utilities $u^\bullet$ of the good across the population. The second is the individual utility $u_i^\bullet \stackrel{\text{def}}{=} u_r^\bullet - \pi^\bullet$ originating from individual utilities $u_r^\bullet$ arising from expected resales revenues, where $\pi^\bullet = \pi^\bullet(s)$ is the price of the respective goods. In the present model these two factors are considered as exogenous ones, while the third one is an endogenous, generic network effect, captured in a contribution $u_m^\bullet$ to the utility. It is convenient to introduce, for all utilities, the **bias** $\Delta x \stackrel{\text{def}}{=} x^A - x^B$ as a measure for the advantage gained by deciding for A rather than B.

Let $\mu^\bullet = \mu^\bullet(u^\bullet)$ be the probability density function (PDF) of the distribution of $u^\bullet$ across the population. The distributions for both goods are taken to be equal and to depend only on the respective **popularities** $p^\bullet \geq 0$, i.e., $\mu^\bullet(u^\bullet) = \mu(p^\bullet, u^\bullet)$. We assume that $\mu(x) = 0$ for $x < 0$, and that μ satisfies the principle of stochastic dominance, i.e.,

$$\mathscr{M}(q,x) \geq \mathscr{M}(p,x) \quad \text{for } p \geq q, \tag{15}$$

where $\mathscr{M}(p,x) = \int_0^x \mu(p,y)\mathrm{d}y$ is the cumulative density function (CDF) of μ. With these settings, the probability that an agent decides to buy A is $\rho^A(\Delta) \stackrel{\text{def}}{=} \Pr(\Delta u + \Delta > 0)$, where the **decision bias** Δ subsumes all other utility contributions to the bias for A. It follows, with the notation $\rho^A(p^A, p^B; \Delta) = \rho^A(\Delta)$, making the dependency of ρ^A on the popularities explicit,

$$\begin{aligned}\rho^A(p^A, p^B; \Delta) = \int_0^\infty \mathrm{d}\mu^A(u) \int_0^{u+\Delta} \mu^B(u') = \\ \int_0^\infty \mathrm{d}\mu^A(u) \mathscr{M}^B(u+\Delta) = \\ \int_0^\infty \mathrm{d}\mu^A(u) \mathscr{M}(u+\Delta-p^B).\end{aligned} \tag{16}$$

In simple models as used below, the distributions $\mu^\bullet$ are given in translation form $\mu(p^\bullet, u^\bullet) = \mu(0, u^\bullet - p^\bullet)$, in which case (16) simplifies to

$$\rho^A(p^A, p^B; \Delta) = \int_0^\infty \mathrm{d}\mu(u) \mathscr{M}(u + \Delta + \Delta p), \tag{17}$$

where $\Delta p = p^A - p^B$ is the **popularity bias**.

With the probability $\rho^A(s) = \rho^A(p^A(s), p^B(s); \Delta(s))$ to buy A at a given total saturation at hand, we can write down the fundamental relation governing the dynamics of the multi-level market in which A and B compete.

$$s^A(s) = \int_0^s \rho^A(s')\mathrm{d}s'. \tag{18}$$

The second element contributing to the decision bias is the agents' *ex ante* estimation of resales revenues and the incentive, thus defining $u_r^\bullet$ and in turn the **resales revenue** and **incentive bias** Δu_r and $\Delta u_i = \Delta u_r - \Delta\pi$, respectively. Due to limited knowledge about the market situation, agents are bound to behave according to a rule of bounded rationality and using partial information. We choose $u_r^\bullet(s) \stackrel{\text{def}}{=} u_r^\bullet(s) \cdot \rho^\bullet(s)$, where $u_r^\bullet$ is the bare resales revenue $u_r^\bullet(s) = \int_s^1 \pi^\bullet / s' \mathrm{d}s'$. Here, $\rho^A(s) = \rho^A(p^A(s), p^B(s); 0)$, and $\rho^B(s) = \rho^A(p^B(s), p^A(s); 0)$ are the probabilities for buying A, B, respectively, governed merely by popularity. That is, agents expect to gain the resales revenue of an undisturbed multi-level market of relative size $\rho^\bullet(s)$. Sellers transaction costs, which can be assumed to be of similar magnitude for both goods, and small for virtual ones are neglected, as well as commissions by which we focus on the competition between the goods, exclusively. The assumptions on the agents' accessible information underlying this Ansatz are i) the price schedules $\pi^\bullet(s)$ are public knowledge, ii) s can be estimated with good precision, as well as iii) $\rho^\bullet(s)$. While

i) depends on the mechanism implemented by the MLIM system, ii) and iii) can be justified to the end that they represent information accessible through *local measurements* within an agent's communication reach. Summarising, this definition of $u_r^\bullet$ represents partially but rather well informed individuals which behave subjectively rational. Further discussion of $u_r^\bullet$ is contained in Sections 5.1 and 5.2.

As already alluded to in Section 2, the dynamics of multi-level markets is very likely to be affected by network effects. In fact, in a completely homogeneous market and in the absence of other externalities influencing an agent's decision, a network effect becomes dominant. For, if resellers of good A, say, are rare then a buyer will be very likely to buy from a reseller of B. In such a situation ρ^A can become negligible and the market completely governed by the **multiplier effect** of resellers of B. We do not presume such an extreme effect to be prevalent, and, since generic utility-theoretic treatments of network effects are lacking except for special cases, we choose an *ad hoc*, moderate multiplier utility $u_m^\bullet \stackrel{\text{def}}{=} \varepsilon s^\bullet/s$ depending on an adjustable parameter ε. This yields a **multiplier bias** $\Delta u_m = \varepsilon(s^A - s^B)/s = \varepsilon(2s^A/s - 1)$ as the single endogenous contribution to ρ^A. With the specification

$$\Delta \stackrel{\text{def}}{=} \Delta u_i + \Delta u_m = u_r^A \rho^A - u_r^B \rho^B - (\pi^A - \pi^B) + \varepsilon\left(\frac{2s^A}{s} - 1\right) \quad (19)$$

the model for the competition of two goods in a MLIM market is complete. Note that (16), (18), (19) present an exactly solvable integral equation for s^A. Will will now examine some special numerical solutions of it.

4. Analytical Results in Two Special Cases

Though the presented competition model is simple, the space of situations covered by it is vast. Input data are the price schedules $\pi^\bullet$, popularity functions $p^\bullet$, and the multiplier factor coupling ε, but also the dependency of μ on the popularities. Here we assume that the latter be of translation form (17), and specify that $\mu(0,u)$ is given by a special form ($f(u;1,2)$) of the commonly used Weibull distribution, see [10], in which case ρ^A takes a simple analytical form. For $\pi^\bullet$ and $p^\bullet$ we specialise to spike functions

$$g(s;m) \stackrel{\text{def}}{=} \begin{cases} s/m, & \text{for } 0 \le s \le m; \\ (1-s)/(1-m) & \text{for } m < s \le 1. \end{cases} \quad (20)$$

Price schedules of spike form offer an early-subscriber discount and a late-adopter rebate, cf. Section 5.3. Technically, they are the simplest price schedules which avoid an initial singularity, thereby minimising the variance with a discrete model, and correspond to markets closing at finite size.

Besides the market shares $s^\bullet$ and the **final shares** $S^\bullet \stackrel{\text{def}}{=} s^\bullet(1)$, the **turnovers**

$$t^\bullet(s) \stackrel{\text{def}}{=} \int_0^s \pi^\bullet(s')\rho^\bullet(s')\mathrm{d}s' = \int_0^s \pi^\bullet(s')\dot{s}^\bullet(s')\mathrm{d}s' = \int_0^{s^\bullet(s)} \pi^\bullet(s^{\bullet\prime})\mathrm{d}s^{\bullet\prime} \quad (21)$$

and the **total turnovers** $T^\bullet \stackrel{\text{def}}{=} t^\bullet(1)$ are important indicators for the economic performance of the competing goods. Note that the maximal turnover that a good can generate is $1/2$ for spike functions. Furthermore, we examine the discrepancy between agents' expectation and the actual resales revenue they can achieve, similarly calculated as

$$v_r^\bullet(s) \stackrel{\text{def}}{=} \int_s^1 \frac{\pi^\bullet(s')}{s^\bullet}\rho^\bullet(s')\mathrm{d}s' = \int_{s^\bullet(s)}^{S^\bullet} \frac{\pi^\bullet(s^{\bullet\prime})}{s^{\bullet\prime}}\mathrm{d}s^{\bullet\prime}, \quad (22)$$

and the resulting actual incentive $v_i^\bullet(s) \stackrel{\text{def}}{=} v_r^\bullet(s) - \pi^\bullet(s)$.

4.1 Free-Rider Phenomena

To counter free-rider phenomena is the main aim behind the conception of MLIM. In fact, the content distribution network of MLIM systems like the Potato system [8, 9] is very similar to the peer-to-peer networks commonly used by free riders. By this rationale, we can compare the performance of a virtual good A with a pirated version B of it in the *same* multi-level market. That is, the popularities are equal $p^A = p^B$ and B is free, i.e., $\pi^B = 0$. Since no confusion can arise, we sometimes drop the superscript A.

Figure 2 shows the plateaus of S^A and T^A in dependence of m and ε. Even without a multiplier effect present, incentives can lead to a non-negligible market share though not dominance. However, significant turnovers are not generated without exploiting the multiplier effect by an initial invitation to enter, i.e., a positive incentive at early times. For multiplier biases $\varepsilon \cong 1$ comparable to the price and other biases, good A can reach market dominance and generate over $1/2$ of the maximum turnover. To maximise turnovers, the price schedule must be aligned with the market growth s^A, which is generally difficult. It can be seen that maximisation of turnover and share are conflicting goals.

The market evolution in this setting is studied in more detail in [10]. The main observations are: i) The simple rule for u_r leads to good estimations for v_r, and in turn v_i. Agents tend to underestimate the resales revenues they can achieve at early times and overestimate them only in an intermediate phase. ii) An early peaking price schedule $m \approx 0.1$ entails an initially high and then steeply dropping incentive bias. iii) A later price peak leads to a smaller, but longer lasting positive initial incentive for A. iv) For late peaking prices ($m \approx 0.9$), Δu_i has a sharp negative peak at high saturations, i.e., a significant entry deterrence for latecomers times.

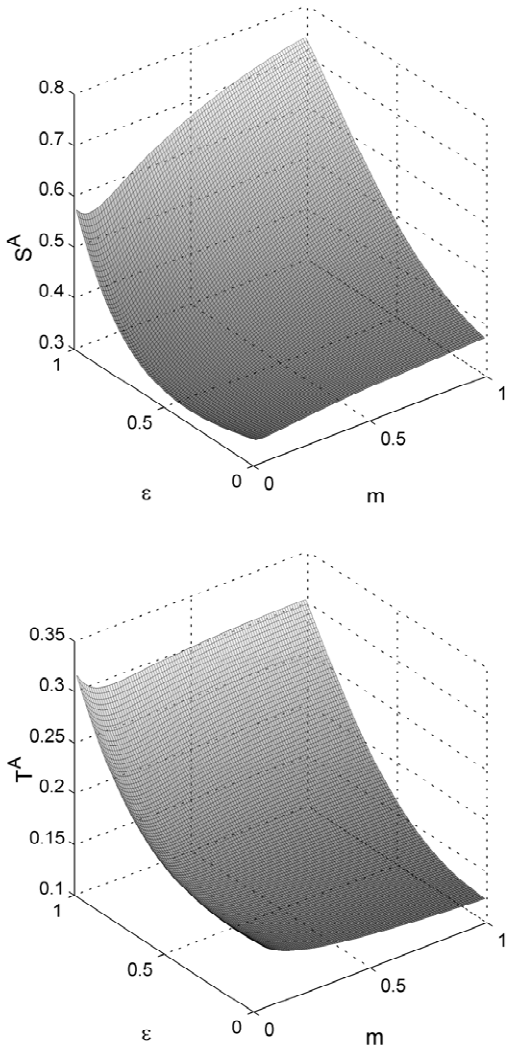


Figure 2. Final shares (top) and total turnovers (bottom) in the free-rider setting.

4.2 Smash Hits and Sleepers

Scenarios for the competition of two goods are manifold within our model and lack of space prohibits a comprehensive treatment. As a familiar example, we considered in [10] the case that good A has a popularity function peaking later than that of B, i.e., A would commonly be termed a 'sleeper' while B can be considered a 'smash hit'. The originator of A would like to counter the slow startup effect due to later popularity utilising an appropriate price schedule, corresponding to various positionings of the peak m^A of his price function. The price function of B is assumed to be centred, $m^B = 0.5$. From the various examples considered in [10] it can be seen that the final share of A is mostly small if the multiplier effect is strong, since then the early rise in popularity of B gives B a persistent advantage. As a central result, to counter this by a long lasting rebate, i.e., a late price peak m^A is in fact possible. The opposite strategy to start the market by an early peaking price and therefore high initial incentive can also work. However in the latter case, the price function of A is misaligned with the market evolution and hampers the generation of turnovers. In conclusion, to optimise the price function of the sleeper so as to obtain good market shares *and* turnovers, is difficult.

5. Discussion and Practical Implications

5.1. Similarity to Pyramid Schemes

Multi-level marketing carries negative connotations and is illegal in special forms known as pyramid selling, snowball systems, chain-letters, etc., under many jurisdictions. The study [13, Vol. II] presents criteria to distinguish between legitimate multi-level marketing and such practises to be considered illicit. In view of them, five arguments can be produced in favour of the legitimacy of multi-level marketing of virtual goods in general, and the MLIM systems within the scope of the present model in particular. First, illicit schemes often require resellers to keep a large, non-returnable stock of the good. The effect of this kind of *inventory loading* is however not present in the case of virtual goods, due to the very nature of information goods. Second, also the marginal costs for their replication and redistribution are mostly orders of magnitude smaller than resale prices and thus transaction costs are largely insignificant. In the Potato system for instance, the processing of resales, including accounting, billing, and charging is fully borne by the central server, for which a percentage of the price is assigned to the system. Third, the compensation plans of illicit schemes often emphasise recruitment of personnel over resale of the good. This is not the case in MLIM where incentives are strictly bound to individual sales of a virtual good of positive pecuniary value. In other words, every agent can at least expect a rebate on the price paid for the good through resales revenues. Fourth, our present model does not allow for down-line payments. Only the directly succeeding level $n+1$ of agents to which the good is directly sold contributes to the revenues of agents at level n. Although, down-line payments for resales are not seen as problematic by [13, Vol.II, page 236], we would argue that they shift incentive payment too much from individual sales efforts to uncontrollable market dynamics. Finally, realistic information on achievable revenues is crucial for the legitimacy of multi-level marketing. The present MLIM model offers in principle the possibility to determine and publish the price and incentive schedule in advance, see below.

5.2. The Free-Rider Problem

Whether MLIM can be successful in meeting the aim to fully replace copy protection measures and conventional DRM is a question for theoretical economy, see [14]. If the good is freely available, as, for instance, in the Potato system, then it is not *a priori* clear that another equilibrium apart from $S^A = 0$ (only free riders) exists. However, the zero-sum condition tells us that an agent partaking in the IM market is on the average not worse off on average than a free-rider, and thus a market of any size is in fact a global equilibrium. Whether such an equilibrium is likely to evolve dynamically is quite a different question, which has been answered affirmatively at least for our simplified dynamical model in Section 4.1.

The free-rider phenomenon is closely connected to the issue of fairness and the economical purport of information. For if the zero-sum condition is *common knowledge*, then rational agents would always choose the free good since they know that later potential buyers with negative (actual or subjective) will do so. This renders the success of real pyramid schemes paradoxical, and shows that the incentive schedule is at most *public knowledge*: There must be agents who know that some others will have a negative incentive but expect them to enter the market nonetheless. This is the reason for modelling the decision mechanism of agents using a rule of bounded rationality, as in Section 3.

It is conceivable that the incentive through resales revenues is insufficient to make MLIM effective against free riders. Originators could combine it with copy protection, or otherwise discriminate the legal version in the MLIM system from illegal copies distributed over P2P networks, e.g., by added value and/or mild forms of copy protection or traceability through fingerprinting etc.

5.3. Dynamical Forward Pricing

A new option arising from the model presented is the possibility, via the inversion formula (7), to dynamically adapt the incentive during the evolution of the market if the originator controls the price as an external parameter. Such **MLIM systems with dynamical forward pricing** can be used to design actual market mechanisms. Dynamical forward pricing is not a new concept for information goods, but has not been widely considered in the context of multi-level markets, neither for virtual nor physical goods.

Figure 1 shows basic possibilities for price functions. The constant price in a) is associated with a strong favouritism of early buyers, and increasingly penalises later ones. A typical example for what is conventionally termed an early subscriber discount is shown in b). Such a price schedule is often used as an initial invitation to enter, i.e., a means to spur the distribution of the good in an early stage, for instance to counteract a slow startup effect. This can become important to counter free-riders, since in their presence early buyers cannot be sure about their potential resales revenues which depend logarithmically on the market size (remember that $v_i(k = sn_\infty)$ scales as $\ln n_\infty$). The price associated with the incentive in b) is monotonous increasing, thus later buyers pay more and receive less incentive, and are thus disfavoured. Example c) improves on b) by letting the price vanish when the market reaches saturation. This v_i combines an early subscriber discount with a rebate for late adopters who finally obtain the good gratuitously, a price schedule which can spur the distribution of the good in late phases, when it may have lost individual utility, e.g., due to dwindling popularity. Assuming that the market has a positive, endogenous growth dynamics in an intermediate phase associated with a high demand, it is reasonable to let the prices peak and lower the incentive in this phase, as in c). Deepness and position of the minimum of v_i can be adjusted almost arbitrarily. Finally, d) shows the effect of a commission on the incentive. In particular it can be seen that the point at which the incentive becomes negative is not significantly shifted with decreasing γ.

For an implementation of dynamical forward pricing, in particular the current size $n(t)$ of the market must be known. This is the case when a central server counts every single acquisition of the good, as, e.g., realised in the Potato system. The market size n_∞, necessary to calculate the saturation $s = n/n_\infty$, is more difficult to determine. Though it could be estimated by market research, comparison with earlier runs of the system, or other means of educated guessing, a more pragmatic solution suggests itself. As in Figure 1 c) and d), setting the price to zero after some finite time, respectively at an *a priori* given n_∞ obtains a condition for *closure* of the market. Though running counter to the aim of maximising the diffusion of the good, the effect on the turnover is limited if the price becomes small enough at high saturations.

Mixed forms of dynamical price settings can be envisaged, e.g., correlation of π with the buying frequency, combined with a frequency or price threshold below which the price is set to zero and the market closed. In any case, designing the optimal price schedule is a complex task, in particular in competitive situations. Then arises, for instance, the additional difficulty that the total market saturation for all goods cannot be determined by a single party.

5.4. Market (In)homogeneity

Multiplier effects are a prominent example for details which undermine one crucial assumption underlying the model presented, namely *homogeneity* of the market. Uniform agents in a structureless market are a good approximation if the number of potential participants is large and consists of a rather homogeneous group of individuals, for

instance with special personal preferences, e.g., musical. However, if the market is biased in the sense that there is a group of agents with systematically higher trading capacities, the assumption breaks down. In reality, large music labels running direct sale web sites are a counterexample where this heuristics is violated. On the other hand, inhomogeneities and multiplier effects carry the imminent danger that an MLIM market can be cannibalised at an early stage by an agent with overwhelmingly high communication capacity, e.g., a popular web site, who could then obtain a practical monopoly. The study of [15] indicates that monopoly creation could be a rather natural effect in E-commerce. While the originator of the good is not too affected by this phenomenon if a commission model is used, the other buyers' incentives are always negatively affected. To what extent the market can be levelled by means of the IM system, e.g., by providing equal communication capacities to all participants, restricting or controlling resale volumes or frequencies, etc., warrants separate discussion.

6. Conclusions

Let us briefly note some directions for further work. On the theoretical side it would be desirable to improve the both the monetary flux model and the competition model to account for, e.g., market inhomogeneities in the former and the influence of further externalities on the agents' decisions in the latter. In particular, a better justified model for the multiplier effect and a proper incorporation of other network effects is wanting. Refined simulations of multilevel markets in the framework of agent-based computational economics [16], can be useful. A proper treatment from the viewpoint of theoretical economics should also answer questions of optimality, equilibria, and their stability. The free-rider problem in MLIM should also be treated in a more theoretical approach using the principal-agent model [14], to describe the effect of the incentive on the moral hazard incurred by the agents. Pragmatically, the most daunting task from the present viewpoint is to ensure equal opportunities for resellers in the market, i.e., to practically corroborate the theoretical assumption of homogeneity — a rather typical problem of genuine E-commerce similar to those faced by reputation systems commonly used by Internet auction houses.

References

[1] C. Shapiro and H. Varian: *Information Rules*. Harvard Business School Press, 1999.

[2] P. Aichroth and J. Hasselbach: Incentive Management for Virtual Goods: About Copyright and Creative Production in the Digital Domain. Workshop 'Virtual Goods' 2003, Ilmenau, Germany, on-line publication, pp. 70–81. http://virtualgoods.tu-ilmenau.de/2003/incentive_management.pdf

[3] E. Becker, W. Buhse, D. Günnewig, and N. Rump, (Eds.): Digital Rights Management — Technological, Economic, Legal and Political Aspects, Lecture Notes in Computer Science 2770, Springer-Verlag, Berlin, Heidelberg, 2003.

[4] H. Kinokuni: Copy-protection policies and profitability. *Information Economics and Policy*, 15:521–536, 2003.

[5] *Information Economics and Policy* 16, special issue, 2004.

[6] D. K. Mulligan, J. Han, A. J. Burstein: How DRM-based content delivery systems disrupt expectations of "personal use". In: *Proceedings of the 2003 ACM workshop on Digital Rights Management*, pp. 77–89.

[7] F. Oberholzer and K. Strumpf: The Effect of File-Sharing on Record Sales: An Empirical Analysis. National Bureau of Economic Research Universities Research Conference "Economics of the Information Economy", 7 and 8 May 2004, Cambridge, MA. http://www.nber.org/~confer/2004/URCs04/felix.pdf

[8] R. Grimm and J. Nützel: Security and Business Models for Virtual Goods. In: *Proceedings of the 2002 ACM Multimedia Security Workshop*, pp. 75–79.

[9] J. Nützel and R. Grimm: Potato System and Signed Media Format — an Alternative Approach to Online Music Business, *Proceedings of the 3rd International Conference on Web Delivering of Music (WEDELMUSIC 2003)*, pp. 23–26, IEEE Press, 2003.

[10] A. U. Schmidt. Incentive Systems in Multi-Level Markets for Virtual Goods. E-print, v2, July 2005. http://arxiv.org/abs/cs.gt/0409028

[11] A. U. Schmidt. Multi-level markets and incentives for information goods. *Information Economics and Policy*, to appear.

[12] G. M. P. Swann. The functional form of network effects. *Information Economics and Policy*, 14:417–429, 2002.

[13] H.-W. Micklitz, B. Monazzahian, and C. Rößler: Door-to-door selling — pyramid selling — multilevel marketing. Study Commissioned by the European Commission. Volume II: Analysis, 1999. http://europa.eu.int/comm/dgs/health_consumer/library/surveys/sur10_en.html

[14] J.–J. Laffont and D. Martimort: *The Theory of Incentives*. Princeton University Press, Princeton, 2002.

[15] S. M. Maurer and B. A. Huberman: Competitive dynamics of web sites. *Journal of Economic Dynamics and Control* 27:2195–2206, 2003.

[16] L. Tesfatsion: Introduction to the special issue on agent-based computational economics. *Journal of Economic Dynamics and Control*, 25:281–293, 2001.

Alternative Distribution Models based on P2P

Martin Schmucker
Fraunhofer IGD
Darmstadt, Germany
martin.schmucker@igd.fraunhofer.de

Peter Ebinger
Fraunhofer IGD
Darmstadt, Germany
peter.ebinger@igd.fraunhofer.de

Abstract

Until now commercial distribution architectures for digital content have been primarily based on centralized systems. P2P networks' capabilities however demonstrate increased reliability, scalability, fault tolerance, load balancing, and performance over centralized solutions. Additionally P2P networks allow the transfer of storage and network costs to their participants. Existing P2P architectures however are "grown" architectures and do not fully exploit available technologies in order to satisfy legal requirements. From a user's point of view it is difficult to understand why the usage of current P2P-networks should be illegal at all.

This article investigates the requirements for a P2P-framework that overcomes the previously described drawbacks. The proposed architecture is a framework for the legal distribution of commercial and non-commercial content via P2P networks. It supports a wide-range of business models ranging from shareable (promotional) content to DRM protected commercial content, ensuring legal exchange without centralized content usage controls.

The proposed framework exploits technological potentials while at the same time maximizing its usability and attractiveness to users. The framework also ensures that consumers act in a legally acceptable manor and any illegal infractions will be flagged. The system achieves this through a process where each peer observes the peers it is exchanging content with thus increasing the probability of identifying infractions. In addition to addressing the requirements of consumers, content owners and distributors alike, this framework can also incorporate technologies that increase its attractiveness to users by providing additional services like collaborative filtering. In this way users are assured that the content they are accessing is both legal and of commercial quality.

In summary, this framework provides a reasonable balance between content owners' rights on content protection and users' requirements on usability and privacy.

1 Introduction

Content owners, providers, and distributors have faced a lot of problems during the last few years. One of their major problems is the illegal distribution of music via P2P-networks: compression technologies and omnipresent broadband Internet access points have allowed computer users illegal content exchange on a scale which is neither measurable nor imaginable. This development was initiated by techies who exploited the technical potential of existing infrastructure and provided this potential as free services to other computer users. The results were content exchange networks which can be misused. In addition, these content exchange networks seduce users to illegal content exchange: the illegality of simple click on content in P2P-clients for downloading or moving files to and from directories is hardly understood by ordinary users.

As a result the content industry spends a lot of effort in building awareness of file-sharing illegality to users. The content industry is also in favor of protecting valuable commercial content with restrictive DRM technologies which has the side effect of deterring potential customers for purchasing content. Some artists do not concern themselves with illegal file-sharing of their content nor about the users' rejecting reaction to DRM and are instead only interested in the promotion of their content. Thus these stakeholders need a legal platform for the promotion of their non-commercial content. Additionally this platform should also support a wide range of business models. One has to be aware that the focus of most artists is on the legal commercial and non-commercial distribution of their content not on usage control which is not in their best interests.

DRM-protected content has to compete with illegal distribution as both provide access to relatively the same content. There are always ways to access unprotected content - at least the analogue hole. This unprotected content can easily be distributed via the Internet as content distribution cannot be controlled (completely) on the Internet [17]. Thus legally offered material must always compete with illegally accessible content.

As a consequence, a legal platform that only allows legal exchange of content without limiting content usage itself is required. Ideally consumers should not experience any limitation to the content usage in such a system. In contrast to existing ”grown” file-sharing systems the content exchange must be on a legal basis. This means that either the content is exchanged legally or any misuse is identified and traced back to the offending user.

The importance of these two requirements - a legal distribution platform neither limiting content usage nor its consumption while supporting a broad range of business models - has up to now not been considered adequately by content owners, providers, and distributors; nevertheless these are the main requirements upon which the presented framework is based.

1.1 Further Requirements

As audio visual data incorporates a huge amount of data, a P2P based distribution system is advantageous as content creators, owners, and distributors can transfer storage and distribution costs to customers.

The security of the presented framework is based on identification of users and tracing of illegal content exchange thus the storage of user related information is mandatory. While privacy is a strong concern for current DRM systems, this is not the case for the proposed system. The difference between the proposed solution and existing DRM systems is that only content exchanges are observed and not the individual usage of content in particular. The storage of personal data is only required for the validation of the exchange process, which is a short time span. Therefore the lifetime of user related exchange data is very short and the proposed framework fully considers privacy requirements.

If a consumer legally downloads content (always the case under the presented framework) only he will decide about the future personal usage of this content. The redistribution of content within the system is restricted by the framework regulations.

1.2 Supported Business Models

Any relevant content distribution architecture must be capable of supporting the distribution of different types of content and the application of a wide range of business models.

Initially, the proposed framework supports two different content categories:

- Content that can be exchanged without limitations: This category comprises promotional content or other content available free-of-charge.
- Content whose usage is limited by a complex license and protected by DRM: Content creators, owners, and distributors are free to use existing DRM solutions - including the so-called light-weight DRM-system or alternative models like the Potato-System. Thus, the presented distribution framework satisfies their commercial interests.

An extension for (unprotected) content that can be received only after paying is possible. As discussed in the outlook in section 5 CONFUOCO also allows the support of a broad range of business models including flat rates (”all you can consume”). Another advantage of CONFUOCO is that collecting societies can be provided with the essential information about the exchanged contents since this data can be transmitted without user related information. Content distributors can freely decide which business and protection model fits best to their requirements. Thus, the use of DRM is not mandatory if artists only require their customers to acquire content legally within the proposed framework.

1.3 Security

As long as trusted computing is not available, perfect security is not possible as each system is under control of its owner. Perfect security however is not required for the presented framework. By using software authentication, software component updates and further technologies, the proposed framework provides a reasonable level of security. The effort for manipulating the proposed solution will be very high and the probability for not identifying individual manipulations is very low. In fact users interested in illegal content exchange will most likely not use the proposed framework for illegal exchange as other available tools are more practical for illegal exchange - such as those on which DarkNets are based[17].

1.4 State-of-the-art

Different distribution technologies (some in combination with protection technologies) have been developed and are available. Peer-to-peer applications include chat, collaboration, white boarding, games, file sharing and content distribution. File sharing in particular has a negative connotation as the most well-known file sharing P2P-networks are used for illegal exchange of content. As stated in [18] “intellectual property is more of an issue in P2P network because there is less built-in control.”

Although different commercial P2P-based solutions are available, many of them concentrate on the B2B area like kontiki [6]. Some of them address B2C-scenarios like Dijjer [4] or Red Swoosh [12]. These solutions aim at the distribution of large files like movies and video games or

software distribution (updates). The benefit of these systems is the transfer of storage as well as distribution costs to customers. The security of each system is based on content encryption and content/user authentication (E.g. in kontiki only dedicated users are allowed to publish content).

For the legal and commercial distribution of audio content most solutions are centrally controlled music download shops such as iTunes [2] or RealNetworks [11]. Although these stores prove that customers are willing to pay for digital music downloads, they mainly map traditional (analog) business models and distribution channels to the digital domain. Only a few commercial P2P-based distribution systems exist. Among the available P2P-based solutions are the Potato-System [10] and Weed [13]. The focus of the Potato-System is actually a business model that was extended with a so-called P2P-Potato-Messenger. The Potato-System awards a commission to people if they pass songs on to friends who buy them. If some users' behavior is "un-cooperative" their P2P-client can be disabled. Similarly Weed is also based on a commission, but the users are only allowed to play the music three times free of charge by using Microsoft's Windows Media DRM [8]. This allows the distribution of Weed files on P2P-networks. Further solutions are announced like Music2Share [21] or Peer Impact [9]. In Music2Share the content is distinguished as public, private or non-authorized. The user can only download content after he paid for it. It "does not take special measures to prevent nonauthorized content from spreading across the network" [21]. During writing this article no details about Peer Impact were available.

Most of the commercial solutions described above integrate digital rights management (DRM) technologies such as [8, 5], which generally impede customers in traditional content usage. This is the reason DRM is also seen as the acronym for "digital restriction management". To weaken these restrictions, different solutions have been introduced such as the so-called "light weight DRM" (LWDRM) [7]. Light weight DRM is based on digital certificates. The idea is to attach information about the person distributing content to the content itself. For this purpose different file formats were defined. Additionally digital watermarks can store user information. Although this system is called "light weight" it still imposes restrictions on content usage.

2 Design Criteria and Architecture

Considering the huge volume of data transferred when downloading audio-visual content, a centralized distribution solution will sooner or later be a bottleneck in the dissemination of content. Distributed systems and particularly P2P-systems allow the transfer of storage and network costs to customers. Other capabilities of P2P systems include reliability, scalability, and performance [18].

Each framework for content distribution has to address specific criteria for success: first, content distribution within a P2P-system must not infringe on IPR. Second, the usage of content distributed within the network must not interfere with traditional content utilization.

The first requirement for any architecture for content distribution is in itself challenging: the potential misuse of content exchange infringing on IPR. A perfect distribution solution will not allow the unauthorized distribution of content.

A truly perfect distribution solution however is only possible if it runs on a trusted device that is not under full control of the user (cf. [14]). Unfortunately for content providers this level of control cannot be achieved as users will neither accept the expensive of such solutions nor will they spend additional money for systems with reduced functional value. The optimal solution considers the protection of IPR, the functional loss and monetary costs for consumers. These requirements can be met by enforcing the customers' liability in the case of misuse which can be implemented with the use of two strategies in tandem:

- Technology must be used within a distribution framework that is able to identify users and the content that is distributed. However under this strategy the identification of content must not be limited to cryptographic hashes. Content based identification, also known as fingerprinting technology, is also mandatory [15, 16, 20, 24]. Combined with "black lists" and "white lists" exchanged content can be limited to authorized content.

- Social issues like community affiliation strongly affect users' behaviors within the group, therefore building a user community with adequate rules is significant for the success of the system.

This article focuses on the technological requirements and architecture of a secure P2P distribution framework, community aspects of the system will be addressed in a future article. The typical use cases have to be analyzed as shown in figure 1. A User can have two roles:

1. As a Content Owner the user inserts content in the P2P-distribution framework. Additionally he can revoke the right to distribute content within the distribution framework.

2. As a Content Consumer the user downloads or exchanges content and "unbags" it from the distribution framework. The last use case is especially important for the usability of content distributed within the network. Content migration from the P2P-system to the outside should be as easy as possible.

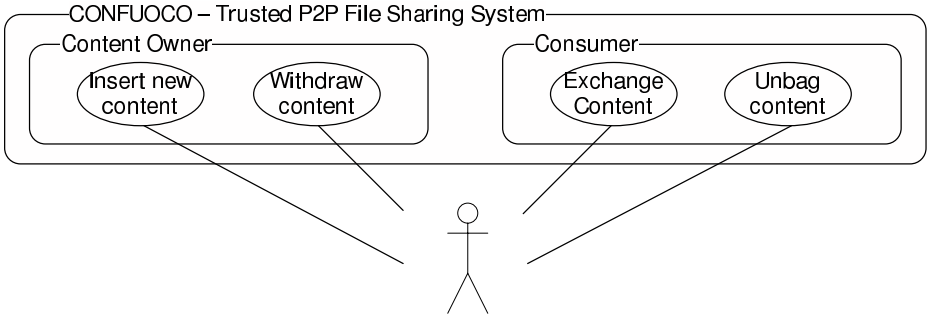


Figure 1. This figure shows typical use cases when exchanging content within the P2P system. As shown a general user can have two roles within the system: Content Provider or Content Consumer. In contrast to existing P2P frameworks one outstanding issue is the revocation of content distributed within the P2P system. This considers musicians' requirement to stop the distribution of certain content if necessary (such as if a musician uses the proposed P2P framework as a promotional vehicle initially and then they want to stop certain content exchange after they have signed a contract with a record label).

For ensuring the authorized distribution of content two strategies have so far been implemented:

- The benefits of a DRMS are limited, and restrictive DRMS has the disadvantage that content usage is impeded which drastically lowers the interest of consumers. Thus users will not accept DRMS solutions and content owners will not reach enough consumers.
- Fingerprinting and watermarking technologies are considered to be passive protection technologies. Up until now these technologies required an external control entity which analyzes the data exchanged within (similar to a police man observing traffic speed) [16, 19, 14].

The required control unit imposes major obstacles and thus we propose to develop a P2P-network which integrates fingerprinting and cryptographic hash technologies in each peer. In this type of P2P-network, each peer acts as the previously described control instance. Illegal content exchange is only possible within a group of 'traitors'[1]. As these groups also exchange content with other peers, their risk of being identified is very high and thus these users will likely use other solutions to exchange content illegally.

The simplified architecture of CONFUOCO is shown in figure 2 and consists of several sub-systems. Each peer has a user interface which controls the login to the P2P-network and the content exchange. User registration and identification is managed by trusted third parties who also manage and validate content exchange.

[1]Traitors refers in this context to people intending to misuse this system for illegal content exchange.

2.1 User Registration and Authentication

During the registration process the User is authenticated by the UserRegistration TTP. This could be done for example through Internet Service Providers (ISP) where the User is already registered, or he could be requested to enter some personal information (e.g. address and phone number). In the second case, this information should be checked for credibility and validated with a confirmation letter sent to the stated address.

Afterwards, a unique identification number is generated as UserID and the User can choose a pseudonym or nickname[2] for identification within the P2P network. The User is requested to set an initial password and the certificates to identify the TTPs are stored on her local client. From now on the User can securely identify the TTPs and vice versa.

The detailed user data is stored at the UserRegistration TTP and only the pseudonyms (UserID and nickname) are submitted to the UserIdentification TTP and added to the list of valid users. During transactions in the P2P network the UserIdentification TTP simply has to check if the pseudonym of a User is in the list of valid users and if the password provided by the User is correct.

2.2 Content Registration

Only registered content can be exchanged within the P2P network. If a User wants to register new content he first

[2]It is important that the User can choose a personal name for his "virtual personality" in the community. This name can reflect some personal attitudes or help other users to visualize something with the User's ID (as it is easier for a user to remember than a number).

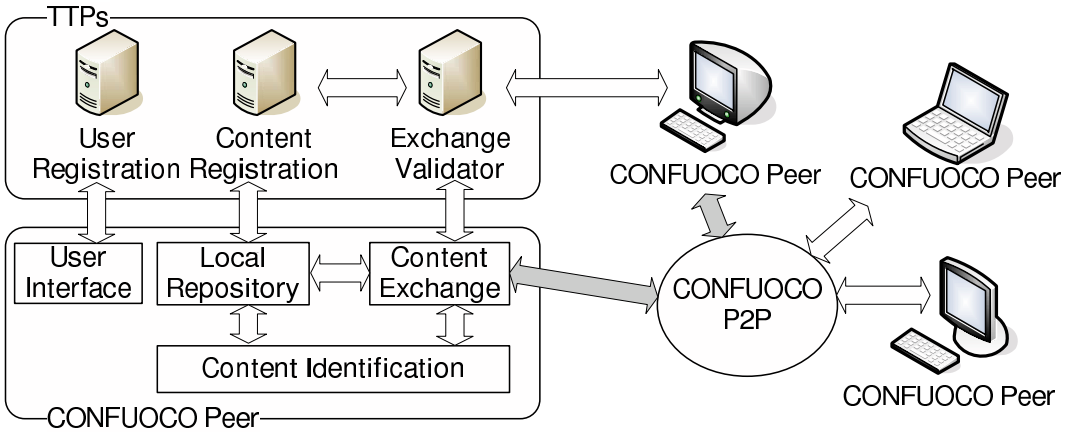


Figure 2. The main entities within the general architecture of CONFUOCO are trusted third parties (TTPs) for user registration and identification, TTPs for content registration and validation, and the peers that consist of several components like user interface, local storage interface, content identification and P2P-networking.

submits the fingerprint (or hash value) of the new content to the ContentValidator TTP to verify that the content is not currently registered (on either the black or white lists).

If the validation process is successful the User submits the new content to the ContentRegistration TTP. This TTP calculates the fingerprint and hash value of the content and verifies that the content is valid.

The calculation of each content identifier depends on the content type. For unencrypted content the fingerprint and the hash value can be calculated and used for content identification, whereas for encrypted content (e.g. DRM protected audio files) only the cryptographic hash value is available.[3]

Following these steps the content can be registered and detailed content data (UserID, timestamp, fingerprint and hash value of the content, optionally: meta-data and validity period) is stored at the ContentRegistration TTP.

Only the fingerprint and the hash value of the content (optionally meta-data and validity period) are submitted to the ContentValidator TTP and added to the white list of sharable content. The User receives a license containing the sharing permission, the identification of the content and the ContentRegistration TTP registering the content.

During transactions the ExchangeValidator simply checks that all content exchanged is registered as sharable using the ContentValidator.

[3]Encrypted content is considered as a binary large object ('blob').

2.3 Peer-to-Peer Client

The P2P-Client is the interface between the users and the P2P-system. It allows users to upload new content, exchange it and transfer content out of the P2P system (e.g. to other devices).

- The *UserInterface* manages communication with Users such as login to the P2P-network or browsing for new content. It also provides a file manager for the insertion of new content and the transfer of existing content out of the P2P system.

  The P2P-client's repository is represented by an ordinary directory of the file system. Users can therefore copy selected content in and out of the P2P-system within an easy to use interface which results in an simple transfer of the content file within the file system.

  No distinction between encrypted objects and unencrypted content is necessary here. The main advantage is that unprotected content can easily be copied to and from other directories, hard discs, or other devices.

- The *LocalStorageInterface* observes changes in the repository directory not caused by a file download. When new content is added a fingerprint value is calculated and the TTP responsible for validation is contacted. Depending on the result of the validation different actions are initiated.

  If the content is

  - already registered and sharable: no limitations apply.

  - already registered and not sharable: it is transferred to the QuarantineWard.
  - not registered: the User is asked if he wants to publish it. If so, the content is uploaded to the ContentRegistration TTP and registered to this User as described in section 2.2.

  Encrypted (DRM protected) content can only be registered by particular users known as Content Owners.

- The *ContentID* component calculates content identifiers depending on the content type. While for unencrypted content the fingerprint and the cryptographic hash value can be calculated, for encrypted content (e.g. DRM protected audio files) only the cryptographic hash value is available.
- The *QuarantineWard* temporarily stores content which must not be shared. This allows the User to delete the files or move them to another folder or device.
- The *MagicTrunk* implements the P2P-functionality like content search and exchange. It initializes the calculation of fingerprints or hash values for the exchanged files.

  Each communicating peer (both the sender and receiving peer) transmits its calculated content identifiers to the ExchangeValidator TTP. This prevents peers from receiving illegal content while it allows the identification of peers illegally transmitting content. Thus each peer observes the communication which increases the identification of manipulated peers drastically.[4]

3 Use Cases

In this section we describe the important use cases for insertion, exchange and content usage. Content usage is simple transfering content out of the P2P system.

3.1 Content Insertion into the MagicTrunk

If a user wants to add new content to the P2P-System he simply copies the new content within his local file system into the folder that contains the shared content. The P2P client identifies the new content, calculates the fingerprint and submits it to the ContentValidator TTP. The TTP validates the content by using black lists (identifying content that is already registered as not sharable) and white lists (identifying content that is already registered as sharable).

1. If the content is registered as sharable the peer receives a license containing the sharing permission.
2. If the content is registered as not sharable the user receives a warning message that informs him about the potential conflict.
   (a) The user accepts the conflict and agrees that the content is removed from the P2P system. The content is moved into the QuarantineWard to give the user a chance to save the content outside the P2P network.
   (b) The user does not accept the conflict and claims that the content is sharable. He uploads the content to the TTP to solve the conflict.
3. If the content is not registered (neither black nor white list) the User is informed that the new content can be registered and the user approves that he is recorded as the legal creator of the content (cf. section 2.2).

3.2 Content Sharing

A user identifies interesting content provided by another user and sends a requests for this content to the other user. The user already storing the content transmits this request together with information about the recipient to the ExchangeValidator TTP who validates the distribution permission.

After the TTP approved the distribution permission the content is sent to the requesting user. The requesting user confirms the new content by calculating a fingerprint and transmitting this fingerprint to the same ExchangeValidator TTP. The TTP verifies that the two fingerprints match.

3.3 Transferring content out of the MagicTrunk

If a user wants to transfer content from the MagicTrunk out of the P2P network he can simply copy the file into another directory. The user can then render the content with the software he likes or store it on other devices or media, e.g. CD or DVD.

4 Implementation Aspects

This article describes only simplified concepts of the CONFUOCO framework and there are additional implementation aspects that must be considered.

- As the presented framework is based on P2P-networks technology, the scalability of the TTPs is an important issue. Centralized TTPs would be a bottleneck to the system. Only distributed TTPs (i.e. a decentralized approach) can completely fulfill scalability requirements.

[4] As this information does not directly identify the user it can be used for example as for alternative fee distribution models.

- In addition to storing black and white lists on the ExchangeValidators, black lists and their updates could be distributed to the peers on a regular basis. Dedicated servers distributing these black list updates is one possible implementation. This allows each peer to verify the content accessible from other peers and thus ensures the users that the published content on their computer is not violating IPR.

- To reach a higher security level, updating software components of the peers is desirable. In a similar fashion to black list updates, software updates could be distributed within the P2P-system or on dedicated servers. Authentication of the components is vital to prevent the users from utilizing malicious code (which could allow the illegal exchange of content or infiltrate the users computers).

- We already emphasized the peer's functionality in observing each other. So each peer is a control instance ensuring the P2P-system's integrity. In addition active wards could verify the functionality of the peers. These wards can randomly offer legal or illegal content and also request content from other peers. This way they can observe the network without any additional means. The wards only need to be known to the ExchangeValidator TTP to make sure that they will not be prosecuted for providing or requesting illegal content to/from other peers.

- The identification of relevant content has an important role in any content distribution system. Besides the typical meta data already accessible in P2P-networks, further possibilities have to be exploited. We suggest the integration of the benefits of a semantic web based approach. This meta data could be created by the community exchanging data. It could increase the value of the exchange framework and also the users affinity to the network.

- The technology not only impedes illegal content exchange, it also addresses social issues which are important and need to be considered. The users should be encouraged to not only use the system for legal content exchange, but also to get involved in the building process of a user community. The optimum solution is reached when the distribution framework is perceived as a user community.

This two-sided control by the peers makes CONFUOCO more reliable and trustworthy to the user than other systems. Even if the users P2P client is corrupted, the malicious software will not be able to share illegal content without detection and notification.

5 Conclusion and Outlook

In this article we showed how to increase built-in control of P2P-networks to address the issue of IPR. Traditional security and software protection mechanisms (such as the verification of software components by using check-sum or (semi-) automatic updates) are used to provide a basic level of security.[5]

The increased protection of IPR is achieved by integrating fingerprinting and cryptographic hash technology in each client, giving each the functionality of a control agent. The clients mutually monitor the communication between each other so even if an individual user manipulates the client software (a general risk for all DRM systems) the mutual monitoring mechanism ensures the identification of these foul players as their participation in the network inevitably involves communication and data exchange with other peers.

It is important to address the fact that "perfectly secure solution is also difficult to achieve with the proposed framework. For example, if groups of users organize themselves into "trusted subgroups" and manipulate the P2P software for illegal content sharing. There are however already scores of commercially available software for building these so-called "DarkNets", so there is little incentive for users to pursue this course of action especially considering the deterrents of their illegal activities being detected and resulting in consequences. In these instances there are more comfortable alternatives for illegal content sharing.

The proposed framework addresses the needs of customers looking for a legal and reasonable alternative in sharing and distributing music, which does not currently exist. The importance of having a legal alternative is clear as typical users do not want to break the law. By using the proposed solution users act within legal limits and do not have to be afraid of heavy legal consequences resulting from unintentional illegal sharing. Nevertheless in our solution users are still liable for their actions within the CONFUOCO framework.

In addition to the proposed technical framework, a social framework has to be supported that improves the users affiliation with the distribution framework. Ideally users will consider such a framework not as a distribution solution, but as a community. Users who become affiliated with this community will be more apt to accept the general rules within and their actions will conform to these rules. Since these rules comply with legal requirements, so too will users actions be legally accepted.

As the main problem today is the identification of relevant but yet unknown content, we suggest utilization of the "collective knowledge" of users to identify relevant un-

[5] This not only protects customers against illegal content distribution, but also against potential software bugs.

known content to other users. The field of semantic web technology covers this area of study in which there are already developments in building distributed semantic webs based on P2P which can be combined with the proposed distribution framework [23, 25, 3].

CONFUOCO allows the distribution of encrypted (DRM-protected) as well as (temporarily) free content. It provides control on the exchange of content and easily allows the uploading and distribution of that content. Different business models can be built upon this framework (e.g. the Potato-System or the Weed system) and other systems, such as the Music2Share approach [21] can be implemented within this design. Rights collecting societies and associations can also benefit as exchange data can be used in calculating the distribution of their fees.

This solution is not only relevant for musicians interested in distributing their content, but also for Internet Service Providers (ISPs) that wish to increase data transfer levels. Besides the distribution of content information gained from user activities (like the files exchanged) are valuable in identifying new trends or to verify the marketing strategy of a record label.

This solution is of course just the beginning and further improvements are possible. In addition to the decentralization and distributed TTPs, privacy can be improved and an increased level of anonymity can be reached [22] (for example with "throw away" temporary IDs or multiple IDs for each user). Nevertheless CONFUOCO provides the mechanisms demanded by numerous musicians like Courtney Love who requested help in addressing fans: *"I'm looking for people to help connect me to more fans, because I believe fans will leave a tip based on the enjoyment and service I provide. I'm not scared of them getting a preview. It really is going to be a global village where a billion people have access to one artist and a billion people can leave a tip if they want to ... Offer some control and equity to the artists and try to give us some creative guidance. If music and art and passion are important to you, there are hundreds of artists who are ready to rewrite the rules."* [1].

This is exactly what CONFUOCO does: providing the needed means for connecting musicians and fans while offering some control and equity to musicians.

References

[1] Courtney love's speech to the digital hollywood online entertainment conference. http://www.cdbaby.net/articles/courtney_love.html, May 2000.

[2] Apple itunes music store. http://www.apple.com/itunes/, December 2004.

[3] The dbin project. http://www.dbin.org/, dec 2004.

[4] Dijjer - an open source p2p web cache for large files. http://dijjer.org/, dec 2004.

[5] The helix platform. https://helixcommunity.org/, December 2004.

[6] Kontiki delivery management system (dms). http://www.kontiki.com/, December 2004.

[7] Light weight digital rights management. http://www.lwdrm.com/, December 2004.

[8] Microsoft windows media - digital rights management. www.microsoft.com/windows/windowsmedia/drm, December 2004.

[9] Peer impact. http://www.peerimpact.com, December 2004.

[10] Potato system. http://www.potatosystem.com, December 2004.

[11] Real player music store. http://www.real.com/musicstore, December 2004.

[12] Red swoosh. http://www.redswoosh.com/index.php, December 2004.

[13] Weed. http://www.weedshare.com, December 2004.

[14] M. Arnold, M. Schmucker, and S. D. Wolthusen. *Techniques and Applications of Digital Watermarking and Content Protection*. The Artech House Computer Security Series. Artech House, Norwood, MA, USA, 2003.

[15] E. Batlle, J. Haitsma, P. Cano, and K. Kalker. A review of algorithms for audio fingerprinting, Jan. 30 2003.

[16] E. Batlle, E. G. Omez, M. Bonnet, P. Cano, and R. D. C. T. Gomes. Audio fingerprinting: Concepts and applications, Jan. 23 2003.

[17] Biddle, England, Peinado, and Willman. The darknet and the future of content protection. In *ACM CCS Workshop on Security and Privacy in Digital Rights Management, LNCS*, 2003.

[18] D. Brookshier, D. Govoni, N. Krishnan, and J. C. Soto. *JXTA: Java P2P Programming*. Sams, 2002.

[19] I. J. Cox, M. L. Miller, and B. Bloom. *Digital watermarking*. Morgan Kaufmann series in multimedia information and systems. Morgan Kaufmann Publishers, San Francisco, CA, USA, 2002.

[20] Herre. Content based identification (fingerprinting). In *ACM CCS Workshop on Security and Privacy in Digital Rights Management, LNCS*, 2003.

[21] T. Kalker, D. H. J. Epema, P. H. Hartel, L. Lagendijk, and M. van Steen. Music2Share - Copyright-Compliant music sharing in P2P systems (invited paper). *Proceedings of the IEEE Special Issue on Digital Rights Management*, 92(6):961–970, June 2004.

[22] A. Kobsa and J. Schreck. Privacy through pseudonymity in user-adaptive systems. *ACM Transactions on Internet Technology (TOIT)*, 3(2):149–183, May 2003.

[23] A. Naeve, B. Wolf, C. Qu, M. Palmer, M. Sintek, M. Nilsson, S. Decker, T. Risch, W. Siberski, and W. Nejdl. EDUTELLA: P2P networking for the semantic web, Jan. 24 2003.

[24] J. Oostveen, T. Kalker, and J. Haitsma. Feature extraction and a database strategy for video fingerprinting. *Lecture Notes in Computer Science*, 2314:117–128, 2002.

[25] G. Tummarello, C. Morbidoni, J. Petersson, P. Puliti, and F. Piazza. Rdfgrowth, a p2p annotation exchange algorithm for scalable semantic web applications. In *P2PKM*, 2004.

An Incentive Based Distribution System for DRM Protected Content Using Peer-to-Peer Networks

Hariharan Rajasekaran
Technische Universitaet Muenchen
Lehrstuhl fur Datenverarbeitung
D-80290 Munich, Germany
email: hariharan.rajasekaran@gmail.com

Abstract

Content owners are adopting various strategies to distribute DRM protected digital content to the consumers in their efforts to reduce unauthorised copying and sharing of content. In addition to offering direct downloads from online web portals, content owners are now trying to co-opt the peer-to-peer (P2P) networks to distribute digital content to the consumers in a cost efficient way. In this paper, a distribution model based on P2P networks, offering monetary incentives is presented. The monetary incentives are dependent on the amount of data uploaded by the consumer to the other peers in a network. The paper discusses why this incentive model is different from other similar schemes proposed in the literature and analyses the process of designing incentives for DRM protected content. This paper also discusses the economics of such a business model and presents an architecture to implement such a scheme.

1. Introduction

The content owners' reluctance to offer alternate methods for digital content distribution and consumption, during the early years of the Internet, lead to the emergence of file sharing networks like Napster. These networks attracted a large number of users resulting in mass distribution of unauthorised copies of copyrighted digital content. After shutting down services like Napster through legal proceedings, the content owners are currently trying to develop new business models for selling digital content on the Internet. Some of the strategies adopted so far are direct downloads from online digital stores (e.g. Apple iTunes) and subscription based services (e.g. Napster-to-go, Yahoo Music). Recently content owners have turned their attention to P2P networks to distribute content. The recent launch of Snocap [17], which is a licensing service that can be integrated to existing P2P networks to track copyrighted content, shows the industry's interest in using P2P for distributing digital content.

In this paper, a distribution model for DRM-protected content based on P2P networks that offers monetary incentives to consumers is presented. In this model, the consumers are awarded monetary incentives based on the amount of data they upload on behalf of the content owner. Though various incentive schemes are discussed in the literature [6, 15, 18, 11] to counter the so called "free rider" problem in P2P networks, the arguments that they use for providing incentives are valid only when the content that is being shared can be obtained for free and do not hold in case of DRM-protected content. This paper outlines the design of an incentive scheme which can be used to promote the sharing of DRM-protected content on a P2P network and discusses the economic aspects of such a scheme and argues why such a scheme would be useful for content owners.

2. Economics of P2P distribution

This section gives a brief overview of the different incentive schemes proposed for content distribution via P2P and discusses the differences between distributing "free content" and "DRM protected" content via P2P systems. The section also discusses why it will be advantageous for the content owners to distribute DRM protected content given the recent proliferation of broadband Internet connections and high capacity digital players.

2.1. Incentive schemes for P2P distribution — An overview

One of the major problems in a P2P network is the unwillingness of the participating peers to share content with other users on the network. This is called the "free rider" problem and this behaviour of the peers has been exten-

sively studied and analysed [3]. The problem arises when the majority of the peers download more amount of data than what they make available for others to download from them. The P2P network in this case is sustained by a few altruistic peers who share their content on the network. The "free rider" problem drastically reduces the effectiveness of a P2P network because, it strains the bandwidth resources of the few peers who share content, and it also reduces the variety of files available for download on the network. To overcome this problem various incentive schemes have been proposed in the literature [18, 2, 5, 7] and recent P2P architectures include the feature where a user simultaneously uploads the content that he is downloading and his download rate for that particular content depends on the simultaneous upload rate (e.g. BitTorrent, eDonkey etc). But even this in-built feature of the architecture is not sufficient to reduce the "free rider" problem as shown by [2]. Hence alternate incentive schemes are needed to improve the performance of the P2P network.

An incentive system which rewards users with increased download speed when they offer improved service by increasing their upload speed is discussed in [2]. A game theoretic model for analysing incentives is given in [5]. [7] defines a utility function of a user in a P2P network and analyses the strategies that a user will take to maximise his benefits by participating in the network. [18] describes an incentive scheme based on increasing or decreasing the value of a single variable called "KARMA". The user is awarded positive KARMA for sharing resources and negative KARMA for downloading resources. The user is not allowed to download content if he does not have enough positive karma. The arguments that many of these schemes use for providing incentives are based on the following assumptions:

- Users incur bandwidth costs when they trade files via P2P networks.
- Hard disk space costs money for the user.
- Users share content on the network for which they might or might not have paid to acquire in the first place.
- Users downloading the content do not incur any other cost other than those related to bandwidth and disk space.

The incentive schemes assume that the user will try to minimise his cost for sharing by reducing his upload rate and hence offer incentives to make the user share more. In the following sections, the economic advantages of distributing DRM-protected content on P2P networks for the content owner are discussed and the new problems arising in designing incentives for such DRM protected content are analysed.

2.2. Case for distributing DRM protected content via P2P

A significant trend seen today in the consumer electronics devices is the emergence of high capacity media players such as Apple's iPod, Dell Juke Box etc. A high end iPod today has a 60GB hard disk and can store up to 15,000 songs. At $0.99 a song, a user has to spend approximately $15000 to fill his iPod which is very unrealistic. Some services like Napster-to-go try to fill this enormous capacity on the hands of the user with subscription based services which give access to the complete music catalogue for a limited period. But the success of such schemes are yet to be determined given their recent launch. Recent surveys in Europe conducted by INDICARE [9] show that about 80% of the users still prefer to download their songs rather than "rent" them for a certain time period. The survey also indicates the widespread use of P2P services by users to acquire music. Hence it is safe to assume that given the rapid adoption of high capacity players and a rather steep cost of filling those devices with digital content, users will resort to downloading unauthorised content from P2P networks. The content owners can prevent this scenario by either resorting to expensive and difficult to implement legal and technical measures or offering downloads at cheap prices, for example lets say at 10 cents per song which is equivalent to the price of a SMS sent on a mobile phone.

It has been reported that even with $0.99 per song, stores like Apple's iTunes are making a loss and it recovers this loss by selling iPods at a premium [14]. Not many content distributors can afford to have such business models and they must find a way to distribute content in a cost efficient manner. The two main contributions to the cost of a digital file are the royalties paid to the content owners and the cost of distributing the file to the consumers. The amount of royalties collected depends on the content owner and is outside the scope of discussion of this paper. But considering the market situations, it is expected that the content owners will benefit by looking at a collection model where lower rates are matched by increased compliance. To increase the attractiveness of DRM protected content, entertainment companies are looking at various business strategies and one idea that is being debated in the industry is releasing movies online at the same time as their theatre release [8]. No online portal can cope with the bandwidth requirements needed in such a scenario in a cost efficient manner and P2P is the only viable solution available to the content distributors to keep costs down.

Maintaining an online portal that caters to a huge number of users is a very cost intensive process even with small sized music files. The cost increases dramatically when high quality video files are involved. This is because an Internet service provider (ISP) charges content providers

based on the amount traffic that their websites deliver and do not offer them flat-rate connections like the ones offered to the end users. To illustrate the amount of money involved in online distribution, we consider the following example. The typical cost of hosting a site with a bandwidth of 30GB per month is approximately $10 [13]. Using this bandwidth a DVD movie averaging 4 GB can be downloaded by around 7 users. Thus it costs approximately $1 to distribute a movie file to a single user without including the costs involved in developing and maintaining the web portal concerned. At this rate, the distribution costs would run into millions when millions of users are involved. Even when a high volume user such as a movie distributor can get substantial discounts in these charges, the cost involved is nevertheless high. In this paper, we try to minimise the distribution costs associated with distributing digital content and propose an incentive based distribution scheme that tries to achieve this goal.

2.3. Designing Incentives for DRM protected content

The changing technological landscape has made some of the arguments for providing incentives to share content obsolete. With the introduction of high-speed flat rate broadband connections and high capacity hard disks and DVDs, the news rules of the game are:

- The bandwidth costs for the user is fixed even if he is not participating in a P2P scheme because of the flat-rate nature of his connections.
- Storage space cost will be negligible with the arrival of High Density hard disks and DVDs.
- If we consider the ideal situation where only DRM protected content is shared on the network, it is safe to assume that the user must have paid the content owner to acquire the content and hence the content is not "free".

Therefore, the major reason for the users not sharing content now is "selfishness", since there is almost no cost involved in sharing. When this is the case for content that can be downloaded for "free", persuading users to share DRM protected content on a P2P network will be even more difficult, considering the fact that the only person who gains an advantage here is the content owner.

Also, distributing DRM-protected content through P2P networks presents a different set of problems in addition to "free riders". Sharing DRM protected content on a P2P network is not very appealing to a user because

- The user has to obtain the license from the content owner to render the content.
- The user can always download such content from the content owner's online portal which does not suffer from low download speeds and fake files—the two most annoying problems in a P2P network.
- The user cannot use DRM protected files for *bartering*, which is a very powerful incentive in a P2P network, where people in possession of rare files can force other people to share their collections in exchange for the rare files. This is because by default a user can always get the DRM protected content from the content owner.

Hence the content distributors can gain a lot if they entice their users to share DRM-protected content thereby minimising their own distribution costs. Some networks like Altnet have implemented schemes where users gain points when they redistribute the so called "Gold files", which are premium content that can be previewed by the user before buying from the content owner [4]. Some papers have suggested the content owners and distributors to improve the search functionalities of P2P networks by providing superpeers that deliver high quality files relevant to the search criteria of the user [16]. In this paper, an architecture for an incentive based distribution scheme for DRM-protected content via a P2P network is presented. The users are offered monetary incentives based on the amount of data they upload on behalf of the content owner. Though the proposed model can be implemented on any P2P network, in this paper, we use BitTorrent type of P2P network to show how such an incentive scheme can be realised.

3 Basics of BitTorrent P2P system

The BitTorrent P2P system has the following critical elements [6]:

Web server provides a meta info file by HTTP. This file is called a ".torrent" file and has the details about the file and its pieces and their checksums. It also contains the URL of a Tracker website. The typical size of a torrent file is a few KiloBytes.

BitTorrent Client is the P2P client that downloads the actual content file using the .torrent file.

Tracker is a non-content-sharing node in the network and is used to track the peers in the network.

Peers are the end users in the network and are of three types:

1. Seeder — The user who has a complete copy of the file.
2. Leecher — The user who is downloading the content from the seeder.

3. Reseeder — A leecher who shares the content after completing the download.

The process of downloading a file is shown in figure 1. Bob, a content owner, creates the torrent file for the movie he wants to distribute online and posts it on a web server. This web server can be Bob's online movie portal. He also a sets up a tracker that can track the users sharing the movie file and makes the actual movie file available for download on a separate computer. The downloaders Ron, Jill and Alice download the .torrent file from the movie portal and use their client software to download the movie file. The client goes to the tracker and find that Bob is the only one who is sharing the file and starts the download from his copy. The BitTorrent protocol makes sure that all the downloaders get a different part of the file from the first seeder (which is Bob here). This way, Bob can send parts of the file only once and the downloaders can share their different pieces to build the complete file. The tracker helps in this exchange process by directing the peers to one another.

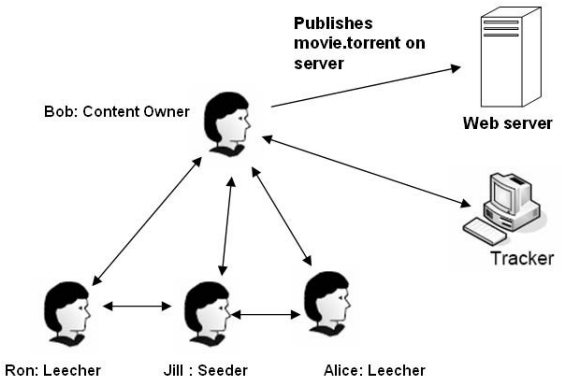


Figure 1. BitTorrent Elements

The BitTorrent system's use of the tracker to co-ordinate the different peers on the network provides an attractive feature for the implementation of an incentive based P2P system proposed here. This is because, every peer in the network calls back "home" every few minutes to search for new peers and update their status on the tracker. This call back feature can be modified to provide a number of other features that can aid in distributing DRM-protected content and implement the proposed model, which is described in the next section.

4 P2P system based on monetary incentives

The BitTorrent P2P network is an efficient way to distribute very large files like movies with very low distribution costs for the content provider. The distribution model proposed here offers monetary incentives to users who share their DRM protected content on the network. For the purpose of explaining the model, we consider a situation where a movie studio decides to open up its entire movie catalogue for online distribution and also plans to release new movies online at the same time as their theatre release.

4.1 Architecture of the proposed model

The proposed model uses the basic BitTorrent protocol and adds additional features on top of it by means of proprietary plugins that handle content owner specific features. The model can be implemented by having the following components

1. BitTorrent Client — The content owner can develop a customised GUI based client that implements the BitTorrent protocol. Another option would be to use existing clients such as Azuerus, BitTornedo etc.

2. Incentive and Management Plugin (IMP) — This is a special software that must be developed by the content owner and is proprietary. The IMP is responsible for keeping track of the data traffic of the user to award compensation. It is also used for billing the user and to retrieve pricing information for content and other information from the content owner. It can also offer search functionality to the user by retrieving content torrent files matching the user's search criteria from the content owner's online database. Another important function of the IMP will be to retrieve the licenses (or Rights Objects) needed to access the protected content. The IMP can offer a list of choices for the license to the user and retrieve them based on the user's choice. The IMP also makes the task of sharing files on the BitTorrent network easy for the user by creating/downloading the relevant torrent files for the shared content.

3. Web server — This is the online portal of the content owner that hosts the torrent files. The user can download these files into the client software to download the actual files.

4. User account — Every user needs to have a user account with the content owner and has to log into that account via the plugin when he uses the client software to download/upload files. Using the account a user can buy credits/points that can be used by the plugin to pay for content that is being downloaded. The credit/point statistics can either be maintained online at the content owner's end and updated by the IMP plugin or can be maintained locally at the client side. In the latter case, the IMP is responsible for securely maintaining this store.

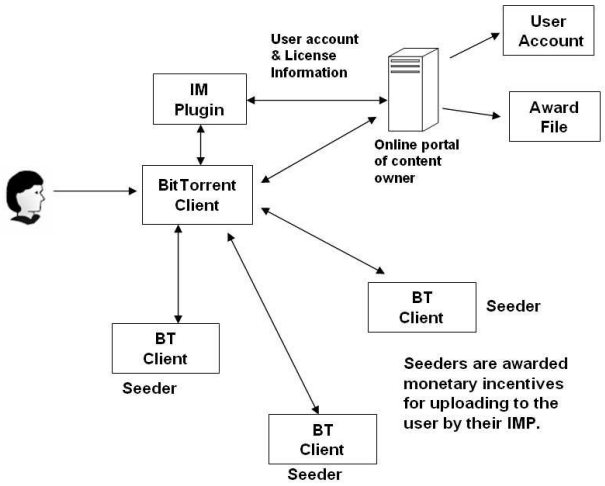


Figure 2. BitTorrent with monetary incentives

4.2 Controlling incentives dynamically

In the proposed system, the content owner can change the incentives offered for sharing particular content files. This is done by maintaining an "award file" that lists the current number of points awarded for sharing particular files. The IMP plugin downloads this file and awards points accordingly. The "award file" has a particular validity period and the IMP plugin has to download a new file once the current one expires. The content owner is able to track the upload rates for files using the tracker and when the demand for a particular file rises (indicated by the rise in number of peers), he can increase the incentives for sharing the file for the seeders so that the transfer rate for that particular file is increased. The change in the award information can be "pushed" to the IMP plugin and the user can change his uploading pattern if he wants to reap more reward points.

In a BitTorrent system, the leechers also upload parts of the same file to other users when the file is being still downloaded. To keep the model economically viable for the content owner, in the general case, monetary incentives are given only to the reseeders. But this rewarding mechanism can be changed dynamically when a new movie is released online by the content owner. In this case, only the content owner has the full copy, and all leechers try to limit the upload bandwidth of the file being downloaded to increase their download rate. This results in the reduction of the overall download rate of the file and the content owner's server is overloaded. Hence for this special case, incentives can be awarded to leechers as well, to increase the overall download rate of the file being distributed. The user can either redeem these incentives as points to buy further content from the content owner or can get his account credited with money.

5 Example Scenarios

This section describes some example scenarios to explain the distribution model. It is assumed here that the user has a user account with the content provider and has bought credits/points. The user also has the proprietary plugin that is integrated with his BitTorrent client.

5.1 Scenario I - Downloading movies and songs online

In this section we consider the case when a user wants to download movies from the catalogue of a movie studio. The user starts the BitTorrent client and the IMP plugin asks for the user's login information. Once the user is logged in, the IMP plugin has all information about the user's previous purchases and the amount of points in his account. The user can use the search functionality to search for movies that he wants to purchase. The plugin returns the relevant torrent files from the studio's catalogue. When the user indicates his choice the plugin prompts him with a message displaying the amount of points needed for the different types of licenses associated with the content. The user confirms his choice and the plugin checks to see if the user has enough credits to complete the purchase and starts the download.

The points in the user's account are reduced once the user chooses to download a particular file. This is to deter users from collaborating to help each other to gain extra points by downloading files without paying the content owner [7]. The download status of the file is logged in the user's computer. This is to ensure that when downloads are resumed after a break, the credit points are not reduced for resuming a current download. This feature can also be used to prevent collusion attacks described above. The user's who are reseeding the file on the network receive bonus points depending on the amount of data they upload. The amount of bonus points depends on an "award file" kept at the content owner's end. The content owner can change this file to reflect the demand conditions for particular files and award points accordingly. When the content file is downloaded completely, the IMP retrieves the appropriate license from the content owner and updates the user account reflecting the purchase. The same procedure is followed in case of small sized files such as huge song collections. When a user has a considerable amount of content that is paid for, this scheme allows him to earn a small amount back which can help him buy more content which benefits the content owner's bottom line.

5.2 Scenario II - Releasing new movies online

In this scenario we consider the case where a movie studio wants to release a movie simultaneously in the theatres and on the Internet. The studio has the option of releasing the movie a few days ahead of the release date online on its website in a protected format and sends the licenses on the day of the release. But this puts a lot of load on its website and there is a very small probability that someone might find a way to unlock the content. If this happens, the loss for the studio will be enormous. Instead, the proposed scheme provides a cost effective method to distribute such content on a large scale in a very short period of time. The typical download time for a movie in a BitTorrent network is normally around 3 to 4 hours when sufficient seeds are available and hence the studio can make the file available just hours before the theatre release.

In this scenario, the studio is the initial seed for the content. The usual download pattern in a BitTorrent network when a new file is introduced is for the users to restrict their upload speeds to gain better download rates. But in case of a new file with only one copy available, this strategy reduces the overall download rate of the file. To some extent this problem is overcome by using the superseed strategy [1] in BitTorrent and making available multiple seeds to begin with. Alternately, the studio can alter the award file to award bonus points to leechers and reseeders. This will provide more incentive for the users to open up more upload bandwidth and hence increase the overall download rate of the file. The changed bonus point status for a file is shown to the user by the IMP plugin so that user can take advantage of the offer. The cost involved in such a points scheme will be much less than having a dedicated file server catering to millions of download requests.

It must be noted that all content that is shared on the network are DRM protected and it is assumed that the user has proper devices that are capable of rendering the content and enforcing the licenses associated with the content.

6 Security of the system

Security is a main criterion when high value digital contents like movies are distributed via P2P networks. Though the content that is distributed in the proposed system are all DRM protected, which means that the content is encrypted and can be accessed only with the appropriate license files, the system does not explicitly prohibit sharing non-DRM protected files. Hence the user can upload files which have their DRM protection removed. The recent legal proceedings at the US Supreme court involving file sharing networks like Grokster etc. did not find any problem with the underlying P2P technology, but it requires the P2P network designers to put safeguards into the system so that copyrighted content is not shared without the consent of the rights holder [10]. This section describes some of the security features in the proposed model.

6.1 Identifying unauthorised content

The digital files shared in the system have embedded digital watermarks and are encrypted by the content owner before being distributed to the users. The user can choose to use the P2P client without enabling the content owner plugin to share the content on the network. If the user shares the content in the encrypted form, there is no security breach and the user is not awarded any points since the plugin is not enabled. On the other hand, if the user shares a cracked version of the DRM protected file, the leak can be traced by the use of the embedded watermarks. When such leaks are detected, it is also easy to stop the file sharing in BitTorrent by shutting down the tracker website that coordinates the sharing process. This method was used to stop many online sites that offered trackers and torrent files for pirated movies [12]. The other scenario is when the user tries to share the cracked content when the IMP is loaded on the client. In this case the IMP first checks for the checksums of the file in case of DRM protected content to identify and award points. When it detects an unprotected content it checks for any embedded watermarks in the file identifying the content owner corresponding to the plugin. If it finds any, it will not allow the content to be shared.

6.2 Eliminating fake copies

One main problem in a P2P network is the availability of fake copies. The system eliminates this possibility by making sure that the torrent files corresponding to the content files are digitally signed by the content owner and the IMP checks this signature before using any torrent file to download content. The underlying BitTorrent protocol relies on checking the cryptographic hashes of every piece of the file that is being downloaded and hence if the torrent file is verified to be correct, the file that will be downloaded using that torrent will be authentic as well.

6.3 Attacks on upload statistics

There is a possibility that a user may succeed in hacking the proprietary plugin and use this hacked plugin to report wrong upload statistics to gain award points. In the proposed system, the points can be awarded either locally by the IMP based on the "award file" or the IMP can send the upload statistics to the content owner in a secure format who then credits the user's account. In the former case, the attack can be prevented by authenticating the plugin when the

user starts the client from the content owner's side and in the later case, the statistics can be analysed based on time stamps to see any unusual upload behaviours.

6.4 Attacks on the P2P client

The model suggests using open source BitTorrent clients to keep the costs down for the content owner. It is possible for a user to change the underlying code of the client to report wrong upload statistics to the IMP. To prevent this, the content owner can first review the code for a particular version of the P2P client for any such hacks and ensure the IMP works only with the executable file created using such a reviewed code. This can be done using cryptographic hashes and other software protection mechanisms.

7 Conclusion

In this paper, a business model and architecture for distributing DRM protected content in a cost efficient way was presented. The model enhances the basic model of a BitTorrent type P2P network to provide features for the content owners to distribute DRM protected content by offering monetary incentives to the users. The security issues involved in the model are also discussed. The paper also analyses why some of the reasons given for providing incentives to avoid the "free rider" problem in a P2P network will not hold when DRM protected content and flat-rate Internet connections are involved. The model presented here is just one of many ways in which the BitTorrent type of P2P network can be enhanced to distribute DRM protected content efficiently. The system can be improved if the user is able to use points bought at one vendor to buy another's content using a single IMP. But this will require the content owners to provide interoperability between their distribution and content protection mechanisms. The paper concludes by noting that selling DRM protected content to consumers, used to getting free content on the Internet, requires innovative marketing and business models and the content owners must experiment with new schemes to reach out to the consumers.

References

[1] Superseed mode. "http://www.bamber.org/wiki/ow.asp?SuperseedMode". Accessed on 10th August 2005.

[2] T. Ackemann, R. Gold, C.Mascolo, and W. Emmerich. Incentives in peer-to-peer and grid networking. *UCL Research Note RN/02/24*, 2002.

[3] E. Adar and B. A. Huberman. Free riding on gnutella. *Technical report, Xerox PARC*, August 2000.

[4] Altnet. "http://www.altnet.com/".

[5] A. Blanc, Y.-K. Liu, and A. Vahdat. Designing incentives for peer-to-peer routing. Technical report, available at "http://www.cs.ucsd.edu/~vahdat/papers/routing-game-tr.pdf.".

[6] B. Cohen. Incentives build robustness in bittorrent. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, USA, 2003.

[7] P. Golle, K. Leyton-Brown, and I. Mironov. Incentives for sharing in peer-to-peer networks. In *EC '01: Proceedings of the 3rd ACM conference on Electronic Commerce*, pages 264–267, 2001.

[8] S. Hansell. Plan to sell movie downloads online. "http://www.theage.com.au/text/articles/2005/07/07/1120704487032.html?oneclick=true.". Published 8th July 2005.

[9] INDICARE. Digital music usage and drm. "http://www.indicare.org/survey", May 2005.

[10] P. Jones. Grokster loses. "http://www.groklaw.net/article.php?story=20050627130232598". Published 27th June 2005 on Groklaw.

[11] H. T. Kung and C.-H. Wu. Differentiated admission for peer-to-peer systems: Incentivizing peers to contribute their resources. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, USA, 2003.

[12] T. Mennecke. Suprnova.org closes. "http://www.slyck.com/news.php?story=626". Published 19th December 2004.

[13] Mymarkdown.com. Web hosting prices. "http://www.mymarkdown.com/specials/101a.html". Accessed on 19th August 2005.

[14] A. Orlowski. Your 99 cents belong to the riaa. "http://www.theregister.co.uk/2003/11/07/your_99c_belong/". Published 7th November 2003.

[15] K. Ranganathan, M. Ripeanu, A. Sarin, and I. Foster. To share or not to share: An analysis of incentives to contribute in collaborative file sharing environments. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, USA, 2003.

[16] S. Singh, S. Ramabhadran, F. Baboescu, and A. Snoeren. The case for service provider deployment of super-peers in peer-to-peer networks. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, USA, 2003.

[17] Snocap. "http://www.snocap.com".

[18] V. Vishnumurthy, S. Chandrakumar, and E. G. Sirer. Karma : A secure economic framework for p2p resource sharing. In *Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, USA, 2003.

Business models for location based access to cultural heritage

L-F Pau , Prof.Mobile business and media ; B.Rees
RSM Erasmus University (Rotterdam School of management)
lpau@rsm.nl

E. Carta
Venis Spa

Abstract

This paper summarizes some findings of the EU eContent project LATCH www.latchproject.net about location based access to cultural heritage .It describes the information life cycle methodology used in a generic quantitative business model covering such cases One key reason for the modeling was to identify the flow of money, branding and IPR dependencies leading to the well known imbalance between cultural heritage collection and the low distribution revenues .In addition, the project has researched a number of alternative business models between the implicated stakeholders , which could be analyzed with the generic model .Two full cases from the project partners in Iceland and Italy served for the real validation of the analysis .

1. Introduction

The business analysis and modeling of cultural heritage raises some major challenges rooted in a number of structural issues. The first is that by definition,"cultural "products and services belong to the world of creation, performance, storage and archival and thus should not be subject to economic rules. The second is that the cultural stakeholders (from creators, to artists, museums, archival institutes) are highly fragmented and have usually begun the digitization of their assets as a means to reach out to others, cultural workers and people interested in culture alike. The third is that a virtual, communications-based society supplements one where people move physically to meet and discover, thus leading to demands across locations and different types of experiences. This paper reports some of the findings of the EU eContent project called LATCH, with the following focus areas:

- Surveying specific aspects and trends in location based cultural heritage digitization and exploitation, aiming at a sustainable business model for all stakeholders
- Documenting important business and technological aspects of a generic quantitative business model developed in LATCH , to support the creation of business plans and the customization of the quantitative analysis in specific cases
- Describing specific cases (Venice digital library, Icelandic manuscript and folkloric recordings), of which one provides enough basic data for a quantitative analysis
- Propose at a conceptual level, but validated by the quantitative analysis, a number of new innovative and sustainable organizational and funding schemes for digital cultural heritage using technology enablers such as LATCH.

But this analysis and modeling really matters most when being used to propose new approaches to the sustainable creation and development of cultural assets. They rely on business processes whereby all parties involved, from creators, archival institutions, to funding bodies (public and private), and culture users (individuals or researchers) find a balance between their desires, funding needs and demand for knowledge. This paper identifies a number of such processes, but also describes a modeling tool allowing for an economic analysis of cash flows, investments, and intellectual property rights across all stakeholders, including some critical emerging players represented by communications operators, value-added

information providers, digital repositories, and also private cultural foundations.

2. Cultural Information Life Cycle Management

This section documents economic, business and technological aspects of the generic quantitative model developed to support the business plan creation and the customization of the quantitative analysis in specific cases

The LATCH business model is built using the concept of information life cycle management (ILM). The Storage networking industry association (SNIA) data management forum (DMF) is harmonizing this definition. ILM is a strategy not a specific product. It comprises policies, processes, practices, and tools used to align the business value of information (in LATCH: cultural information) with the most appropriate and cost effective infrastructure from the time information is created through its final disposition or archiving. In short, ILM is a strategy by which storage and reformatting resources are allocated depending on the business value of the data being stored. The various information management stages can be categorized into information classification (by ontology such as the LATCH ontology and file types), information policy (including DRM), information management (in LATCH repository and semantic Web) and tiered infrastructure (such as storage systems and value added distribution e.g. wireless).

3. Generic location based cultural heritage business flow and modeling tool

This section describes the LATCH generic business-modeling tool for cultural heritage information accessed from various locations. This includes gathering: awareness information from anywhere, pre-visit trip preparation information from the home or work place, live experiences and wireless information on-site during a site visit, post-visit information via media and the Internet, and finally research data by researchers at their place of work. As the sustainability can only be analyzed if all major stakeholders are accounted for, the list therefore includes: creators, archival institutions, post-production and digitization, repository databases (XML) and software tools, value-added information providers, billing and digital rights management agents, communications transmission operators. It obviously also includes the central state, local government, cultural private foundations, cultural site owners, cultural site operators, media and local publicity , all parties modeled in the tool . A very important aggregated set of stakeholders is made up of all the derived and collateral beneficiaries of the existence and localization of cultural heritage. This includes; hotels, tourist shops, restaurants, media shops, local transport, etc. Analysis reveals that these parties benefit in a major way from cultural heritage but contribute very little to its sustainability, even when tax payments (VAT) are taken into consideration. The modeling methodology for the tool is one of cultural information life cycle management Section 2) and business flows. The model itself gives the key flows of cash flow, investments and intellectual property rights between all above stakeholders, with the assumptions made.

The following paragraphs describe how essential parts of the LATCH quantitative business model have been built up or the assumptions the tool depends upon.

3.1 Business model flows

Three budgets are maintained regarding one cultural site and a low number of related cultural repositories: a longer term investment budget, a short term 1 year operational budget focusing on cash flows between stakeholders, and a copyright and intellectual property right (IPR) balance. Three types of flows are considered: operational income or expense, investment income or expense, usually used for longer term investment purposes, and IPR rights. These budgets are maintained in the tool with book balancing within each of the investment and operational budgets, that is a revenue from a stakeholder is normally a +, while an expense to a stakeholder is normally a -. IPR rights are split into assignor/owner, with normally "+" labels, and assignees/licensees with normally a "–" label.
No account is taken of the balance sheet between assets and liabilities in this analysis, where only operational cash flow between parties is studied. Consequently, any investment budget surplus linked to the stated cultural activities is assigned as an excess internal operational surplus.

3.2 Cash flow balancing principles

Essentially all public bodies, cultural site owners and operators, as well as private foundations, must operate on a neutral cash flow basis, without cash flow surplus or deficit. Local authorities must balance; if not their deficit or surplus is sent to public authorities, as they set the VAT and other tax rates .Cultural site

operations must always balance, with publicity adjusted accordingly. These rules are implemented in the tool.

3.3 Value added enhancement operators

The value added enhancement operator (VASP) functions as a cultural information roaming and clearinghouse, as well as a secondary information enhancement agent. It performs:

- Multi-language conversion
- File type conversion and formatting
- Transforms requests in a visited cultural network to a request in a home cultural network

It does not do cross settlement for transmission services or content rights, all of which are handled by an independent DRM and billing agent.

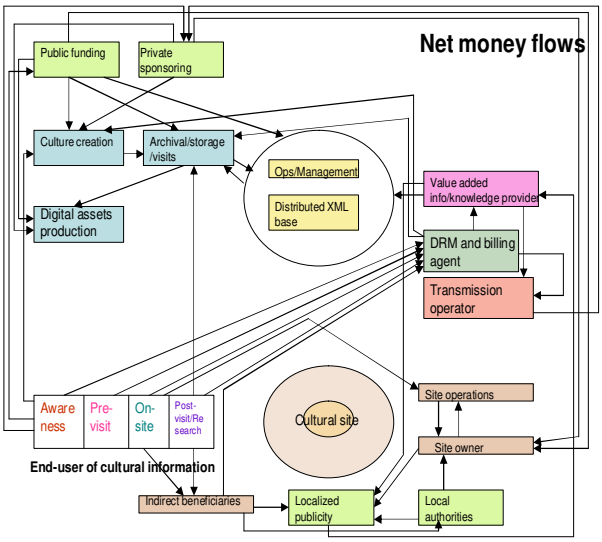


Figure 1: Business flows in cultural heritage information distribution

3.4 Some content charging assumptions

All awareness content is free of charge. Internet access is charged ultimately by telecom operator/ISP, with no share for the VASP, transparent through the billing agent.

Pre-visit content is enhanced by the VASP. The VASP owns all the dependent enhancement rights. Out of total access and content charge billed by billing agent, a part goes to the operator to cover his transmission costs, while the remaining part (set by a splitting ratio) is retained by the VASP. The VASP though must provide a portion of revenue to the cultural creator who owns the primary rights.

All the on-site WLAN access and content income collected by the billing agent is returned to the operator, who has adapted the content to WLAN distribution. The VASP enhances the content. The ratio split is set so that the VASP is rewarded adequately by the operator for enhancements made to the content.

3.5 Effect of publicity

The propensity of users in different phases of a cultural cycle varies depending on the total spending on publicity by the local publicity agency .Furthermore, it is assumed that end users can receive push publicity during a value-added IP session as well as during a GPRS session. They are not charged for data traffic pushed, ultimately borne by the publicity agency, after a split between the transmission operator and the value-added provider (who owns the corresponding user sessions). In this way, the LATCH business model is one of the very few mobile business models to account directly for both mobile advertising and for the effects of multi-access value added content

3.6. Price Elasticity of Software Licenses and their Overall Effect

This effect is to acknowledge the fact that access to cultural heritage, for most by non-subsidized users, is subject to tight spending restrictions including software. Price elasticity is in-built into the tool ; one is for the client software, in direct relation to end user spending ceiling on this category, and the other for server software in relation to systems professional middleware (in that value added suppliers may decide to invest in other application sectors instead).

4. Innovative business models

This section surveys a number of alternative and often complementary routes to change the imbalance identified and modeled in Section 3, using some key technologies (wireless networks, XML ontologies, life cycle management) and a number of LATCH business innovations. The first is about changing the operating rules for private foundations, which should become less regional and national, but begin competing with states in their initiatives and fund collection as cultural multinationals (as some museums are already starting to do).

The second class of LATCH business innovations are linked to new ways of users and researchers pulling cultural content with associated publicity, as opposed to the traditional mobile advertising push

models which are very detrimental to knowledge discovery and privacy in cultural interests.

The third class of LATCH business innovations pertain to the realization that cultural assets are often not branded and even less covered by commercial law attributes such as trademarks, leading to the notion of cultural trademark or "culture mark". Fueled by branding and such trademarks, soft approaches need to be exploited in building up trust when cultural heritage is not just experienced via visits or performances, but in a virtual experience context where there may be distrust of channels and cultural actors.

The fourth class of LATCH business innovations rests on the power, mostly of wireless ubiquitous communication, in eliciting communities and knowledge sharing. It is proposed that access to cultural heritage content via modern communications be bundled into cultural interest communities via membership schemes offering these bundles. In this way sustainability can be created and defended via direct stakeholders - the members. They also get access privileges to cultural knowledge and bulk rates on the various overheads such as storage, software, communication tariffs, visitors' fees, media, etc. Finally, this section shows how the LATCH technical and business innovations together can find their way as much needed components for personalized cultural tourism IT platforms - to mobile games. The report also points at extending by bundling the universal service provisions found in many communications laws and regulations about basic services (including Internet), to access to publicly owned cultural heritage content for education and research.

5. Business case: LATCH Icelandic cultural heritage content

To validate the LATCH analysis and business model, two sets of cases were collected by two sets of LATCH project partners. One set pertains to the establishment of digital libraries across Italy and in Venice, mostly for audio content. The second concerns the digitization, the distributed XML archiving software and the distribution of ancient Icelandic manuscripts and of some folk tale recordings. This second case is fully supported by a business analysis, tariffs, latent demand estimation, and the full location based life cycle management and access to this heritage information.

5.1. Description

LATCH includes links to the resources of the Arni Magnusson Institute collection. This XML catalogued collection comprises nearly 200'000 pages and 250 hours of audio data. The XML catalogued collection so far is made of 143 vellum manuscripts plus 102 other manuscripts representing an approximate total of 200 000 JPEG images of approximately 1 Mbyte. The collection also includes 250 hours of audio data.

Stakeholders: As the Arni Magnusson Institute collections are preserved at public expense, and are itself of great value to a number of other institutions, this case must first identify stakeholders in turning this cultural heritage to a business proposition:

-The researchers at the Arni Magnusson Institute, who have analyzed and even collected (in the audio case) the material and would like their work to be better known about the icelandic and nordic culture of those days.

-The institutional stakeholders in the Institute ,i.e. the University of Iceland and the Ministry of culture, who would want a "return on investment" in terms of research assets and global interest in icelandic culture; they represent in essence students and young researchers on one hand, and the general public on another hand

-The indirect beneficiaries of the above efforts, such as the Tourist Board , cultural associations in Iceland and elsewhere (due to emigration) specializing in icelandic culture ,and possibly some tourism agencies having special tours complementing nature tours by cultural tours

-Software and integration companies, such as Raqoon, who see a business opportunity in promoting Semantic Web services and XML technologies in servers and client devices

-Data warehousing sites, in Iceland or elsewhere, which could perform the service of storing the content

-Media companies possibly interested in special works of art (from radio broadcasts to special handbooks/guides) to be able to assemble changing and interesting parts of the collections by thematic interests or by physical locations

-Possibly some smaller villages would want to show to their rare visitors what content (audio or manuscript references) were collected in the neighborhood.

Business Parties: Since the content is based at the moment on Icelandic manuscripts and audio data, the business targets at this time are:

- Mobile operators in Iceland, of which there are two: Siminn and Og Vodafone; their business benefit would be in using the LATCH case as a first instance of wireless access to cultural content; once this trial would be in place, and users be identified, it would be easy to add other collections
- Other cultural Institutes in Iceland wishing to make their content more widely applicable. A list is provided by the LATCH project.
- LATCH partner Raqoon, as a software technology product (server and client) supplier, and possible integrator; here again the experience gained via the LATCH feasibility project could be extended to other collections and domains

Given the completed nature of the demonstration software, it could be made operational within a matter of weeks for a trial.

5.2. LATCH Icelandic case business analysis

The LATCH business model (Section 3) allows to make some estimates ,allowing to determine a sustainable equilibrium for all stakeholders based on some latent end user count and on spending pattern assumptions .This end user behavior in turn affects public investment as well as foundation based investments in said cultural heritage, which again determine the provisioning cost and tariff levels .The derived actual demand then determines site visits, revenues to indirect beneficiaries, and thus tax revenue to public bodies and publicity investments around the brand created by the cultural heritage.

Latent Demand: The propensity of users to access the Icelandic manuscript and audio recording content at different phases would be estimated as:

-in awareness phase (media, broadband): 26 %

-in visit preparation phase (broadband): 100 % if content is on line

-on-site visits to either museum holding the manuscripts or sites where audio recordings were made (wireless information, physical visits): 39 %

-post-visit research (broadband): 68 %

-research mostly at the Institute and affiliates (physical visits): 5 %

Has been determined a total number of latent users of 10 000 persons globally. There would be at most 4,7 instances per year of access by either means mentioned above and in either phases , representing an affordable bundle for end users in view of their spending pattern on cultural virtual goods . Media purchases would be very limited. Beyond communications subscription access fees ,actual users would at most spend 50 Euros/ year on that type of Icelandic content , split between content access charges, donations to foundations and/or membership in cultural bodies involved in the same type of heritage.

Value added service providers: Driven by the split in content access types, and especially by the propensity to use wireless content, location information, and value added Internet content (such as research reports on the folkloric songs), it is determined that revenue/year for all value added service providers is around 333k Euro / year .The average operational profit margin would be about 50 %.

*Operator :*It is determined that operator would derive transmission revenues of about 50k Euros specifically from such a content distribution , with an operational profit margin of 90 %.Besides , subscription fees may benefit operator much more significantly ,to the extent that all value added service providers only offer access via that one operator ; value of needed subscriptions has a value of 702 kEuros/year .Also provisioning for access to other content is probably the largest upside element for the operator .

*Public authorities :*Would be expected to fund this cultural heritage, by a diversity of means and channels (all taken into account ,including losses on site visits and maintenance) .However ,due mostly to derived income , especially tax and VAT ,they would still recover this investment (actually : positive profit margin of a few percent)

Private foundations: Are assumed to contribute with about 41k Euro to the same cultural heritage, for tasks falling under their prerogatives.

Content creators : In the specific instance of the stated Icelandic heritage ,there is no royalty identified in for the manuscripts, but there may be for the audio recordings (either to the folk story tellers or to the parties making the recordings, incl. the Institute) .The rate is set at 15 % for all end user and commercial access (excluding research) But obviously the content creators and bodies such as the Institute ,are the biggest budgetary beneficiaries of the LATCH Icelandic content life-cycle content distribution, as they receive the bulk of public funding plus additional revenues .,for a total of 721 kEuros/year ;they do have to spend though on digitization , archival, technology services ,etc .

Publicity: Besides publicity by the museum or body storing the audio recordings, the value added service providers and indirect beneficiaries are supposed to contribute modestly to the publicity needed for the public awareness around this heritage, at the rate of 2 % of their related income, for a grand total of 97 kEuros/year.

Technology provider: Would derive total revenue of about 240 kEuros/year, of which 195 kEuros from software licenses, and an operational profit margin of about 80 % (taking R&D subsidies into account). It is here assumed that the price for the end user access client software (XML Semantic Web access, fixed and wireless) is 49 Euros, and that value added service providers would have an annual license fee of 1000 Euros each for the server .It has been shown that the high client software price is actually an hinderance to demand .The technology providers is also deriving storage service and indexing revenues

Cultural site owners: In the case of the stated Icelandic content, site owners are mostly the museum holding the manuscripts, plus some local places linked to the folk tales .Their income is assuming an entry ticket of 10 Euros, with a reduced propensity after the first visit due to the static nature of the collection. This highlights that museums with new displays and collections have a major advantage in getting returning visitors funding all their collection's fixed costs .The total annual site owner budget is assumed to be 1,4 MEuros, with losses picked up by public authorities .

Indirect beneficiaries: such as hotels, publishers, guides, museum cafes, would see their revenues increase by 1,2 MEuros/year due to the flow of physical visits and content sample purchases.

References

[1] Alderuccio D., Bordoni L., Loreto V., - "Data compression approach to monolingual GIRT Task: an agnostic point of view" Working Notes for the CLEF 2003 Workshop, 21-22 August, Trondheim, Norway
[2] Berlecon Research Workshop "Mobile Stadtinformationsdienste", Berlin, 14 Jan 2000 www.berlecon.de
[3] Capra F., "The Hidden Connections" Redwing Book Co., Taos, NM, USA, 2001
[4] Castells M., - "The rise of the network society: economy, society and culture Vol I, 2nd ed" - Blackwell Publishing, Oxford, UK, 2000, 624 p., www.blackwellpublishing.com
[5] Elkington J., - "Cannibals with Forks" Published by New Society Publishers, Gabriola Island, BC, Canada, 1998
[6] Dr. Robinson., - "The Natural Step Story" New Society Publishers, Gabriola Island, BC, Canada, 2002
[7] Eccles R., Nohria N, Berkley J, "Beyond The Hype" Beard Books, Frederick, Maryland, USA, December, 2003
[8] Flores F., Winograd T., "Understanding Computers and Cognition", Addison-Wesley Professional, Boston, MA, USA, 1st edition January 1, 1987
[9] Hinze A., Voisard A., "TIP, Location and time based information delivery in tourism" Proc.Intl conference on spatio-temporal databases (SSTD3), LNCS 2750, Springer, 2003, Berlin, Germany, pp 489-507 http://www.inf.fu-berlin.de/inst/ag-/publications/2003/hinze_voisard_sstd03.pdf
[10] Hui Kai-Lung, Png I. P. L., - "On the supply of creative work: evidence from the movies" School of computing, National University of Singapore, Paper provided by Economics Working Paper Archive at WUSTL, 2002
[11] Maturana H., Varela F. "The Tree of Knowledge", Shambhala, Boston, MA, USA, Revised edition, March 31, 1992
[12] Mobile tourist guide, IST-2001-36004, m-ToGuide, EU IST 5th Framework project 2003, www.motorolatele.com/MOTOnow
[13] Nattrass B., Altomare M., - "Dancing with the Tiger: Learning Sustainability", New Society Publishers, Gabriola Island, BC, June 1, 2002
[14] Pashtan A., Heusser A., Scheuermann P., - "Personal service areas for mobile Web applications" IEEE Internet computing, nov-dec 2004, 34-39
[15] Pau L-F, "The business challenges in communicating mobile or otherwise", Erasmus Institute for research in management (ERIM), Rotterdam, , www.erim.eur.nl, nov.2002
[16] Senge P., "The Fifth Discipline", Bantam Doubleday Dell Publishing Group, Inc., New York, 1990
[17] Robert K.H., "The Natural Step: a framework for achieving sustainability in our Organizations", http://www.naturalstep.org/ date not known
[18. Bhavani T., "Managing and mining multimedia databases", CRC Press, Boca Raton, USA, June 2001, ISBN: 0-8493-0037-1
[19] "An environment to bring together collections, services and people in support of the full life cycle of creation, dissemination, use and preservation of data, information and knowledge". http://bengross.com/dl/
[20] "Definition and Purposes of a Digital Library" , Association of Research Libraries (October 23, 1995) which became IFLA

Panel on European Accessible Information Network, EUAIN

Organised by

EUAIN and FNB

Coordinated by

David Crombie , FNB, The Netherlands

Neil McKenzie, FNB, The Netherlands

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

Open Source and Open Standards as tools for accessibility

David Crombie[1], Roger Lenoir[1] Benoit Guillon[2], Neil McKenzie[1]
[1]*FNB, The Netherlands*
[2]*Association BrailleNet, France*
Dcrombie@fnb.nl, rlenoir@fnb.nl, bguillon@snv.jussieu.fr, nmackenzie@fnb.nl

Abstract

The concept of Libre and Open Source software development is not a new one, but much of the infrastructures and development frameworks are now of a mature enough level that Open Source Initiatives provide a valuable tool for solutions in the field of accessibility. The first part of this paper outlines some of the ways that accessibility can add new markets for Open source software initiatives and how these initiatives can be valuable tools for achieving design for all. The second part outlines and overviews some of the activities under which these initiatives meet.

1. Introduction

Accessible design has been around in the form of Design For all[design for all] for some time, and although the take up has been slow, there are major developments now taking place which incorporate accessibility lower in the virtual supply chain. This makes mainstream software more usable and also smooths the uptake of new technologies for visually impaired and other specialist markets

There is a now a move towards a more macro level definition of what is required for accessible solutions. The focus, while still on technology and ensuring solutions are available to those who need them, is also moving towards areas such as standards and dissemination of specialist knowledge into the mainstream. This is where Accessibility becomes a process rather than a product. [20]

In order to ensure that this focus is engrained in all the different facets of technology development which make up our modern world, it is essential that this expert knowledge is available in multiple forms.

This paper outlines the part that open source developers play in this emerging field and overviews the initiatives through which these developments are carried out.

2. Accessibility and Open Source initiatives

2.1 Accessibility

The core objective of accessible solutions is to communicate some sort of information in an alternative means. Accessibility should therefore be seen as a communication issue. This often requires recreating context in order to convey meaning, or presenting information in a new innovative manner.

Accessibility, while being a niche market to many is a huge subject. The subject requires awareness of developments in many separated areas. As an example important Accessibility work goes on at a development level.[14], at a lobbying level, [13], at a standards level[24], [11], and at an organizational level, [10]

Accessibility solutions can be anything from small plugins or command line tools for performing a simple but important task, to complete redefinition of system for processing information which requires the integration of several different actors in a system.

2.2 Open Source software solutions

Open Source software is at some levels an organisational tool, at others a development tool, and to others a philosophy. It's very nature depends on it's ability to be all or some of these things to many different types of people.

Analysis of the available projects on Sourceforge[21] shows that open source software

solutions can range from a small student project uploaded as part of the skills to be learned from that particular part of the curriculum to some of the largest software projects ever conceived. [25]

Open Source solutions have gained such success in recent years through the adoption of Apache[4] , Firefox[15] and other solutions within the mainstream that it cannot be argued that they are not here to stay.

3. Open source accessibility

3.1 Access to technology

One simple aspect in which Open source Initiatives, projects and technologies contributes to accessibility is the idea of access to technology. Very often designers and implementers of accessible solutions adapt simple tools in order to meet some very specific user need. On a very simple level, making software open source allows these designers to have access to software which can be used as a foundation to build small solutions on. The solution possibly wouldn't have been viable had they had to build from scratch.

Open source makes accessible code, which improves accessibility by increasing the source base which other developers start from in order to build technologies or adapt technologies for other users needs. On a more macro level, this allows users and developers of the code to have access to the deep infrastructures within the code, which supports the concepts of design for all[9], accessibility from scratch[7] and the idea of integration of accessible notions earlier in the information processing chain.

If the infrastructure is open then it provides the foundation for someone else to do the work in making the system accessible from the ground up without necessarily affecting the output of the system. i.e. retrofit accessibility.

3.2 Close Collaboration with standards

One of the biggest challenges faced by designers and implementers of accessible solutions is that of incorporating their solutions with the main stream in order to ensure scalability, extensibility and future relevance. As a result the accessibility world embraces standards, as they take away some of the problems in dealing with niche markets and very specific user needs.

A similar trend can be observed in the open source world, where there is a realisation that standards are needed for the uptake of open source solutions on a more mass scale. Endeavors such as the free standards group[11]ensure that standards exist to provide an easier transition in meeting user expectations and requirements.

There are several types of standards but they can usually be split into two categories; de jure standards and de facto standards. De Jure standards are standards that have been formalized by a standardization body. De facto standards are standards which have been adopted on such a wide spread scale due to their usefulness or reliability. The Open Document format whaich was recently standardized by OASIS is a de Jure or formal standard where as Microsoft Word is a De facto standard, in that you can expect everyone with a computer to have access to it.

Standards can be split into two further types; proprietary standards and open standards. Proprietary standards are developed and maintained by a commercial company who decide on changes in the standard, whereas open standards are developed and maintained in public where anyone can contribute to changes and the progression of the standard.

3.3 Building Communities

In accessibility world the use of communities to exchange information and best practice has been widespread for some time. Through necessity, users sought out and made contact with the other people around the world who take part in their niche activity. With the advent of the virtual world, these communities can now flourish online using the technology available, and as a result you find that there are several very specific communities which are able to bring together almost all the people with an interest in that field form around the world, be it virtually[1][5]or actually.(Summer camps for Braille Music in Austrailia)

Open Source projects rely on similar communities. Sourceforge is one of the best tools for managing these communities as it is essential to the smooth running of open source projects. The community building aspect of these virtual organizations is only just beginning to be understood [22]. People with different skill bases, from different countries for different reasons come together to unite in a common goal.

It would advantageous to both the accessibility world and to the open source world if the similarities in these emerging communities can be mapped onto each other to ensure smooth knowledge transfer from player to player and user to technology.

3.4 Similar organizations structures

[1] http://www.resonare.org

The field of accessibility, while recognized as one of the few growth areas in IT[1], is not traditionally a commercial field. For years organizations producing materials for the blind and visually impaired have relied on hard working teams of volunteers. These volunteers become experts in a niche field and offer their time to impart this knowledge and help create materials in accessible formats.

These volunteers are also found in the virtual communities which feed the open source projects all over the internet. Traditionally open source initiatives were based around volunteer communities, who were working not for financial reward, but for the future of technologies. This fantasy has now been enhanced with a more viable business opportunity which can bee seen in open source development, and people are now employed by companies to take part in the open source initiatives, but much of the organisational structure and attitudes intrinsically exist which are similar to the volunteer networks in accessibility.

In a similar way, the idea of the Guru within a niche market who has a very specialist knowledge can be seen in both the accessibility world and the open source world. People become established as "the" expert in a very specific field of a particular technology, standard or format.

This happens in the accessibility world, where a format such has Braille Music has many different flavours, and it is relatively easy to become the expert in one of those flavours if little work has been performed previously. In the open source world, a similar thing happens when work is carried out by one person on a very specific driver or file set.

The challenge is to ensure that the organisation and work distribution is such within these structures, that this expert knowledge can be modeled and passed on more easily than the guru being available 24/7 by email.

Open source initiatives therefore have experience in modeling expert user knowledge which is required to be incorporated within the system in order that the system can be built upon for the future.

4. Continuing work

There is much work under way in in making the ICT world a more accessible domain. This now takes place at many different levels as the challenges in accessibility become more widespread requiring integration at different levels and in different fields. This is where the focus of accessibility is less on the product and more on the process [20]

4.1 standards activities

Standards are a very important part of accessibility(see section 3.2). As a result there are several open standards initiatives which are taking place at the time of writing. There is, as mentioned above, the Free standards group Accessibility workgroup [12] there are also several other initiatives which either deal with accessibility directly or provide standards which are deamed to improve technology for integration with mainstream solutions. It is important that Open Source developers are aware of these as they provide a means to ensure an extra audience of end users to any information processing systems.

4.1.1 DAISY//NIMAS

DAISY(ANSI/NISO Z39.86) is one of the success stories of the Accessibility world. DAISY defines a NISO (http://www.niso.org/) standard for created structured audio books for the visually impaired. The Standard is used by almost all orgainsations producing materials for the visually impaired and there are several tools available which incorporate the standard for producing audio books[8] [2]

The DAISY Consortium was established in 1996 by a number of not-for-profit organisations and institutions serving blind and visually impaired persons. As of today, it consists of 12 Full Members and 57 Associate Members worldwide, as well as 23 Friends (for-profit companies who share interest for the DAISY format).

Using XML text files and MP3 audio files, with the DAISY format we can create a range of text only, fully synchronized text and audio and audio only books that are fully accessible and navigable for blind and visually impaired users as well as persons with other disabilities, such as dyslexia.

There is now a [standard emerging called NIMAS.(National Instructional Materials Accessibility Standard)[16] . Specifications for NIMAS 1.0 are already available, and it includes both a Baseline Element Set that has to be met by any publisher producing files according to the NIMAS standard, and a set of optional elements. The NIMAS format has had a lot of support from the publishing industry due to it's simplicity as a format.

A NIMAS set of files includes an XML source file based on the DAISY 3 standard, a Package File with information about all the files in the package and a PDF file with embedded images as a reference of what

the printed edition looks like. Using the same standard will greatly benefit to all publishers and all specialised agencies in the US when it comes to produce books in alternative format in a timely manner.

4.1.2 Web Content Accessibility guidelines(WCAG)

WCAG[24] 1.0 became a W3C(World Wide Web Consortium) recommendation in 1999. Currently there is a working draft of WCAG 2.0, The guidelines aim to provide best practice to web developers to allow them to create web content which is accessible to users with a variety of disabilities.

The basic tenets of accessible web design are to make sure that the information can be viewed through alternative means which do not detriment the mainstream solution. This is mainly through alternative description in content such as images. The guidelines explain how to implement multimedia in an accessible environment, so that developers don't have to sacrifice images or video to make their content accessible.

In order for Web Accessibility standards to be implemented on a wider scale technologies used for content creation must encourage their implementations. As an example NVU[17], the open source web editor encourages the user to enter an alternative text tag when inserting an image in a web page. This should be encouraged, as it leads the designer through the process, making the incorporation of accessibility less painful.

4.1.3 The Open Document format

On May 1st, 2005 the *Open Document Format for Office Applications (OpenDocument)* became an OASIS standard. The open document is the standardized version fo the XML format currently used by Openoffice.org[19]

As it is an XML based standard, it is very good news for accessibility, as it provide a means of separating presentation from content in an accessible and open manner.

There is currently an open source project under way through the EUAIN consortium, which will provide technologies for processing the Open Document format into accessible formats. [18] The OpenAIP hopes to provide some technologies which demonstrate what is possible in the filed of accessible document processing

4.1.4 Metadata Standards

One area receiving some attention within the accessibility world is that of metadata. In information processing, metadata becomes the essential tool for deriving information for interchanging this information with other formats. It therefore follows that Metadata is an essential tool for processing accessible information, where many formats rely on context and surrounding information in order to build alternative representations of the data.

The CEN/ISSS Workshop on Metadata for Multimedia Information – Dublin Core was held in 2004, organized by the Dublin Core Accessibility Working Group (part of the DC Metadata Initiative). This workshop put the stress in both multilingualism and accessibility. In its final recommendations, the Workshop advocated for the creation of a new element for Dublin Core metadata, *DC:Accessibility*, to describe accessibility of resources and services.

The aim of the CEN/ISSS MMI-DC workshop was to identify and investigate the ways in which metadata can help achieve efficient and future-proof solutions to accessibility. It is assumed that this encompasses the provision of adequate access to information for people with disabilities and for everyone in a multilingual and multicultural environment. In order to make this perceived information useful, it must be represented within an architecture which allows the accessibility requirements to be questioned in more than one way. Such an architecture must enable both the core system to adapt to new and changing representation requirements, and to allow (theoretically) infinite user requirements.

4.2 Open source initiatives for Accessible Information Processing

One of the hardest parts of open source systems design is bridging the gap between a working beta and a system useful to end users who didn't take part in the construction of the system. One objective of the EUAIN project is to ensure that there is awareness of the open source solutions for accessibility which are available.

A Quick scan of sourceforge throws up several open source software projects for accessible. The main problem is making people aware of what is available and uniting the fragments of various projects to create sustainable solutions for the future. It is hope that one output from the EUAIN project can be an analysis of the various projects available.

4.3 Accessible Systems

In order to facilitate the paradigm shift of moving accessibility to the Macro level, it is essential that large scale systems exist which can support the various technologies which will operate and interoperate within

these frameworks. There are several initiatives which take place to make these frameworks possible.

4.3.1 MPEG

The MPEG standards are revised and approved under the supervision of SC29, the Committee that deals with the standardization of the coding of audio, picture, multimedia and hypermedia information within the Joint Tecnhical Committee 1 (JTC1) of ISO (International Organization for Standardization) and IEC (International Electrotechnical Commission), two of the leading organizations that produce international standards (together with ITU, International Telecommunication Union).

At the last MPEG Meeting(73rd meeting in Poznan, Poland), FNB launched a Core Experiment to answer a call to enhance the MPEG 4 technology for representing music notations through a Symbolic Music Representation standard. This core Experiment will lay the ground work for implementing an Accessible Music Production Suite within this technology.

MPEG have an official liaison with ISO/IEC JTC 1 Special Working Group on Accessibility. JTC 1 establishes a Special Working Group on Accessibility which aims to:

- determine an approach, and implement, the gathering of user requirements, being mindful of the varied and unique opportunities (direct participation of user organizations, workshops, liaisons)
- identify a mechanism to work proactively between meetings to make forward progress
- gather and publish an inventory of all known accessibility standards efforts
- identify areas/technologies where voluntary standards are not being addressed and suggest an appropriate body to consider the new work
- track public laws, policies/measures and guidelines to ensure the necessary standards are available
- through wide dissemination of the SWG materials, encourage the use of globally relevant voluntary standards
- assist consortia/fora, if desired, in submitting their specifications to the formal standards process

It is hoped that the EUAIN consortium and MPEG's involvment in this activity ensures that the results and outcomes have the desired effect and can be dissemuinated into the industrial mainstream.

4.3.2 Accessible Xoops

The combination of accessible content management systems, accessible desktop systems and content modalities with internalised notions about accessibility, can be used to form a new generation of information processing environments. Because of the presence of explicit entities that can be used to represent the User (perception) models on one side, and content and application models on the other side, we can experiment with new interaction schemes. These new interaction schemes will, because of the knowledge preservation process that is included in the approach, create a consistent body of information including real-world applications for education purposes.

From the knowledge technology perspective, the development from traditional content consumption schemes towards knowledge consumption and even understanding consumption that may be stimulated to emerge. After all, understanding can be considered the dynamic systemic overview one can obtain of all the facts, the interactions between facts and the interactions between these interactions and it's surroundings. The mere process of conceiving and creating this systemic overview can be considered education to oneself. Allowing a system to include multiple perspectives on that systemic overview and additionally allow that system to create associations between these multiple viewpoints for any relation to be explicated, stimulates the emergence of mutual understanding. In other words, a system that facilitates communication from scratch.

In order to implement these ideas on a wider system level, a project has been develop which builds on the Xoops portal system[27] to create tools for creating accessible content management systems. The project is currently in the development stages, but as it reaches maturity results will be posted on the EUAIN Web Portal

4.4 Communicating and educating

One of the key efforts in the dissemination of knowledge about accessible design, accessible standards, and accessibility projects is communication and networking. There are several initiatives that tackle parts of this area and it is important that people are made aware of them.

4.4.1 EUAIN

EUAIN This project aims to promote e-Inclusion as a core horizontal building block in the establishment of the Information Society

by creating a European Accessible Information Network to bring together the different actors in the content creation and publishing industries around a common set of objectives relating to the provision of accessible information.

The authors of this paper are EUAIn consortium members and much of the information in this paper is also available on th EUAIN web portal.

4.4.2 AIIM and PDF access

AIIM (The Association for Information and Image Management) [3] has been the leading international organization focused on helping users to understand the challenges associated with managing documents, content, and business processes. AIIM holds a number of standards committees and working groups that draft recommended practices for different activities. These drafts are being reviewed and revised until they agree the document is ready to be submitted to ANSI for approval.

The PDF-Access Working Group deals specifically with how PDF documents can be fully accessible. The starting point is that PDF files can be accessible, they just need to incorporate a number of guidelines on how to convey the information that traditionally has been only useful for sighted users. Also, the committee will look for ways to make PDF exportable to XML and NIMAS in the USA.

According to the working group, a PDF file will be considered accessible once it is tagged, all text within the file is searchable, has a logical read order, contains alternate descriptions for all the images included in it and is navigable.

4.4.3 EdeAN

Created in accordance with one of the specific goals of the e-Europe 2002-2005 Action Plan, EdeAN is mainly engaged in raising the profile of Design for All and emphasizing its importance in achieving greater e-Accesibility.

EDeAN consists of a number of Special Interest Groups that are relevant to the EUAIN objectives:

- *Benchmarking*. Development of a structured mechanism for consensus creation and assessment of available knowledge on DfA.
- *Standardisation*. Contribute to pre-standardisation activities aiming to support groups active in European standards making (e.g., CEN/ISSS DfA Worksop on "DfA and Assistive Technologies in ICT"), as well as international standardisation bodies, in developing implementation strategies for DfA.

Proactive assessment. An effort towards providing an account of technological developments likely to take place in the future, with the intention to assess both the impact of these developments on certain user communities, and the validity of DfA as an instrument for proactively accounting for technological accessibility. This SIG will be concerned with future aspects of the emerging information society, and will aim to inform and validate DfA as a proactive philosophy of design.

5. Conclusion

It is hoped that the information provided in this paper serves to provide information into both Open Source Development for accessibility and Open Standards which are useful for accessibility.

The most important thing is the unifiction of the fragmented knowledge which this very wide area covers and anyone who is interested in being part of any of the work presented here should contact the authors

7. References

[1] “Accessible Technology in Computing—Examining Awareness, Use, and Future Potential” Research Commissioned by Microsoft Corporation and Conducted by Forrester Research, Inc.,

[2] Amis adaptive multimedia information system: http://www.amsiproject.org

[3] The Association for Information and Image Management: http://www.aiim.org/

[4] The apache software foundation: http://www.apacche.org

[5] Blind Programming discussion list: http://www.blindprogramming.com/]

[6] Braille Music Discussion list: http://www.resonare.org/

[7] Crombie, D., Lenoir, R., and McKenzie, N., (2004) Accessibility from scratch : how an open focus contributes to inclusive design, Proceedings ICCHP,

Lecture Notes in Computer Science, Vol (3118). Springer-Verlag, Berlin Heidelberg New York

[8] The DAISY consortium: http://www.daisy.org

[9] European Institute of Design fro all: http://www.design-for-all.org/

[10] European Design for all e-accessibility network: http://www.e-accessibility.org/

[11] Free Standards group: http://freestandards.org/

[12] Free standards group, accessibility work group: http://accessibility.freestandards.org/

[13] ISO/IEC JTC 1 Special Working Group on Accessibility (SWG-A)

[14] KDE accessibility project: http://accessibility.kde.org/

[15] The Mozilla foundation: http://www.mozilla.org/

[16] national instructional materials accessibility standard: http://nimas.cast.org/

[17] NVU: http://www.nvu.com/

[18] The OpenAIP Project: http:://www.openaip.sourceforge.net

[19] Openoffice.org: http://www.openoffice.org

[20] Regan b. "Web Accessibility and Design: A Failure of the Imagination", http://www.designfor21st.org/proceedings/proceedings/plenary_regan.html

[21] Sourceforge: http://www.sourceforge.net

[22] Wagstrom, P.A. Herbsleb, J.D., Carley, K. "A social Network Approach to Free/Open Source Software Simulation" Open Source Systems 2005, Genova, Italy.

[23] Web Accessibility Initiative: http://www.w3.org/WAI/

[24] Web content Accessibility Guidelines: http://www.w3.org/TR/WAI-WEBCONTENT/

[25] Weiss, D. "Quantative analysis of Open Source Projects on Sourceforge" Open Source Systems 2005, Genova Italy.

[26] World wide web consortium: www.w3.org

[27] Xoops project http://www.xoops.org

EUAIN Overview

David Crombie[1], Roger Lenoir[1], Neil McKenzie[1]
[1]*FNB, The Netherlands*
Dcrombie@fnb.nl, rlenoir@fnb.nl, nmackenzie@fnb.nl

Abstract

Accessible information processing streams are non standardised and fragmented. Almost every processing chain which includes accessibility relys on quick fix steps and stop gap solutions. The EUAIN project began last year as part of the eInclusion thread within the IST research programme of the European commission. EUAIN aims to promote e-Inclusion as a core horizontal building block in the establishment of the Information Society by creating a ***European Accessible Information Network*** *to unite the different participants in the content creation and publishing industries' supply chain of content in accessible mediums.*

The paper describes some of the processes behind the EUAIN Project and overviews what actions have taken place so far. There is then an outline of activities planned for the future and information on how interested parties can get involved in these activities.

1. Accessibility as a communication tool

Accessible solutions are required for anyone who requires assistance in using the mainstream solution. This could be because a user is blind, visually impaired, or impaired in some other way, and the term print-impaired is often used in this context. Accessible solutions range from small assistive applications (such as a screen magnifiers), to full scale operating systems and screen reading environments. The traditional problem with accessible solutions is that they are normally implemented as an afterthought or a piggy-back solution. This results in solutions which are not fully integrated (or not well integrated) with the mainstream solutions. These independent applications are then at a disadvantage whenever software versions or operating systems are updated. In order to make this integration process easier, and provide more intuitive designs for the future, it is essential that "design for all" and accessible design methodologies are widespread. Standard, policy and legislation also helps ensure that accessible designers have a solid standard to meet to ensure future-proofing.

Notions of "accessibility" are normally equated with the adaptation and conversion of digital content, where this content can be made available. On a European level, and indeed often on a national level, much of the existing expertise on creating accessible adaptations of digital content is of a highly distributed nature. Within specialist organisations supporting print impaired people; or within university research laboratories; or indeed within publishing houses, many automated tools have been designed and implemented at least partially to execute the necessary adaptation procedures. However, each automated tool has its own, highly specific, field of application. Furthermore, the knowledge required to build these very specific tools is equally distributed, so that there is currently very little re-use of either tools or knowledge. The content provider's perspective on digitisation is further complicated by security issues. In the modern environment driven by the internet for content dissemination, security is a vital issue for publishers. DRM is a complex problem for all content holders. Every publisher's content, client base and requirements are different, which often results in a personalised set of requirements for each case. As a result, agreements on accessibility are often negotiated on a case-by-case basis. Naturally, publishers have to be confident that any digital format is being delivered through secure gateways to only the people who are intended to receive it.

2 A Broader Perspective

Accessibility can also be viewed from a wider angle. Being able to see content in whatever modality; perceive its context; and attach a useful meaning to it requires that the user be able to access this content, its context and relevant software application in a way that meets that particular user's consumption preferences. These preferences may become requirements over time - we all get older. Being able to attach useful meanings

to content is what lies at the very basis of the preservation and education of thought. Attaching useful meanings to content underpins the basis of culture, commerce and civilisation. Being able to access software and the content and the potential for understanding it unleashes, requires us to be able to gain access to software and not be hindered by huge costs, complexity, lack of support and additional barriers.

Designing a more inclusive world requires a more *Open Focus*. This *openfocus* can be achieved through an interplay of practical solutions conceived by greater co-operation between science and philosophy; technology and industry; and community and education. The traditional route to solving a problem requires that 'expert knowledge' is built onto the subject at hand before the problem is tackled. This layered knowledge is built upon until the expert points of view are focused almost exclusively on the solution. However, incorrect or inappropriate knowledge may lead to an intellectual dead-end. The solution to this dead-end is to take a step back, or in structural terms to move to a perspective with a higher level of abstraction. This can be described as an *openfocus*.

In order to model user requirements central to the users' needs rather than relying on existing tools, a change of focus is required. This may often seem to be counter-intuitive, apparently undermining the focused concentration usually associated with such fundamental problems. This *openfocus* requires thinking in a non-standard abstraction layer, and can actually be very easy if the mind is trained to overlook fundamental beliefs. The *openfocus* approach refers to a generalisation of the problem by looking at the bigger picture, which may involve considering different perspectives. These perspectives are often far removed from the perceived core of the problem, such as user requirements or product aesthetics, but they can lead to generic and robust solutions. Under *openfocus*, abstraction leads to a better understanding of the situation by contributing information from multiple viewpoints.

Transition & Convergence

Given the differences between the traditional approach to accessibility and the wider view outlined above, we are in something of a transitional phase at this time. From the software producer, business community and the Open Source System community we see a move towards the inclusion of accessibility features into systems, tools and the programming languages themselves as system wide core functionalities (examples being KDE, GNOME, and Java Accessibility). From the accessibility community we see a move towards more advanced and abstract descriptions of the procedures involved in moving from 'common' content towards content that is processed to be granted accessible certification. A good example of such a move is the Web Accessibility Guidelines 1.0 and 2.0, which provide detailed guidelines on how to (re)structure and enhance websites and their content to ensure a sufficient level of accessibility.

The transitional stage described above involves relatively slow change when compared with general exhilarating technological developments. However, this relatively slow pace also creates an opportunity to take a step back and observe all the individual processes that touch upon the notion of accessibility. This allows us to explicate similarities and possible complementarities, a process of convergent gradualism if you like. The opportunity then arises to synchronise various efforts in the accessibility arena and offer them to end-users and business as a 'package'. Such a package contains scientific knowledge about accessibility, as well as technological knowledge about how to implement such notions. This package also contains detailed descriptions of the requirements of the end users, producers and distributors of content, as well as tools aiming towards market segments that rely on these requirements. Such an approach that aims to unify 'common' content, system, service and tool provision and the more 'specialised' content, system, service and tool provision, can be called Accessible Information Processing (AIP).

However, what is still clearly needed is a focal point to bring these disparate initiatives together. The European Accessible Information Network will provide this cohesion by addressing the key areas and issues which are of common concern to all the actors in this area. This is an ambitious goal, but the convergence described above makes this both worthwhile and achievable.

3 Current Implementations

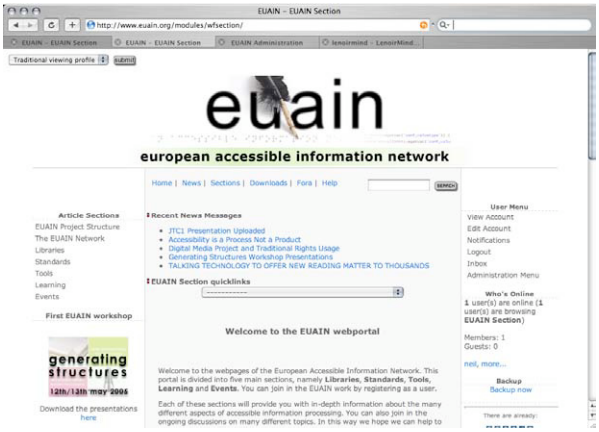
3.1 The EUAIN Web Portal

The EUAIN Web portal [8] is the central knowledge management. The portal is both a dissemination tool and an implementation of many of the design for all concepts accessible information processing techniques which are output from the EUAIN project.

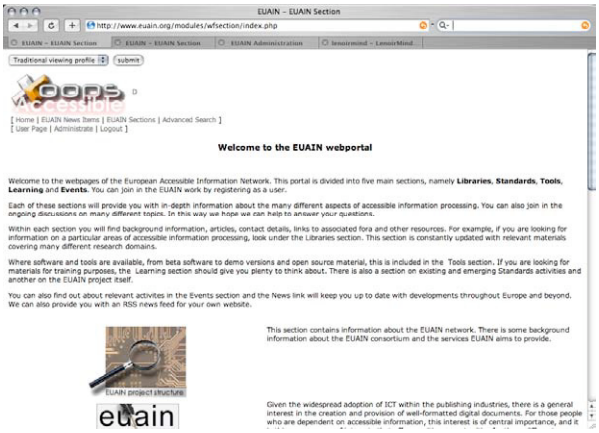
The portal use the theme engine of the Xoops portal system to separate presentation of the information from the in formation(Content) itself. This provides the

portal with the ability to personalize content and present it in different manners to different users. This is made possible through content adaptations which can take place on the meaningful user models within the framework.

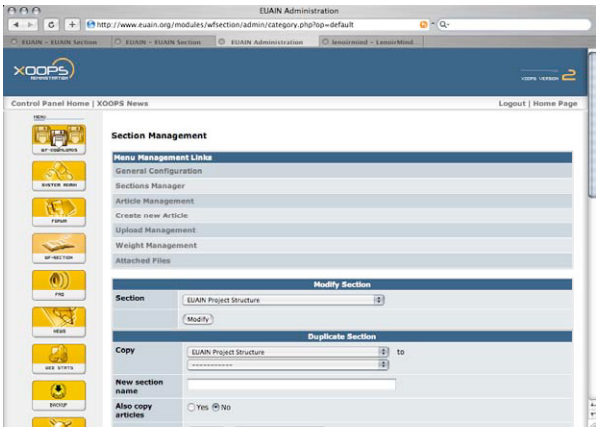
The portal uses a access privilege system(With the theming engine) to provide different levels of complexity to different users. There is the normal front page:



The same information presented with an accessible theme, for screen reading or use with other adaptive technologies:



Behind this, and controlling the content, users with admin privledges are able to access the admin section with WYSIWYG editors and content management tools:



The EUAIN portal as been very well received, with over 150 registered users, active forums and over 500 posted News items. There have also been several downloads of the dissemination materials which have been made available so far in the projec

3.2 First EUAIN Workshop on Generating Structures

The First EUAIN Workshop on Generating Structures took place in Brussels in May. The Generating Structures Workshop provided an opportunity to hear from many different voices and communities of interest. During the first session, some of the ideas and concepts behind the relatively new approach which lies behind the EUAIN network were presented. In many different countries, interesting initiatives running for several years demonstrate the need and value of greater access to sources of information.

During the second session, several publishers who have taken practical steps to widen access to their content, with some tangible results gave their opinions on how the actions of EUAIN can improve Accessible Information processing in industry. In addition, we heard about several practical tools which can be used to make information accessible and to provide a reusable level of structural flexibility. This is particularly important in this age of multi-channel publishing.

The third session focused on standardisation activities and the concepts and ideas which are being implemtned through the CEN/ISSS workshop on Document Processing for Accessibility

The final session covered the interface between accessibility and DRM issues. This is an area which is of increasing importance, particularly as content creators move towards more complex multimedia formats and the need for greater structural integrity emerges. A combination of technical protection mechanisms and trusted intermediary communities does indeed seem to point the way forward.

The workshop was very successful and it is hoped that the success can repeated in the future workshops.

3.3 CEN/ISSS Workshop on Document Processing for Accessibility

The diversity of perspectives taken on accessible information processing is reflected in the project structure itself, with a workpackage examining standards and through the specific inclusion of a proposed CEN/ISSS Workshop [10], called Workshop on Document processing for accessibility(WSDPA), which was established in may as a CEN/ISSS Workshop under the existing CEN Workshop Agreement procedures. The Secretariat for the Workshop was established through the CEN National Member, the Nederlands Normalisatie-instituut [11] (NEN).

Although the Workshop on Design for All (WS/DFA) has been completed, we would hope to build on the outcomes from this work, and in particular, to provide a forum where RTD projects can present their work with a view to standardisation. This would involve the ongoing creation and verification of standards on the provision of accessible information for the different workpackage themes (standards, protection, production and tools and distribution).

EUAIN will necessarily also be involved in standardisation activities with other groups. In particular, EUAIN would aim to raise awareness of the DAISY format for digital audio [12], currently a NISO standard but likely to become an ISO standard in the relatively near future. As noted above, it is now possible to see DAISY 3.0/NISO z39.86 as the de facto XML standard which can allow content creators significantly to enlarge their markets through the adoption of this inclusive format.

The European Council, in 1994 [13], stressed the need to create a general and flexible legal framework at Community level in order to foster the development of the information society in Europe. Important Community legislation to ensure such a regulatory framework is already in place or its adoption is well under way. Copyright and related rights play an important role in this context as they protect and stimulate the development and marketing of new products and services and the creation and exploitation of their creative content. A harmonised legal framework on copyright and related rights, through increased legal certainty and while providing for a high level of protection of intellectual property, will foster substantial investment in creativity and innovation, including network infrastructure, and lead in turn to growth and increased competitiveness of European industry in the area of content provision and information technology.

4 Future Work and how to participate

All activities which take place through the EUAIN project are covered on the Web portal[8]. The second EUAIN workshop will take place in November of this year. Information on Call for papers, registration and receiving materials following the event are available on the portal.

Interested parties can get involved in EUAIN on many levels. These are not restricted to people within the European Union, and users from countries outside Europe are encouraged to post on the fora and download deliverables and dissemination materials. Involvemnt in the processes can take any form; from posting news items and discussion items on the for a to attending meetings and presenting positions. The most important thing is to get involved in these emerging and continuing work.

5 Conclusion

When we examine the current situation it is clear that there are a number of good initiatives in the area of accessible information provision, but these initiatives are fragmented and examples of successful implementation are neither widely disseminated nor clearly understood by the different stakeholders in the information publishing chain. If these fragments can be brought together in discussion of standards and practices, then the previous efforts become more worthwhile by being incorporated into a cohesive framework for a wider audience. The EUAIN project will tackle these issues in a systematic manner and will provide solutions which can only be achieved by examining these problems at a European level. In so

doing, any outcomes are immediately applicable for all European member and candidate states.

References

1. Miesenberger, K., (2001) Convergence in Electronic Libraries, Cultural Heritage and Service Provision for Print Disabled People: Austrian Literature Online, in: Marincek, C., Bühler, Ch., Knops, H., Andrich, R. (ed.): Assistive Technology – Added Value to the Quality of Life , Ljubljana, IOS Press, Amsterdam
2. CEN Workshop Agreement 14661:2003; page 7. This project focused on producing guidelines for standardisers which complement CEN/CENELEC Guide 6.
3. ICT Standards Board Project on "Design for All and Assistive Technologies in ICT, How ICT standards can enable all people in daily living", CEN/ISSS Open Meeting on Design for All and Assistive Technology, Draft Report: http://www.cenorm.be/isss/Workshop/Design-for-All/Default.htm
4. IMS (2003). IMS Learner Information Package Accessibility for LIP Version 6.0. Available at: http://www.imsglobal.org/accessibility/index.cfm
5. http://www.design-for-all.org
6. www.e-accessibility.org
7. www.edean.org
8. www.euain.org
9. Crombie, D., Lenoir, R., and McKenzie, N., (2004) Accessibility from scratch : how an open focus contributes to inclusive design, Proceedings ICCHP, Lecture Notes in Computer Science, Vol *(forthcoming)*. Springer-Verlag, Berlin Heidelberg New York
10. See www.cenorm.be
11. See www.nen.nl
12. ANSI/NISO Z39.86-2002 (ISSN: 1041-5653) Specifications for the Digital Talking Book. See www.daisy.org

Managing electronic content for adaptation to the reader's profile: project MultiAbile for the inclusion of e-Learners with disabilities.

Thimoty Barbieri[1], Antonio Bianchi[1], Licia Sbattella[1], Ferdinando Carella[2], Marco Ferra[2], Piergiovanni Santini[2]

[1]*Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy* [2]*Web Learn srl, Strada Padana Superiore, 317 – 20090 Vimodrone (Mi), Italy*
thimoty.barbieri@polimi.it, abianchi@elet.polimi.it, licia.sbattella@polimi.it, f.carella@weblearn.it, m.ferra@weblearn.it, p.santini@weblearn.it

Abstract

The need to provide better accessibility to networked information is widely recognized and expressed in a number of initiatives and government regulations such as W3C WAI and Section 508 of US Rehabilitation Act. Most part of these guidelines focus on making more accessible the "navigation" or "publishing" environment. Project MultiAbile (http://www.multiabile.it) is an ongoing research and implementation effort to create an example of web application that adopt a new approach to accessibility question: while accessibility of the navigation/publishing environment is a necessary foundation, we advocate that much more attention should be devoted to the creation, aggregation and processing of electronic content, together with the care for the environment which is used to publish it to users. Multiabile is a distance learning environment dedicated to the general public, including people in situation of disability, offering a range of different channels and modalities to access the content and to study it. Some of the different modalities that have been considered to access the content are: improved accessibility for screen readers; dynamic transcoding to a dialogue-based navigation over the telephone, using VoiceXML speech synthesis and recognition; real time TTS transcoding using high quality engine; sound/tactile description of images (e.g. cartographic data) using force-feedback mice or pads. The content is also offered in two different textual versions, an original version and a re-processed version, in order to obtain high readability and intelligibility. The edited version is created in a way that privileges the use of words from a basic dictionary (the Italian "Vocabolario di Base"), and that aims to obtain a high GULPEase index (an index of complexity in the structure of text). In order to encompass the variety of channels and modality with which the user can "read" content from the platform, we devised a metadata schema with which to organize the elements of the provided content, and a creation and editing workflow to process the content before publication in the learning environment. The resulting content is also structured in a way to be compliant to SCORM standards, so it is viable to process for multichannel and multimodal access also pre-existing SCORM contents. Users can describe their —profile" œ whether they use a screen reader, or prefer to have a dialogue-based navigation, or if they prefer a simplified version of the text before confronting themselves with the original version. The user profile is used to dynamically adapt the content to the user, facilitating user accessing it. In the paper we describe the general data structure devised within the project to accommodate the different access modalities to the content, and to allow the user to access and read it dynamically according to his/her profile. We explain how this structure impacts on the workflow of creating and editing of the content, or of repurposing and adapting preexisting content in SCORM-compliant format.

As a conclusion, we advocate that this workflow and content structure can pave the way to offer the same content to different channels and with different modalities (including emerging channels, such as Digital TV, and specific modalities, like for example automatic rendition to sign language), limiting the effort of repurposing the content by means of automatic transcoding and transformation algorithms.

1. Foreword and Motivations: Electronic Content Accessibility

The need to provide better accessibility to networked information is widely recognized and expressed in a number of initiatives and government regulations (W3C WAI, Section 508, etc.). However, most part of these guidelines focus on making more

accessible the "navigation" or "publishing" environment (in the sense of the digital system through which the user can obtain the desired information, for example a web site through which the information is offered) more than facilitating the use and the comprehension of the electronic content within. Accessibility guidelines, in their current implementation, are mostly used as prosthetis, alongside other assistive technology devices (screen reading software, Braille displays, etc.), with web sites or publishing environments featuring an "accessible" portion alongside the "normal" part of the environment. While accessibility of the navigation/publishing environment is a necessary foundation, we advocate that much more attention should be devoted to the creation, aggregation and processing of electronic content, more than to the environment which is used to publish it to users. Improved strategies at content level allow to empower the user offering him/her several alternatives to navigate and peruse the content itself, turning the way accessibility is currently implemented in electronic systems from a "prosthetics" notion to a truly multichannel and multimodal paradigm [2, 3, 12].

Focusing on how electronic content is structured, more than on the web site in which it "lives", allows to dynamically repurpose it with transcoding and transformation procedures over different communication channels and interaction modalities (e.g. displaying on a web page with AAA-conformance, reading aloud with TTS over the phone [4], efficently presenting contents to Braille displays or converting it in an accessible/readable PDF eBook, providing information via a Digital TV screen). This way, once the user has selected the content he/she wishes to access, he/she can select the preferred channel and plurality of modalities. In a scenario of content adaptability over the channel chosen by the user, it is crucial that the user can express, in a formalized way, the features which are required in accessing the electronic content. This information, collected and organized in a general model, the *user profile*, is the base to describe the interaction modalities between an electronic publishing system providing contents and its user, both with or without disabilities .

In this paper, we present this approach, that we followed during the development of project "MultiAbile" (http://www.multiabile.it). We created a "publishing multimodal platform", which can deliver electronic contents in different forms, based on a simple user profile described by the user himself. The content was described using a simple DTD devised with these objectives in mind [12], and aggregated following SCORM standards to provide the functionalities of a multichannel, multimodal e-Learning platform which could offer improved accessibility to the teaching material for the learner with a disability.

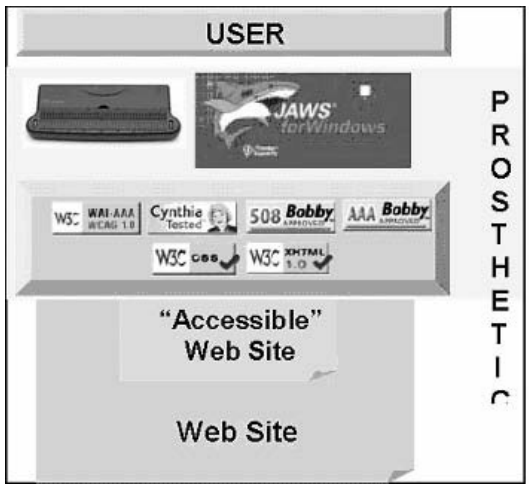


Figure 1 - Focusing on electronic contents accessibility (right)

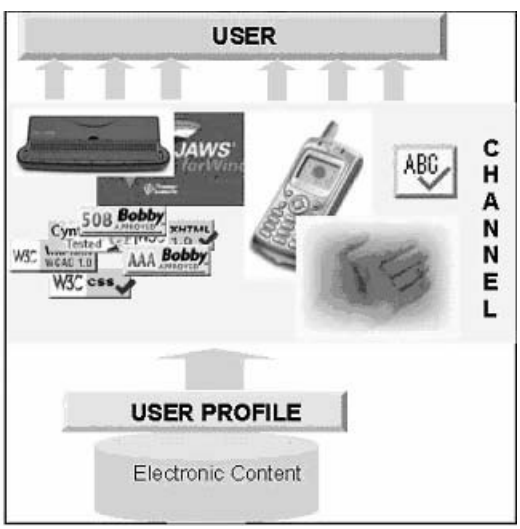


Figure 2 - over accessibility of the "publishing container" (left), it is possible to provide the user with a disability with an adaptive, multimodal electronic publishing environment that better adapts to his/her needs

MultiAbile can be considered a "multimodal electronic publishing platform" because it presents electronic contents to the users using different channels and with different and multiple modalities,

according to the defined user profile. [6]. Currently the channels supported by the implemented prototype are the web and the telephone, with possibility to transcode dynamically (with an XSLT-based approach) electronic contents to AAA-compliant HTML readable with Internet Browser, a Browser Screen Reader, or a Braille Display; VoiceXML with TTS/ASR navigation system accessible over the phone; Macromedia Flash cards with pictures, animations and colorful layouts. Over these channels, the user profile determines a number of modalities with which content is aptly optimized to better suit the user's needs [10].

These modalities include:

- different levels of language and text structure complexity for the same content, measured with a cognitive complexity index (refined and tested only for Italian language, at this time);
- coupling of information with audio and vibro-tactile cues to support navigation into content and exploration of contents of a graphical nature;
- use of mouse/hand gestures to improve or facilitate navigation into content;
- reading of single sentences or paragraphs with high-quality synthesized TTS (Text-To-Speech).

These modalities, most of which can be used simultaneously mainly using the web channel, are designed to help people with different disabilities finding the way which best suits their preferences, in accessing the same structured electronic content. In our early tests, we are validating the benefits of a multimodal publishing environment with:

- users with visual disability, which benefit from using the TTS functionalities, the audio/tactile cues and the alternative informations through Braille display; - users with acoustic disability, which benefit from the availability of a simplified version of the text (in structure and lexicon) which helps to overcome difficulties with the written word; -dyslexic users, which benefit from the possibility to have the structurally simplified version read by a high quality TTS system; -users with cognitive disorders (problems in attention or sequencing), which may benefit from a highly graphical, multimedia rendition of the content using Macromedia Flash cards.

2. A Reference Model for the User

People with disabilities have been traditionally grouped into profiled categories: blind, deaf, physically impaired, and users with learning disabilities. This approach has had its effectiveness, focusing on specific needs. Nevertheless, this approach has been growingly criticized by associations of disabled people for its narrow vision and for the approach with sets the person with a disability in a markedly separated context. During these years different ways to represent the situation of disability have been worked out. In the 1980s WHO's ICDH2 proposed a linear approach distinguishing among impairment, disability and handicap, but still conserving a causal approach. Confrontations between social, medical and associative institutions have produced the ICF, [5] International Classification of Functioning disability and health. This document has the great merit of introducing a more integrated representation of the health state of a person; its four areas are: body functions, body structures, activity and participation and environmental factors. ICF has been thought as an operative tool, useful for comparing different levels of intervention both on individual and social politics level and for evaluating the dynamic of a person's functions. We have taken inspiration from this approach in designing Multiabile's user profile, using attributes borrowed from ICF for representing the functional attributes of the user. This choice has the advantage of focusing on a relatively neutral description of the characteristics of the user, a description that may be adapted to any user, and not specifically to people with disabilities. This is perceived, and indeed it is, as an important factor of social integration. The problematic side of this approach has to do with the availability of the information needed for a complete profiling, partially related to privacy issues. Obviously the eagerness of the user to supply information about him/herself is related to the perceived benefits from accessing and using the information publishing system. To increase this perception, users may express a set of preferences, both about device availability and about his/her intentions to use, for example, a multimodal or mono-modal interface, or the guided or fluent dialogue version of the vocal interface.

Figure 3 - An example on how the user can set his/her profile in the publishing environment to configure multimodal access to electronic content

3. Content Structure and Creation Workflow

To create and organize content in a way that could be dynamically adapted to the user's profile, we adopted a predefined content structure and an electronic content management workflow based on XML and XSLT, with a final step consisting in packaging learning contents into modules according to the Content Aggregation Model provided for in the SCORM standard [1]. The general purpose of the content development framework originally was to allow rapid development of e-learning content modules. Describing electronic content in an XML format allows a complete separation between the content itself and its presentation, making it possible to present the same electronic content in different contexts or on different media. During the production of the first adaptable content, we designed a streamlined chain of activities to obtain a controlled production flow, with a high production throughput.

The general workflow can be described as follows: usually the original contents are supplied by authors in different forms (usually a plain Word document). Starting from this plain document, an editorial curator creates a "storyboard template", reorganizing the provided content marking titles, reorganizing in paragraphs, deciding the general layout of the content for presentation in multimedia format, classifying sentences as more relevant and less relevant (for content restructuring), proposing synonyms and simplified versions of sentences, etc. This "storyboard" is designed keeping as reference specific "semantic placeholders" that are related to the elements present in the XML data structure, defined in a Document Type Definition. The DTD is organized in a way to be kept as simple as possible (to maintain an error-free and effective content management), but still to be able to capture all necessary content structure information to allow a correct repurposing of the content across channels and modalities. In a second step, content editors receive the storyboard (usually a text document) and proceed creating an XML version of the content. The creation consists in starting from a new XML template using a normal XML editor (e.g. Altova's XML Spy), then the storyboard placeholders are mapped into XML elements. Starting from the storyboard, the XML content module is generated in at least two versions, one using original texts and contents, and another simplified in language, lexicon and structure. The language simplification is verified against a language complexity index, which we have elaborated combining different complexity analysis techniques (the GULPEase index, combined with the use of words taken from a Base Vocabulary of the language and a complexity analysis at morpho-syntactical level) already existing for the Italian language [9]. The following DTD shows the data elements which are used to organize the content.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--DTD generated by XMLSPY v2004 rel. 3 U
(http://www.xmlspy.com)-->
<!ELEMENT game EMPTY>
<!ATTLIST game path CDATA #REQUIRED
visible CDATA #REQUIRED >
<!ELEMENT item EMPTY>
<!ATTLIST item id (voce1 | voce2 | voce3 |voce4 |
voce5 | voce6) #REQUIRED status CDATA
#REQUIRED >
<!ELEMENT module (title, subtitle)>
<!ELEMENT page_content (#PCDATA)(emphasize)>
<!ELEMENT page_title (#PCDATA)>
<!ELEMENT quiz EMPTY>
<!ATTLIST quiz path CDATA #REQUIRED visible
CDATA #REQUIRED >
<!ELEMENT sco (module, sco_name,sco_content,
sco_menu)>
<!ATTLIST sco version CDATA #REQUIRED >
<!ELEMENT sco_content (sco_page+)>
<!ELEMENT sco_menu (item+)>
```

```
<!ELEMENT sco_name (title)>
<!ELEMENT sco_page (page_title,page_content, quiz, game)>
<!ATTLIST sco_pageid (1 | 2 | 3 | 4 | 5 | 6)#REQUIRED >
<!ELEMENT subtitle (#PCDATA)>
<!ELEMENT title (#PCDATA)>
```

When the document is complete, it is validated against the DTD: this ensures that the document is both well formed and valid. In this developing process, it is high unlikely that errors can occur without generating some validation errors. The XML editor assists the developer in creating structured content and the conformance between the presentation template adopted by each channel and DTD structure description enables an easy flow from content to presentation, maintaining a clear separation that makes possible the reuse of the XML content files in other applications/publishing environments. Once content is structured following the present DTD, it can be transformed using an XSLT transformation stylesheet, using a predefined template into which the content itself will presented (e.g. a graphical frame for Flash presentations, a template web page for HTML, a voice navigation prompt system for VoiceXML, a standard page layout for PDF documents, etc.).

Once content is transformed into one or more version, the third step in the process is to package it as a learning object (commonly called SCO in the SCORM course structure) into a single deliverable SCORM package. This SCORM package is ready to be installed into a regular e-Learning platform, or into a publishing environment which allows multimodal interaction with the content itself, such as the MultiAbile site.

4. Content packaging and transformation

The aggregation of the learning content is made by creating an XML document named “imsmanifest.xml” in which one defines the organization of the module in sections, lessons, quizzes and so on, and specify all resource files needed to use the course (XML files with the actual content of each learning object, images and so on). All learning objects and course assets are finally archived together with imsmanifest.xml in a compressed ZIP file that can be uploaded to the e-learning platform that manages the contents. The course is thus published for registered users - keeping track of user access, session time, lesson location, quiz scores and so on, according to the SCORM v1.2 “Runtime Environment Specification”. Using XML as the basic language to define content is the most convenient way to deliver the same course among different channels and to provide to the user different modalities of interaction with the system. One of the most interesting modality of interaction provided by the MultiAbile LMS is the vocal user interface realized using VoiceXML, that is delivered through a voice gateway that gives access to the platform by using a normal telephone. Content is synthesized by a special software and user can interact with the system using his/her own voice and DTMF codes by typing on the telephone keypad. The vocal interface is created by transforming learning objects contained in XML files and imsmanifest.xml using XSLT. The VoiceXML document created in such a way provides several commands to make the interface more accessible: a user can listen to the content, going back and forward through paragraphs, repeat the current paragraph, pause and restart reading, and so on. The manifest file provides the user with an index of the course and an easy way to access learning objects. The LMS tracks user activity so that he/she could switch between graphical and vocal interface keeping his/her completion status and scores.

An XSL transformation is applied to the XML source file to obtain a new document in XHTML format. This transformation produces a WCAG-AAA compliant document. With the XSL transformation, the system takes the content part of the XML source and transforms its logical elements into HTML elements, in compliance with accessibility guidelines. Layout elements are implemented using CSS classes and identifiers, tag syntax is compliant to XHTML specifications, and every visual information is provided with its alternative text or title, to convey information also in a non-visual context. Picture elements in the content are transformed in image tags in the HTML document, with appropriate alternative text and title attributes; table elements are converted in HTML table, with summary attributes and separation between header and body. The transformation takes into account other two groups of information from the XML source: the page list and the paragraph list (as in the following screenshot). Items in the page list are linked to other pages of the current Learning Object. Paragraph items are instead linked to anchors in the same page (useful for enhance accessibility for users using screen readers or Braille displays). Every significant navigational link has an access key and follows a predefined sequence which can be traversed by means of the TAB key, to simplify navigation with the keyboard.

Figure 4 - An example of the electronic content automatically transcoded to AAA-compliant XHTML, a comfortable format also for users utilizing a screen reader and/or a Braille display.

4.1. High quality TTS tool for users with dyslexia

We have realized a web interface that let users (especially people with dyslexia) interact with the content through its vocal form, synthesized by a TTS engine, to support comprehension of otherwise difficult written text.
The need of having a voice as much natural as possible, has carried us to choose a TTS engine that uses "Unit Selection" concatenative technique applied to a wide range of sound samples to obtain high quality voices.

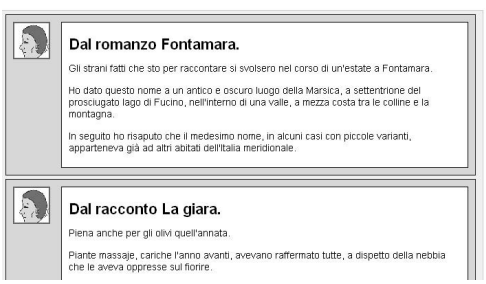


Figure 5 - each paragraph may be heard pressing a button, simplifying user's access to the content

Figure 5 shows the first modality to access to TTS engine via web page: clicking on a button, user receives an audio file (wav or mp3), containing synthesized voice form of the text. The text being read is highlighted in order to facilitate the user to locate it. The user may also want to have a portion of the text of the paragraph read, just selecting it and then pressing the reading button.

Another speech facility is supplied by the language assistant, whose interface is shown in figure 6. User can insert free text and specify some parameters for transformation (language, voice type, pitch, etc). The text is synthesized and send back as an audio file.

Figure 6 - the language assistant interface

5. Conclusions and Future Work

Within this project, we were able to efficiently repurpose existing teaching contents into a general XML schema, and repurpose it both to industrial Learning Management Systems and to the MultiAbile multimodal environment, making them accessible also to learners with disabilities using different channels. We experimented with content simplification to make the content even more accessible, and we realized how it was important to adopt a complexity metric which could give immediate feedback on the real simplification level which was achieved, at an early stage of the content preparation workflow, without having to test the content directly in the publishing environment. In the following stages of our activity, we are organizing courses for teachers and tutors, educating them in the use of assistive devices for users with disabilities, making them able to assist the learner with a disability in using assistive technology. We will subsequently make available to end users sample courses via the accessible platform, with the assistance of skilled tutors, evaluating the impact of making available electronic content in a more accessible way, in respect to traditional web platforms. We expect that the approach of standardizing the structure of the content, along with the adoption of the SCORM aggregation model, will facilitate repurposing and republishing of even more electronic content already

available, thus broadening the availability of accessible content for users with disabilities.

10. References

[1] Advanced Distributed Learning, "Definition of SCORM standards", *http://www.adlnet.org/*

[2] Bond, A., & Granlund, M. (2002). "Surfing the net: Persons with learning disability using assistive technology for information adaptation", *Tizard Learning Disability Review*

[3] Brewer J., "How People With Disabilities use the web"; *http://www.w3.org/WAI/EO/Drafts/PWD-Use-Web/Overview.html*

[4] Brown Michael K, Stephen C. Glinski, Brian C. Schmult (2001) "Web Page Analysis for Voice Browsing", *WDA 2001*

[5] ICF, International Classification of Functioning, Disability and Health, World Health Organization, 2001, *http://www3.who.int/icf/icftemplate.cfm*

[6] L. Sbattella, R. Tedesco, "Profiling users in 'Virtual Campus.' The adaptability of an advanced and distributed learning environment". *ITHET 2004*, Istanbul (Turchia).

[7] M. Cesarini, S. Guinea, L. Sbattella, R, Tedesco. "Innovative Learning and Teaching Scenarios in Virtual Campus", *ED-MEDIA 2004*, Lugano (Svizzera).

[8] Minsu Jang et al., "Web Content Adaptation and Transcoding based on CC/PP and Semantic Templates", *WWW03*, Budapest

[9] Pietro Lucisano, Maria Emanuela Piemontese, "GULPEASE: una formula per la predizione della difficoltà dei testi in lingua italiana, in «Scuola e città», 3, 31, marzo" 1988, *La Nuova Italia*;

[10] T. Barbieri, A. Bianchi, L. Sbattella, "Multimodal Communication for Vision and Hearing Impairments. Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment", *CVHI 2004*, Granada (Spain).

[11] W3C Device Independence Activity, "Multimodal Interaction Activity", *http://www.w3.org/2001/di, http://www.w3.org/2002/mmi*

[12] Yip Chung Christina, Gertz Michael, "Sundaresan Neel (2002) Reverse Engineering for Web Data: From Visual to Semantic Structures", *Proceedings of ICDE 2002*

AXMEDIS Conference

Tutorial Notes

a part

Coordinated by

Paolo Nesi, University of Florence, Italy

Kia Ng, ICSRiM, University of Leeds, UK

Jaime Delgado, Universitat Pompeu Fabra, Barcelona, Spain

Atta Badii, IRC, University of Reading, UK

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6

MPEG Standards Enabling Universal Multimedia Access

(Tutorial Description)

Christian Timmerer and Hermann Hellwagner
Klagenfurt University, Department of Information Technology (ITEC)
{christian.timmerer, hermann.hellwagner}@itec.uni-klu.ac.at

Abstract

Over the last decade, a wide spectrum of (multimedia) content has become available to an increasing number of users who desire to access it through various devices and over heterogeneous networks. Interoperability is the key for enabling transparent and augmented use of (multimedia) content across a wide range of networks and devices. Standardization efforts within the Moving Picture Experts Group (MPEG), in particular MPEG-7 and MPEG-21, aim to provide appropriate tools for achieving this goal of Universal Multimedia Access (UMA).

This tutorial provides, in the first place, the concepts of UMA and corresponding MPEG-7 metadata tools built to support these concepts. Subsequently, the vision, an overview, and the state of the art of the emerging MPEG-21 Multimedia Framework are given. Finally, MPEG-21 Digital Item Adaptation (DIA) tools which implement the "Terminal and Networks Characteristics" key element within the whole framework are illustrated in detail. The goal of MPEG-21 DIA is to achieve interoperable transparent access to (distributed) advanced multimedia content by shielding users from network and terminal installation, configuration, management and implementation issues.

1. Introduction and Motivation

The information revolution of the last decade has resulted in a phenomenal increase in the quantity of content (including multimedia content) available to an increasing number of different users with different preferences who access it through a plethora of devices and over heterogeneous networks. End devices range from mobile phones to high definition TVs, access networks can be as diverse as GSM and broadband networks, and the various backbone networks are different in bandwidth and quality of service (QoS) support. In addition, users have different content/presentation preferences and intend to consume the content at different locations, times, and under altering circumstances.

Substantial research and standardization efforts have aimed at supporting *Universal Multimedia Access (UMA)* [1][2] (Figure 1) which attempts to comply with the scenarios indicated above. The primary goal of UMA is to provide the best quality of service (QoS) or user experience with regard to the actual circumstances. In order to achieve interoperability, the Moving Picture Experts Group (MPEG) supports the concepts provided by UMA by means of normative description tools specified within the two recent standards, MPEG-7 [3][4] and MPEG-21 [5][6].

Today, there are many technologies in place to establish an infrastructure for the delivery and consumption of multimedia content. In practice, however, several elements of such an infrastructure are often stand-alone systems and a big picture of how these elements relate to each other or even fit together is not available. Therefore, MPEG-21 aims to provide an open framework for multimedia delivery and consumption. The vision of MPEG-21 is to define an open multimedia framework that will enable transparent and augmented use of multimedia resources across a wide range of networks and devices. The intent is that the framework will cover the entire multimedia content delivery chain encompassing content creation, production, delivery, personalization, consumption, presentation, and trade. During this tutorial, we will present the vision of MPEG-21, provide an overview of selected parts of MPEG-21 supported by corresponding examples and use case scenarios, and give the current state of the art.

The emerging MPEG-7 and MPEG-21 standards address UMA in a number of ways. The overall goal of the content description (i.e., metadata) standard MPEG-7 is to enable fast and efficient searching, filtering, and adaptation of multimedia content. In MPEG-7, the scalable or adaptive delivery of

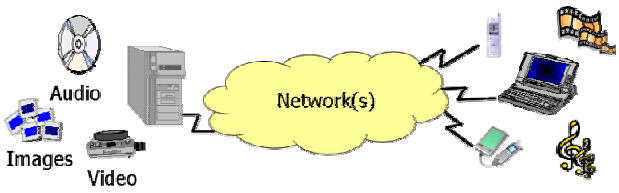


Figure 1 — Concept of UMA [6].

multimedia (in other words, UMA) is addressed by description tools for specifying transcoding hints, content variations, user preferences, usage history, space and frequency views, and summaries [7]. MPEG-21, in particular part 7, Digital Item Adaptation (DIA) [8], addresses the description of the multimedia usage environment, which includes devices and networks, among others. As such, it implements the terminal and network characteristics key element outlined in the vision (part 1) of MPEG-21 [9]. The goal of DIA is to achieve interoperable transparent access to (distributed) advanced multimedia content by shielding users from network and terminal installation, configuration, management, and implementation issues. Therefore, MPEG-21 DIA provides normative description formats (i.e., tools) enabling the construction of device and coding format independent adaptation engines [10]. Device independence is achieved through a unified description model providing information about the user characteristics, terminal capabilities, network characteristics, and natural environment in an interoperable way. Coding format independence is accomplished by means of bitstream syntax descriptions and adaptation QoS information. The former, bitstream syntax description, describes the syntax of a media bitstream using XML, in terms of packets, layers, and headers but not on a bit-per-bit basis. The resulting XML document is transformed according to the usage environment properties and is subsequently used to generate an adapted version of the actual bitstream. The latter, adaptation QoS, provides means to select optimal parameter settings for media content adaptation engines that satisfy constraints imposed by terminals and/or networks while maximizing QoS. As such, it provides the parameters for transforming the aforementioned bitstream syntax description.

In this tutorial we give a detailed overview of these descriptors and tools. However, they will not only be addressed theoretically but also by means of concrete use case scenarios and comprehensive examples.

The remainder of this overview is organized as follows. Section 2 provides an overview of the UMA concept and the challenges involved. In Section 3 we give a brief overview of MPEG-7 and provide details about its UMA tools. The main part is Section 4 which describes the MPEG-21 multimedia framework with a focus on UMA, i.e., Digital Item Adaptation. This paper is concluded in Section 5 which also provides future perspectives.

2. UMA: Concept and Challenges

As briefly outlined above and sketched in Figure 1, *Universal Multimedia Access (UMA)* denotes the concept that any multimedia content should be available anywhere, anytime, on any device, tailored to the user's needs and preferences. It is highly desirable that the access is transparent and convenient in the sense that the user does not have to trigger or even does not notice that content negotiation, adaptation, and/or personalization has to take place to enable high quality media consumption under the given circumstances. It is worth mentioning that a similar notion, termed *Device Independence*, is being pursued by the W3C in the Web context: its objective is access to a unified Web from any device in any context by anyone [11][12].

Given that (1) the multimedia content base is steadily growing and becoming richer, e.g., w.r.t. coding format or interactivity, (2) user and usage needs and preferences may vary widely, (3) there is a growing diversity in the devices used to access multimedia content, and (4) heterogeneous networks with dynamically changing characteristics may have to be traversed during content delivery, the realization of UMA represents a major problem in multimedia research and standardization. The following specific challenges are involved, which result in corresponding "building blocks" of a UMA-enabled system:

- Authoring rich, multimodal, scalable content.
- Providing multimedia content descriptions, i.e., metadata conveying, e.g., the available content variants or adaptation options.
- Providing descriptions of the delivery context, i.e., of user preferences, usage environment, device capabilities, and network characteristics.
- Negotiating, selecting, converting, adapting multimedia content (on both syntactic and semantic levels), and/or personalizing graphical user interfaces and applications.
- Managing and enforcing digital rights.

An overall important aspect is that solutions addressing these challenges must be interoperable such that, e.g., a network gateway can properly handle both the content description and device capabilities and deliver – potentially after adaptation – a conformant, suitable content variant to the device for play-out to the user. Standards and standardization bodies

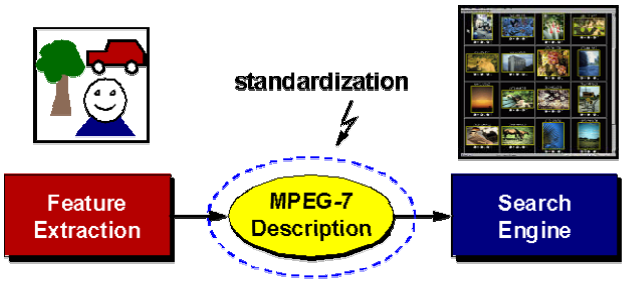


Figure 2 — Scope of MPEG-7 [3].

therefore play a central role for UMA to become reality.

This introductory part of the tutorial gives a brief overview of MPEG's efforts and achievements in addressing the challenges listed above. (Also, similar activities by the W3C in the Web context are pointed out.) The remainder of the tutorial and this paper will then specifically deal with the recent MPEG-7 and MPEG-21 tools that may be employed for UMA systems.

3. MPEG-7

3.1. Introduction

The Moving Picture Experts Group (MPEG) is mostly known as a pioneer in providing future-proven standards in the field of audio-visual encoding, compression, and transport schemes such as MPEG-1, MPEG-2, and MPEG-4. These standards found their way into applications like video/audio storage, broadcasting, streaming, or object-based manipulation of multimedia content. The next step in MPEG's standards evolution is content description facilitating technologies like automatic indexing, multimedia search engines, content-based retrieval, or personalization and summarization. The name of the resulting standard is *MPEG-7*, the *Multimedia Content Description Interface* [3][4]. MPEG-7 provides means for media logging, enterprise content management, repurposing, or even basic support for content adaptation. In general, MPEG-7 is informally referred to as providing "the bits about the bits" which makes multimedia content as searchable as text is today. The scope of MPEG-7 is limited to the description schemes (DSs) and descriptors (Ds) whereas its use, e.g., search engines or feature extraction, is open for industry competition (Figure 2).

However, the MPEG-7 standard to a certain extent provides support for UMA as well, which is summarized in the subsequent sections. This paper covers only UMA aspects of MPEG-7; for a more complete overview the reader is referred to [3]or [4].

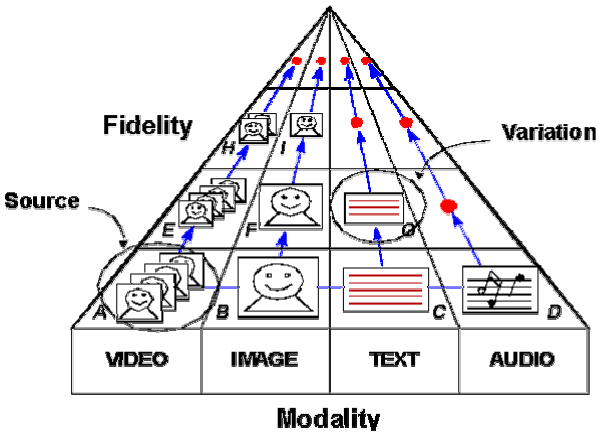


Figure 3 — Variations of a source multimedia content [7].

3.2. MPEG-7 UMA Tools

3.2.1. Variation DS

The most prominent MPEG-7 tool with regard to UMA is the *Variation DS* which describes variations of the multimedia content, e.g., low-resolution versions, summaries, different languages or even modalities of the source content, as depicted in Figure 3. This tool allows servers or proxies to select the most appropriate variation of the multimedia content according to the usage environment taking into account network conditions, terminal capabilities, and user preferences. The variations are described by a fidelity value, indicating the quality of the variation as compared to the original version, as well as the type of the variation, e.g., summary, abstract, color reduction, spatial reduction, and so forth.

3.2.2. View DS

The *View DS* describes a structural view, partition, or decomposition of a multimedia signal in space, time, or frequency. In general, views of such signals correspond to low-resolution views, spatial/temporal segments, or frequency sub-bands. An example for a space/frequency graph is depicted in Figure 4. The nodes correspond to different space and frequency views of the frame where each "S" transition indicates spatial decomposition while each "F" transition indicates frequency or sub-band decomposition.

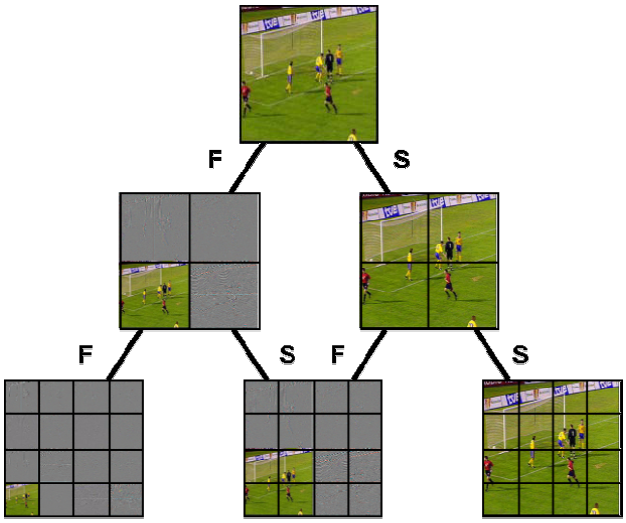


Figure 4 — Space/frequency graph describing the decomposition of an audio/video signal in space and frequency [7].

3.2.3. Summary DS

The *Summary DS* provides means for describing compact executive summaries of multimedia content in order to facilitate browsing and navigation through the multimedia content. The browsing or navigation can be performed in a hierarchical or sequential manner. The former organizes summaries hierarchically by describing different levels of temporal details. The latter composes sequences of images or video frames which are possibly synchronized with audio.

3.2.4. Media Information DS

The media features of the encoded multimedia data are described by the *Media Information DS*. It provides information about the modality, format, coding, and so forth. In general, the media information DS identifies the master media from which different profiles or instances can be derived. Media profiles are referred to as different encodings, storage or delivery formats of the master media whereas media instances represent different instantiations of the master media as physical entities by means of an identifier and locator.

3.2.5. Semantic DS

In some applications the structure of the content is less important than its semantics. Therefore, the *Semantic DS* provides an alternative approach describing the semantics of the multimedia content in an abstract way based on events, objects, places, and time in narrative worlds. For example, the Semantic DS can describe something like "U2 live in concert in Vienna's Ernst Happel Stadium on July 2, 2005".

3.2.6. Transcoding DS

The *Transcoding DS* has been introduced in order to achieve the transcoding objectives, namely maximize the content value (e.g., quality), minimize the transcoding complexity (e.g., delay), and meet the usage environment constraints (e.g., terminal and network capabilities). Therefore, transcoding hints can be used to specify the relative importance of segments, regions, objects, or audio-visual multimedia content. The spatial resolution hint specifies the maximum allowable spatial resolution reduction factor for improved perceptibility. Furthermore, the shape hint specifies the amount of shape change in the media and the difficulty hint defines the transcoding difficulty for a particular media. Finally, motion hints specify motion un-compensability and motion intensity information. The former provides the amount of new content in a segment or region and the latter the motion intensity of a segment or region.

4. MPEG-21 Multimedia Framework

4.1. Vision, Strategy, and Tools

The aim of the *MPEG-21* standard, the so-called *Multimedia Framework*, is to enable transparent and augmented use of multimedia resources across a wide range of networks, devices, user preferences, and communities, notably for trading (of bits). As such, it provides the next step in MPEG's standards evolution, i.e., the transaction of *Digital Items* among *Users*.

A *Digital Item* is a structured digital object with a standard representation and metadata. As such, it is the fundamental unit of transaction and distribution within the MPEG-21 multimedia framework. In order words, it aggregates multimedia resources together with metadata, licenses, identifiers, intellectual property management and protection (IPMP) information, and methods within a standardized structure.

A *User* (please note the upper case “U”) is defined as any entity that interacts within this framework or makes use of Digital Items. It is important to note that Users may include individuals as well as communities, organizations, or governments, and that Users are not even restricted to humans, i.e., they may also include intelligent software modules such as agents.

The MPEG-21 standard currently comprises 18 parts which can be clustered into five major categories each dealing with different aspect of the Digital Items:

```
<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS"
  xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS">
 <Item>
  <Descriptor>
   <Statement>
    <dii:Identifier>
     urn:uma:tutorial:digital_item:T800
    </dii:Identifier>
   </Statement>
  </Descriptor>
  <Choice minSelections="0" maxSelections="1">
   <Selection select_id="TERMINATOR_1"/>
   <Selection select_id="TERMINATOR_2"/>
   <Selection select_id="TERMINATOR_3"/>
  </Choice>
  <Item id="Terminator_1_Item">
   <Condition require="TERMINATOR_1"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator1.mpeg"/>
   </Component>
  </Item>
  <Item id="Terminator_2_Item">
   <Condition require="TERMINATOR_2"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator2.mpeg"/>
   </Component>
  </Item>
  <Item id="Terminator_3_Item">
   <Condition require="TERMINATOR_3"/>
   <Component>
    <Resource mimeType="video/mpeg"
     ref="rtsp://my.movies.com/arni/terminator3.mpeg"/>
   </Component>
  </Item>
 </Item>
</DIDL>
```

Document 1 — Example DID declaring a Digital Item of the fictive Terminator trilogy.

declaration (and identification), rights management, adaptation, processing, and systems aspects which are described in the following.

4.2. Declaring and Identifying Digital Items

A Digital Item is a structured digital object with a standard representation, identification, and metadata. The standard representation of Digital Items is defined by a model which describes a set of abstract terms and concepts and is expressed by the XML Schema based Digital Item Declaration Language (DIDL) [13]. The resulting XML document conformant to DIDL is called Digital Item Declaration (DID). The DID may contain several building blocks as defined in DIDL which defines the structure of the Digital Item. A brief overview of the most important building blocks is given in this paper, for further details the reader is referred to [5] or [6].

The *Item* comprises a grouping of sub-items or components. In general, an item can be considered as a declarative representation of a Digital Item. Note that an item without sub-items can be considered a logically indivisible work and an item that does contain sub-items can be considered a compilation.

The *Component* defines a binding of a multimedia resource to a set of descriptors which provides information related to all or parts of the resource. These descriptors will typically contain control or structural information about the resource such as bit rate, character set, start points, or encryption information.

The *Descriptor* associates information with the enclosing element, i.e., its parent (e.g., item) or following sibling (e.g., component). The information can be itself a component (e.g., thumbnail of an image) or a textual statement.

The *Resource* is defined as an individually identifiable asset such as video, audio clip, image, or textual asset. Note that the resource must be locatable via an unambiguous address.

Digital Items are configurable through the so-called choice/selection mechanism. A *Choice* describes a set of related *Selection*s which can affect the configuration of an item. As such it provides a generic and flexible way for multimedia content selection based on certain criteria defined by the Digital Item author. Such criteria may include rights expressions and/or usage environment constraints.

Another important aspect of MPEG-21 is the identification of Digital Items. The Digital Item Identification (DII) standard provides means for uniquely identifying DIs and parts thereof [14]. However, it is important to emphasize that DII does not define yet another identification scheme; in fact, DII facilitates existing schemes such as International Standard Book Number (ISBN) or International Standard Serial Number (ISSN) and specifies means for establishing a registration authority for Digital Items.

An example DID is shown in Document 1. The Digital Item is appropriately identified utilizing DII and provides three selections. Note that each of the three selections may contain further DIDL elements with more detailed information regarding each selection but omitted here due to space limitations. The sub-items conditionally refer to one of the selection identifiers and comprise the actual reference to the media resource.

4.3. Expressing Rights

Digital rights management (DRM) support within MPEG-21 can be divided into three parts, namely the Rights Expression Language (REL), the Rights Data Dictionary (RDD) and IPMP Components.

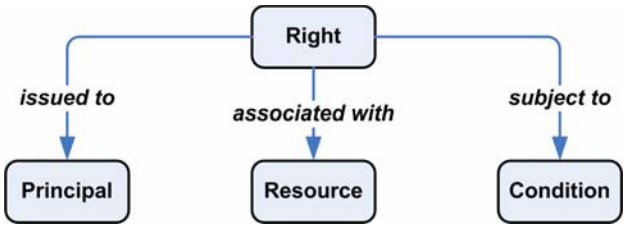


Figure 5 — REL data model [15].

The *REL* is a machine-readable language that can declare rights and permissions on digital resources [15]. The main goals of REL can be formulated as supporting guaranteed end-to-end interoperability by providing a standard way to express rights/interests and a standard way to express grants of rights. The former is used for protection of digital content as well as privacy and use of personal data. The latter specifies the access and use of controls for digital content by honoring the rights, conditions, and fees specified in the rights expressions. The REL data model is shown in Figure 5 which contains four basic entities. The right defines the action (or activity) or a class of actions that a principal may perform on or using the associated resource und given conditions, e.g., time, fee, count, territory, freshness, integrity, marking, signed-by, and so forth.

The *RDD* comprises a set of clear, consistent, structured, integrated, and uniquely identified terms to support REL [16]. The goals of the RDD are twofold. On the one hand the RDD provides a standard way to describe the semantics of terms based on their relations to other terms. On the other hand, the RDD supports mapping/transformation of metadata from the terminology of one namespace (or authority) into that of another namespace (or authority).

The *IPMP components* specify how to include IPMP information and protected parts of Digital Items in a DIDL document [17]. It deliberately does not include protection measures, keys, key management, trust management, encryption algorithms, certification infrastructures or other components required for a complete DRM system. Currently, the IPMP components consists of two parts, the IPMP DIDL providing a protected representation of the DID model, and IPMP information schemes defining structures for expressing information relating to the protection of content including tools, mechanisms, and licenses.

4.4. Adaptation of Digital Items

A vital and comprehensive part within MPEG-21 and with regard to UMA is part 7 of the standard, referred to as Digital Item Adaptation (DIA), which specifies normative descriptions tools to assist with the

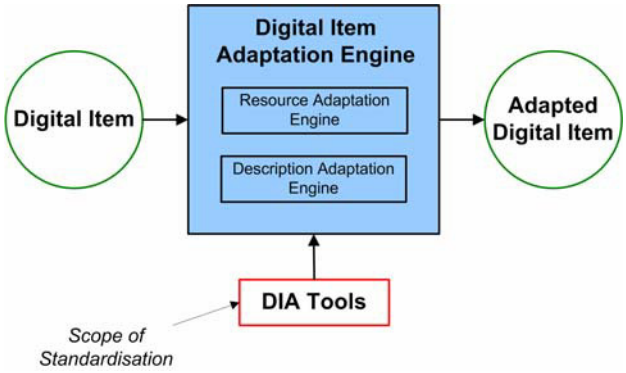


Figure 6 — Concept of MPEG-21 DIA [8].

adaptation of Digital Items [8][18]. In particular, the DIA standard specifies means enabling the construction of device and coding format independent adaptation engines. The high-level architecture of DIA is depicted in Figure 6. As shown in the figure, only tools used to guide the adaptation engine are specified by DIA, the adaptation engines themselves are left open to industry competition.

4.4.1. Tools Enabling Device Independence

The tools allowing for device independence are generally referred to as *Usage Environment Description (UED)* tools which include terminal capabilities and network characteristics as well as user characteristics and the characteristics of the natural environment. Such descriptions provide a fundamental input to any adaptation engine and a selection is briefly reviewed below.

The concept of terminal in DIA is rather generic and is representative for all kind of devices within the delivery chain including also server and intermediary network nodes. The terminal capabilities can be classified into three categories. First, the *codec capabilities* define the encoding and decoding capabilities of a terminal. As such, the supported codecs of the requesting device can be identified which may result in one or more transcoding steps of the original multimedia content. Second, the *input-output characteristics* comprise display capabilities (e.g., resolution or color capability), audio output capabilities (e.g., frequency range or number of output channels), and user interaction inputs (e.g., keyboard or touch screen). These tools control the presentation layout or the user interface of the multimedia content. Third, the *device properties* cover a wide range of tools including power and storage characteristics, and CPU benchmark measures, among others. The power characteristics include information such as remaining battery capacity which may be considered by a sending

device in such a way as to adapt its transmission strategy in order to maximize the battery lifetime. Storage characteristics (e.g., transfer rate or size) may influence how the Digital Item may be consumed, e.g., whether it needs to be streamed or locally stored. The benchmark tool enables the description of the CPU performance which could be used to infer a device's capabilities of handling a certain type of media possibly encoded at a certain quality level.

For *network characteristics* two major categories can be identified, namely static capabilities and dynamically varying conditions. The former, the network capabilities, include attributes describing the maximum capacity and the minimum guaranteed bandwidth the network can provide. Additionally, information about in-sequence delivery and how erroneous packets are handled can be signaled using this tool. The latter, i.e., network conditions, provide means for describing the currently available bandwidth, error, and delay. The main objective of these tools is to enable improved transmission efficiency and media quality optimization w.r.t. network constraints.

The *user characteristics* enable a variety of applications including adaptive content selection as well as personalization. Therefore, DIA provides means for describing general information about the user as well as her/his preferences and usage history which has been re-used from the MPEG-7 tool set. Furthermore, the presentation preferences (e.g., format or modality of the multimedia content) belong to the user characteristics. Other important aspects of the user are accessibility (i.e., certain visual or auditory impairments) and location (i.e., mobility and destination) characteristics. The former allows for adaptive delivery of multimedia content according to a user's impairment whereas the latter is important for location-based services.

Finally, the *natural environment characteristics* pertain to the physical environmental conditions around a user such as lighting conditions, noise level, or time and location where Digital Items are consumed or/and processed.

4.4.2. Tools Enabling Coding Format Independence

In order to cope with today's diversity of existing scalable coding formats, e.g., MPEG-4 or JPEG2000, a generic adaptation approach for these coding formats is desirable. DIA's response to this desire is the *Bitstream Syntax Description (BSD)* tool providing means for describing the high-level syntax of a bitstream, e.g., how the stream is organized in terms of frames, layers, or packets, utilizing the Extensible Markup Language (XML). Therefore, the Bitstream Syntax Description Language (BSDL) based on XML Schema defines restrictions and extensions taking the structure of multimedia formats into account. In addition to BSDL, the DIA standard defines two generic processors for parsing a BSD and generating the corresponding bitstream and vice versa. Furthermore, a generic BSD (gBSD) based on BSDL is defined which has been specifically designed for being used within constrained and streaming environments. It uses predefined elements guaranteeing format independence. Additionally, gBSD provides semantically meaningful marking of bitstream segments, hierarchical gBSD structure, flexible addressing schemes, and intrinsic support for distributed adaptation in terms of multi-step adaptations.

The actual bitstream adaptation can be divided into two logical steps. The first step transforms the (g)BSD (e.g., using the Extensible Stylesheet Language for Transformations (XSLT)) according to the parameters derived from the usage environment properties. The second step adapts the bitstream by means of the transformed (g)BSD according to the definition of the (g)BSDtoBin processor as specified in the DIA standard. Please note that both steps can be and should be combined for efficiency reasons.

However, the (g)BSD-based adaptation approach is only one step towards coding format independence. An integral part of media adaptation for UMA is providing the optimal adaptation parameters with respect to the UED, taking into account QoS information of the multimedia content. Therefore, DIA specifies two tools that meet the above requirements, namely the *AdaptationQoS (AQoS)* and *Universal Constraints Description (UCD)* tools. AQoS specifies the relationship between, for example, device constraints, feasible adaptation operations satisfying these constraints, and associated utilities (or qualities) of the multimedia content. The UCD enables users to specify further constraints on the usage environment and the use of a Digital Item by means of limitation and optimization constraints; e.g., the UED might describe a 1,280 x 1,024 pixel resolution display and the UCD constrains this further by informing the adaptation engine that only 70% of this area is available while the frame width and height of the multimedia content should be maximized.

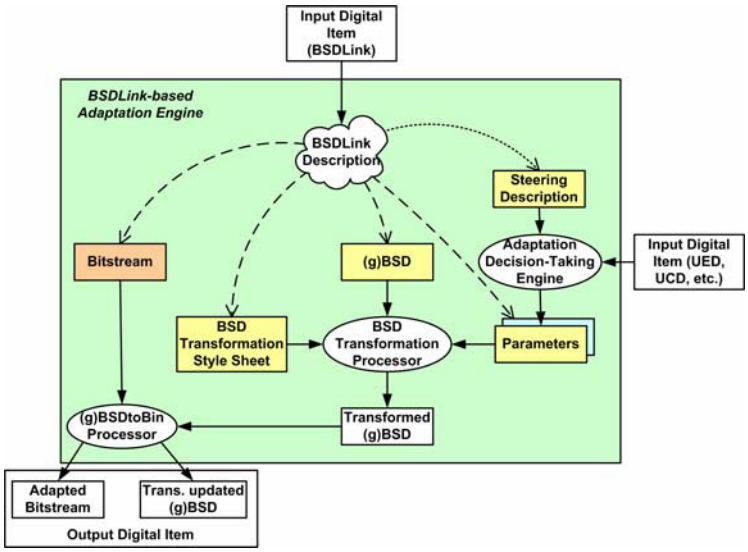


Figure 7 — The big picture of device and coding format independent multimedia content adaptation [6].

The big picture of device and coding format independent multimedia content adaptation is illustrated in Figure 7. Another tool, the *BSDLink* which is also specified in DIA provides references to the required information assets and links the steering description (i.e., AdaptationQoS) to the (g)BSD-based adaptation approach. In particular, the output parameters of the adaptation decision taking process are linked to the input parameters of the (g)BSD transformation.

4.4.3. Miscellaneous DIA Tools

MPEG-21 DIA specifies also additional tools enabling metadata adaptation, session mobility, and the configuration of a DIA engine.

Metadata has become more and more popular resulting in several adaptation issues. First, as the content is adapted, the associated metadata must be modified accordingly. Second, in cases where the metadata is transmitted and consumed, scaling may be required due to terminal and network constraints. Third, a user may be only interested in certain parts of rich and detailed descriptions of the content, i.e., these parts must be filtered or extracted in an efficient way. Finally, if multiple sources of metadata for the same media resource exist, an efficient mechanism for integrating all these assets into a single description may be required. DIA specifies tools to assist with all of the above issues which are generally referred to as *Metadata Adaptation*.

The *Session Mobility* tool refers to the transfer of configuration state information of a Digital Item, i.e., the instantiation of choices and selections within the DID, from one device to another one. This enables the Digital Item to be migrated to, and consumed on, a different device in an adapted way. Note that the application state information, i.e., the information specific to the current application, may be also transferred.

The *DIA Configuration* tools are used to help guide the adaptation process considering the intentions of a DID author. One means by which this is achieved is by allowing authors and providers of Digital Items to specify useful DIA descriptions that would help to either configure the DID or adapt the resources according to the usage environment in which they will be consumed. Another means are particular tools that have been specified that guide the DID configuration process.

4.4.4. Conversions and Permissions

The first amendment to DIA facilitates the description of fine-grained media conversions by means of the conversion name and its parameter, which could be used to define rights expressions to govern adaptations in an interoperable way [19]. The conversion descriptions can be used to identify suggested conversions for particular resources, or to describe terminal capabilities in terms of its supported conversions. The name of the conversion references (specialized) RDD terms which also defines its parameters. Furthermore, the conversion descriptions can be included into so-called permitted DIA conditions and change constraints enabling seamless integration of DIA tools in DRM-aware environments.

4.5. Processing of Digital Items

The declaration of a Digital Item defines its structure, but still a DID is static. The question what happens when a DID arrives at a terminal remains unanswered so far. The *Digital Item Processing (DIP)* standard, MPEG-21 part 10 [20], allow Users to add functionality to a DID: on receipt of a DID a list of Digital Item Methods (DIMs) that can be applied to the Digital Item is presented to the User. The User chooses a Method which is then executed by the DIM Engine (DIME).

DIMs provide a way for Users of a Digital Item to select preferred procedures by which the Digital Item should be handled. Note this is done at the level of the Digital Item itself and it is not intended to be utilized for implementing the processing of media resources

themselves. As such, DIMs are basically a "list of operations" specified in the normative DIM Language (DIML) for which ECMAScript has been selected. DIMs may utilize Digital Item Base Operations (DIBOs) or Digital Item eXtension Operations (DIXOs). The former is a set of normative basic operations on which DIMs are built analogous to a standard library of functions of a programming language. DIBOs are atomic operations and are defined using a normative, high-level interface. The latter is used as an extension mechanism enabling interoperable execution of user-defined operations. Currently, Java is exclusively used for DIXOs but C++ bindings are under development within the first amendment of the DIP standard.

4.6. MPEG-21 Systems Aspects

The MPEG-21 systems aspects include the MPEG-21 file format (.mp21), the binary format for XML-based MPEG-21 descriptions, and an activity just started named Digital Item Streaming.

The DID declares a Digital Item using XML which includes textual assets such as metadata or licenses and references to the actual multimedia content. Nonetheless, the DID is declarative and does not provide a physical container including all the assets of a Digital Item. Additionally, it is not possible to embed binary content within XML in an efficient way – base64 encoding results in approximately 33% overhead. The *File Format* has been defined to solve these issues and is based on the ISO base media file format [21] which has been extended by specific MPEG-21 requirements such as the 'meta' box with the mp21 metadata handler.

Another important issue for MPEG-21 is the efficient storage and transport of MPEG-21-based metadata. Therefore, MPEG's binary format for metadata (BiM) [22] has been adopted for the *Binary Format* of the MPEG-21 standard. BiM defines an alternative schema aware XML serialization format which adds streaming capabilities to XML documents, among other useful features.

Finally, the incremental delivery of a Digital Item in a piece-wise fashion and with temporal constraints such that a receiving peer may incrementally consume the Digital Item has been acknowledged within MPEG-21. Therefore, the *Digital Item Streaming* activity has been launched providing answers on how to fragment XML documents which are heavily used within MPEG-21, how to assign time information to those XML fragments, and how to specify the streaming of different components of a Digital Item. It is expected that adequate answers to the aforementioned questions will be provided within this new part 18 of MPEG-21.

5. Conclusion and Future Perspectives

The tutorial introduces the notion of Universal Multimedia Access (UMA) and the challenges involved in achieving this desirable mode of accessing and consuming multimedia content. The Moving Picture Experts Group (MPEG) has addressed the UMA challenge by the recent MPEG-21 (Multimedia Framework) family of standards, but MPEG-7 (Multimedia Content Description Interface) provides descriptions tools supporting UMA as well. The tutorial gives an overview of the concepts incorporated in the relevant standards, provides examples of their usage, and demonstrates reference and utility software that makes use of these achievements.

Standards usually provide only a framework ensuring interoperability while there are still open issues and research challenges left. For UMA we briefly discuss two issues which we would like to see being addressed in the near future. First, today many utility measures for certain media modalities (e.g., visual or audio content) exist. One of the most important challenges w.r.t. UMA is an investigation towards a common utility measure across different modalities for a diverse set of available scalability dimensions (e.g., spatial, temporal, signal-to-noise ratio, or color) taking into account the variety of user preferences and characteristics. We believe such utility measures are indispensable for a satisfactory and universal multimedia experience. Second, MPEG standards supporting UMA specify only the format and deliberately exclude transport, exchange, negotiation, and management issues. These issues have to be solved in practice when building end-to-end systems facilitating UMA concepts by using traditional network protocols. On the one hand, the information of the usage environment needs to be captured and transported to the content provider either directly in a request for content or as a separate message. At the content provider side, this context information needs to be managed appropriately. On the other hand, content-related (timed) metadata has to be transmitted along with the media data it describes resulting in various synchronization and efficiency issues.

While thus basic technology enabling UMA is in place in terms of recent multimedia (metadata) standards, it is still open whether or not and how these standards will be adopted by industry to create UMA-ready content and applications. Given even today's heterogeneity of multimedia content, devices, and

networks, fast and wide deployment of this technology would be highly desirable in order to make future multimedia systems and applications easy and enjoyable to use.

6. Acknowledgments

The authors are most grateful to John R. Smith for co-authoring the first UMA tutorial presented at WWW'2004 and providing most of the MPEG-7 related material; Rik Van de Walle, Fernando Pereira, and Ian S. Burnett for providing parts of the general MPEG-21 related material; and Anthony Vetro for providing parts of the MPEG-21 DIA related material.

7. References

[1] R. Mohan, J. R. Smith, and C.-S. Li, "Adapting Multimedia Internet Content for Universal Access", *IEEE Transactions on Multimedia*, vol. 1, no. 1, pp. 104-114, Jan.-Mar. 1999.

[2] A. Vetro, C. Christopoulos, and T. Ebrahami, eds., *IEEE Signal Processing Magazine, special issue on Universal Multimedia Access*, vol. 20, no. 2, March 2003.

[3] S.-F. Chang, A. Puri, T. Sikora, and H. Zhang, eds., *IEEE Transactions on Circuits and Systems for Video Technology, special issue on MPEG-7*, vol. 11, no. 6, 2001.

[4] B.S. Manjunath, P. Salembier, and T. Sikora (eds.), *Introduction to MPEG-7: Multimedia Content Description Language*, John Wiley & Sons Ltd., NJ, 2002.

[5] F. Pereira, J. R. Smith, and A. Vetro, eds., *IEEE Transaction on Multimedia, special section on MPEG-21*, vol. 7, no. 3, June 2005.

[6] I. Burnett, R. Koenen, F. Pereira, and R. Van de Walle (eds.), *The MPEG-21 Book*, Wiley, 2006 (to appear).

[7] P. Salembier and J. R. Smith, "MPEG-7 multimedia description schemes", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, 2001, pp. 748-759.

[8] A. Vetro and C. Timmerer, "Digital Item Adaptation: Overview of Standardization and Research Activities", *IEEE Transactions on Multimedia*, vol. 7, no. 3, June 2005, pp. 148-426.

[9] ISO/IEC TR 21000-1:2004, *Information technology — Multimedia framework (MPEG-21) — Part 1: Vision, Technologies and Strategy*, 2004.

[10] C. Timmerer and H. Hellwagner, "Interoperable Adaptive Multimedia Communication", *IEEE Multimedia Magazine*, vol. 12, no. 1, January-March 2005, pp. 74-79.

[11] World Wide Web Consortium, *Device Independence Activity*, available at http://www.w3.org/2001/di/.

[12] World Wide Web Consortium, *Mobile Web Initiative*, available at http://www.w3.org/2005/MWI/.

[13] ISO/IEC FDIS 21000-2, *Information technology — Multimedia framework (MPEG-21) — Part 2: Digital Item Declaration*, N6927, Hong Kong, China, January 2005.

[14] ISO/IEC 21000-3:2003, *Information technology — Multimedia framework (MPEG-21) — Part 3: Digital Item Identification*, 2003.

[15] ISO/IEC 21000-5:2004, *Information technology — Multimedia framework (MPEG-21) — Part 5: Rights Expression Language*, 2004.

[16] ISO/IEC 21000-6:2004, *Information technology — Multimedia framework (MPEG-21) — Part 6: Rights Data Dictionary*, 2004.

[17] ISO/IEC FCD 21000-4, *Information technology — Multimedia framework (MPEG-21) — Part 4: Intellectual Property Management and Protection Components*, N7196, Busan, Korea, April 2005.

[18] ISO/IEC 21000-7:2004, *Information technology — Multimedia framework (MPEG-21) — Part 7: Digital Item Adaptation*, 2004.

[19] ISO/IEC 21000-7:2004/FPDAmd 1, *DIA Conversions and Permissions*, N6937, Hong Kong, China, January 2005.

[20] ISO/IEC FDIS 21000-10, *Information technology — Multimedia framework (MPEG-21) — Part 10: Digital Item Processing*, N7208, Busan, Korea, April 2005.

[21] ISO/IEC 14496-12:2005, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*, 2005.

[22] ISO/IEC FCD 23000-1, *Information technology — MPEG-B — Part 1: Binary MPEG format for XML*, N7251, Busan, Korea, April 2005.

Author Index

Axmedis 2005, edited by P. Nesi, Kia Ng, J. Delgado ISBN 88-8453-354-6