StreamOnTheFly: a Peer-to-peer Network for Radio Stations and podCasters

Roland Alton-Scheidl[1], András Micsik[2], Máté Pataki[2], Wolfgang Reutz[3],
Jürgen Schmidt[4] and Thomas Thurner[5*]

[1]*Vorarlberg University of Applied Sciences*
roland@alton.at

[2]*Computer and Automation Research Institute of the Hungarian Academy of Sciences*
micsik@sztaki.hu

[3]*Vorarlberg University of Applied Sciences*
wolfgang.reutz@fh-vorarlberg.ac.at

[4]*strg.at*
juergen@strg.at

[5]*Team Teichenberg*
thomas@teichenberg.at

[*] *Authors' names are in alphabetical order*

Abstract

The StreamOnTheFly network demonstrates new ways of management and personalisation technologies for audio. The architecture is based on a decentralized network of software components using automatic metadata replication in a peer-to-peer manner. The network also promotes a new common metadata schema and content exchange format. Content reuse and content exchange is made possible by StreamOnTheFly in several use cases.

1. Introduction

With the deregulation of the radio market in the 90s, not only commercial stations entered the aether, but also a number of non-profit radios. Community radio stations would like to be able to archive, exchange and stream content over the Internet, but appropriate tools have not been available at an affordable price. They rely on exchanging content and expressed their demand to set up a robust distributed infrastructure with a technological solution which is flexible, effective and low-cost.

In 2002, the StreamOnTheFly project with the participation of Public Voice Lab (Austria), MTA SZTAKI (Hungary) and Team Teichenberg (Austria) gained grants from the European Commission's Information Society Technology Program to build a middleware application for radios with various front ends.

After two years of development and user trials, StreamOnTheFly provides an open-source, station-management interface, a programme exchange platform and a portal engine for radio programme reuse. All software has been built with Open Source Technologies subject to Open Source Licences, providing flexible and cost-efficient solutions for both small community radio stations and larger broadcasters.

StreamOnTheFly is focused upon next-generation audio content management and broadcasting offering a customisable community radio program. StreamOnTheFly demonstrates new ways of management and personalisation technologies for audio. Many programs produced at community radio stations are too valuable to be broadcasted only once. The StreamOnTheFly consortium has created recycling technologies, making them available standalone or networked for syndication.

It has been demonstrated at various events, such as IST Copenhagen in 2002 or at Ars Electronica 2003. Nowadays, StreamOnTheFly has gained popularity among a number of central European radio stations and is now also serving the podCasting community.

This paper shows our technical approach, demonstrations and exploitation opportunities.

2. StreamOnTheFly Architecture

The project aimed to create a network of radio content that develops in an organic and dynamic way, in decentralized structures that build synergies on an international, community-based level. The content is offered by local producers or stations to the network, but it's the international communities who choose
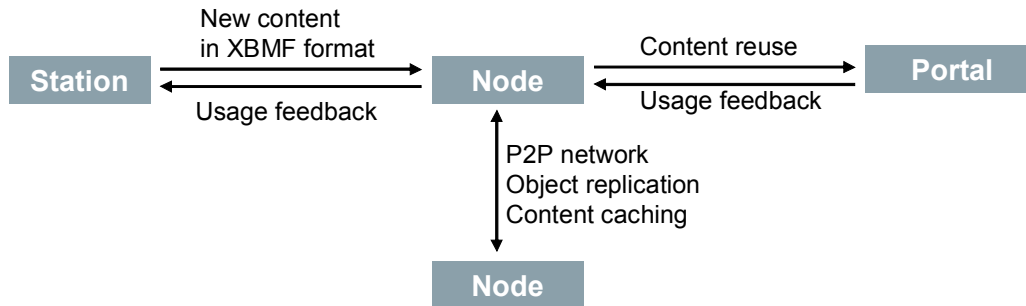
Figure 1. Basic components of StreamOnTheFly network

which content is presented and how it is presented. International publishing on the net is not a form of archiving or recycling anymore, but a motivation to communicate with a whole new group of people.

The three main components of StreamOnTheFly are (Figure 1):

- Station control: A web application to manage all aspects of a radio station, from scheduling to recording, archiving and publishing selected programmes to Node Servers.
- Node server: A distributed network of servers for content. It can be seen as a distributed archive or digital library providing the necessary infrastructure for the dissemination of content and for the collection of user feedback. Users may browse for interesting materials, and compile personalised radio streams for listening.
- Portal server: Through easily administered portals, the audio content of the network can be presented through various interest filters. This tool could, for example, be used to set up a portal dedicated to film reviews.

For example, Orange 94.0, a community radio in Vienna, may send some of its broadcast material on the node running in Austria. These programmes are archived at the node and become accessible for the whole StreamOnTheFly network for a given period of time. A German radio may find a programme on the Austrian node that they want to rebroadcast locally. They find the contact for this programme, in this case Radio Orange and ask for their permission for rebroadcast. If there is a portal in France for the Turkish people, its portal editor may include some of the Turkish programmes of Radio Orange in the recommendation list.

The StreamOnTheFly network tries to serve the whole lifecycle of a radio programme, starting with the creation and ending with various (re)uses and discussions of the programme. The three main components of the network are introduced in the following subsections.

2.1. Nodes

Nodes implement the basic infrastructure for the network. Each node hosts a set of radio stations. For each station the node archives radio programmes with a rich set of metadata and other associated content.

Users get transparent access to the contents of the whole network at any node. They can browse through radio stations or topic trees or search in the archive. Continuing our previous work [8], an advanced query facility is integrated which enables users to construct and save complex queries. Queries are built step by step by composing atomic expressions (e.g. broadcasted before year 2004) and joining them with logical AND/OR. In atomic expression the query target can be any field of the metadata scheme. The possible syntax offered for the user changes according to the type of the expression target (string, date, number, enumeration, etc.). These queries can be saved with a name in the users' profiles, and later re-executed easily.

Interesting items can be collected into a playlist. These playlists can also be used as a kind of personalized radio. Users may listen or download audio content in different formats (e.g. MP3, Ogg/Vorbis) and in different bit rates or quality.

The network gathers and forwards information about radio programmes back to their creators. This involves the collection of various statistics (download, listen, etc.), the collection of comments, references and rating information.

Editors have a comfortable 'console screen' where they get an overview about all their programmes, they can publish new programmes and manage existing ones. Access privileges can be controlled on various objects: stations, series, programmes etc.

2.2. Stations

A station management tool has been developed by the project, which connects radio stations to the

network. This tool provides an easy way to feed a node with radio programmes. Furthermore, it helps in the scheduling work of the station, and leads the creators' hands before and after the broadcast of a radio programme. Authors may add metadata to their programmes and publish them on a node.

The station management tool handles editors and other users of a radio with proper access privileges. Programs and series can be scheduled for broadcast by the creators or editors, and station managers control the weekly broadcast plan. This tool provides most of the basic required software support for a radio.

A recording tool can be used when no prerecorded audio is available for a radio programme (e.g. live shows). This software records the broadcasted audio of a station, and automatically cuts the stream into individual radio shows. There is a desktop upload tool under development, enabling creators to assemble metadata and audio files into a program package and send it directly to the node from their personal computer.

2.3. Portals

In the StreamOnTheFly project a radio portal offers a selection of radio programmes in a custom design. Portals may serve ethnical groups, topics or geographical regions by selecting relevant content from the whole network. Additionally, a portal can offer editorials and more readings or pictures about the selected programmes. Visitors of a portal may listen to programmes individually or listen to a continuous stream playing the list of selected programmes over and over again. Visitors may also rate, comment and discuss programmes. Comments and ratings made on portals are forwarded to the nodes, and the comments are delivered to the creators of the programme via e-mail.

The StreamOnTheFly portal is a specialized CMS, thus the main activities are managing content and managing appearance/design. Appearance is controlled using a special HTML based WYSIWYG editor, which can be used from any web browser, without installing any extra programs, components or plugins. Basically, the user can arrange a table as the skeleton of the portal page, where each cell is filled with design elements or content.

2.4. Content and Metadata Format

Community radios and other smaller audio producers are in need of common data exchange solutions. XBMF (Exchange Broadcast Binary and Metadata Format) was planned and established based on these user needs, with the participation of producers. The requirements did not allow for the complexity of emerging formats supported by large broadcasters [1,2,3].

XBMF enforces the coupling of metadata with content, and also creates possibility for providing content in different audio formats, or attaching different media (images, text) to the audio content.

Standardisation efforts have been performed in collaboration with public broadcasters resulting in the XBMF metadata description for radio shows [6]. The metadata part of XBMF is based on the latest status of Dublin Core [4] and uses the application profile building rules for metadata schema definition. XBMF metadata is also compatible with the recommendation of the European Broadcasting Union (EBU) on core metadata set for radio archives [5].

XBMF can be used for transfer and archival of radio programmes coupled with metadata. This is a simple container format, in which metadata (in XML format), content files, and other associated files can be packaged together.

Technically, an XBMF file is a compressed directory structure, where naming conventions and placement of files identify the data they convey. Implementation of software for handling XBMF files is straightforward, and existing open software can be used.

The metadata schema for XBMF can be seen as an application profile of Dublin Core for radios. This profile is constructed from DC elements and the StreamOnTheFly element set according to the rules implemented for metadata schema registries in the CORES project [11].

3. Bridges to and from StreamOnTheFly

As new applications of the StreamOnTheFly suite appeared, several new connections were established with other systems (Figure 2). The basic goals of these connections are

- to enable the reuse of the content stored in the network,
- and to enable sending new content into the archive, thus migrating or merging existing archives.

For the goal of content reuse RSS [16] proved to be the most supported approach. Up to now, Manila, Plone and Typo3 content management systems have integrated StreamOnTheFly content via RSS feeds.

RSS feeds of the StreamOnTheFly node not only offer the latest content from radios, series or people, but also provide all important information for single radio shows including metadata, listening URLs and statistics. With this foundation, new frontends can be
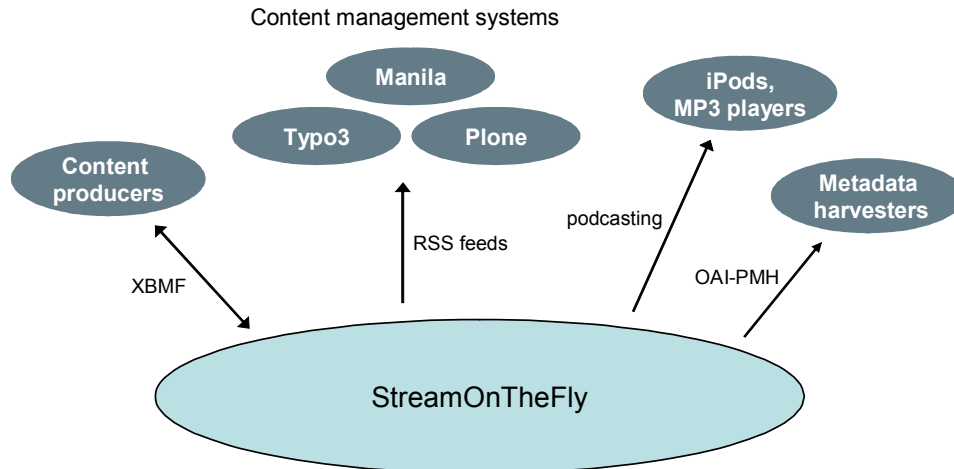
Figure 2. Bridges between StreamOnTheFly network and external systems

built easily for the integration of node content into weblogs, static webpages or portals

An example for connecting StreamOnTheFly with a portal engine is the case of Campus Radio "Achwelle". The Vorarlberg University of Applied Sciences supports the StreamOnTheFly network and maintains a node for its campus radio. Radio Achwelle started broadcasting in 1997. In 2005 the distribution channels were extended with a 24/7 livestream and a livestream with all archived broadcasts. For these new possibilities a new website was implemented using Plone, an open source, standards compliant (XHTML, CSS, WAI-AA) content management system.

XBMF and the XBMF import feature on the node are used to upload data from Plone to the node.

For the display of StreamOnTheFly content from the node on the Plone CMS we use RSS 2.0 feeds. In order to extend the amount of exported information, a new namespace was defined for RSS 2.0 which allows to access single data elements of the stations, series, programmes and even queries. This namespace supports sending very detailed information also about available media files.

This bridge is also included in the Plone product "PloneRadio", developed under an Open Source license (GNUPL 2). PloneRadio [15] allows the display of up to two independent livestreams in two qualities each in a convenient portlet with display of currently playing track with album cover art and a list of the last played tracks. The other part of the project enables the use of Plone as a portal to StreamOnTheFly nodes, allowing established community platforms to deploy nodes as media archive.

RSS 2.0 gave rise to a new phenomenon called podcasting. With special software one can download not only the new headlines like in traditional RSS, but also audio files for radio shows. Users select series they like and download new episodes directly to their iPod or other MP3 listening device. After that, they can listen their favourite programs anytime anywhere. Since early 2005, StreamOnTheFly is capable of serving content for podcasters.

As an experimental feature, playlists or query results created by users can also be offered via RSS or podcasting. This way not only series editors, but any user may create his/her own sharable personalised radio.

An Open Archives Initiative (OAI) data provider [17] has also been implemented for the node, which is another way of offering data export for external systems.

For the goal of new content archival, XBMF acts as a base technology. XBMF can be compiled by any application following the open specification. Nodes have a special directory for incoming new or modified content as XBMF. Any usual network protocol can be used to upload XBMF into this directory, for example rsync, ftp or scp. Incoming XBMF files are periodically processed by the node software, radio shows are accordingly created or updated, and audio files are automatically converted to recommended formats. There are several ongoing efforts to implement desktop tools for the creation and upload of XBMF to StreamOnTheFly nodes.

4. Current Usage

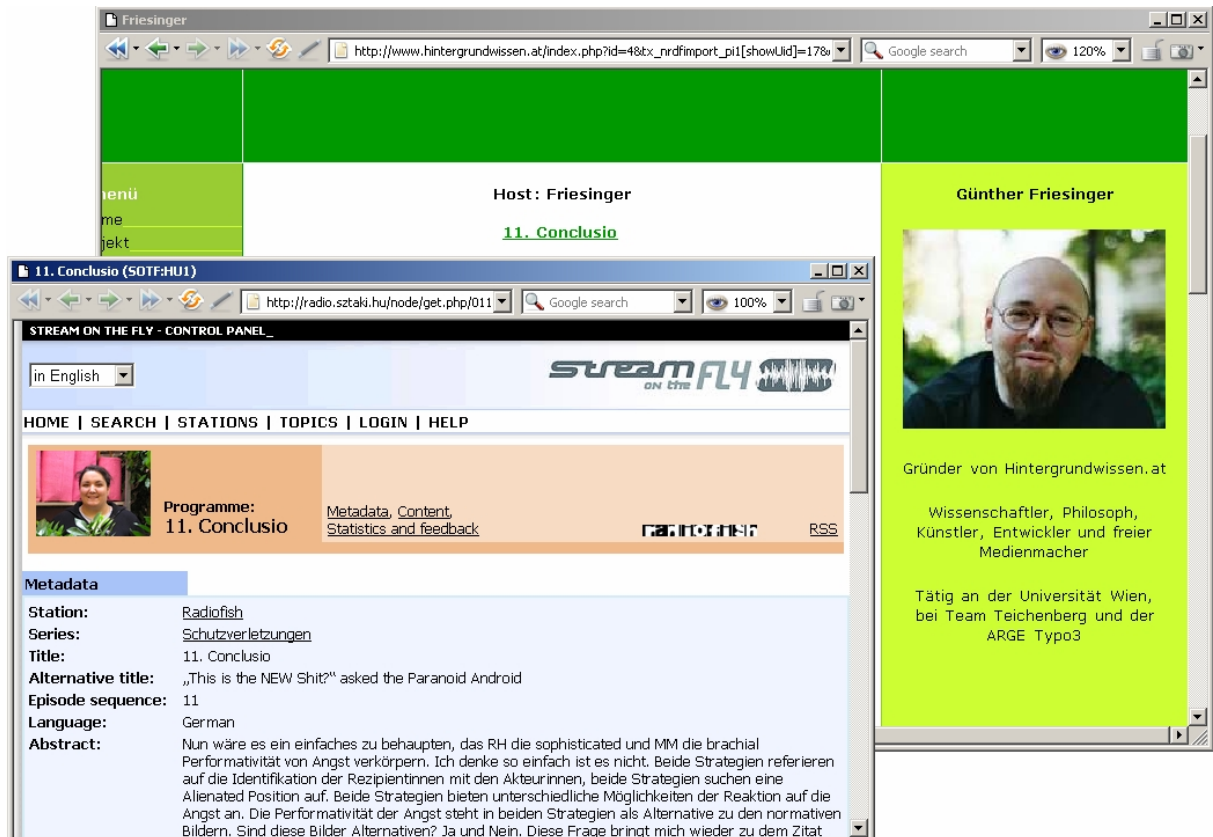The StreamOnTheFly network is operational since October 2003. The core of the network (excluding

Figure 3. Audio content in the archive and reused on a portal

experimental and stand-alone nodes) is accessible in 4 languages, and contains more than 400 hours of audio content (http://radio.sztaki.hu).

The first real use case of the system is at the Technical Museum in Vienna. Medien.welten is the permanent exhibition on the development of media at the museum. To illustrate the convergence of radio and computers, StreamOnTheFly is used to record the complete Ö1 Programme, segment and categorize it automatically and publish it as a 30 day archive in the exhibition. Due to copyright issues this node is not available outside the museum.

The Volkswirtschaftliche Gesellschaft uses StreamOnTheFly technology to run a webradio station at www.schulradio.at. The project started late summer 2004, and will include up to 10 involved schools in the "lower Austrian region".

For the third time, Christian Berger's team had broadcasted live from the Frankfurt Book Fair. In 2004, the whole webcast and archive was realized with StreamOnTheFly technology. The Frankfurt Book Fair is Europe's largest come-together of the European literature scene. Literadio is a joint project of German and Austrian community radios, and covers lectures, round tables, speeches and interviews in a daily webcast from 10am to 17pm live from the studio on Literadio's stand. The shows were stored at a node, and special integration with other portals was done using the RSS feed of the node. Another similar project based on StreamOnTheFly was the live radio broadcast at the 2002 IST conference in Copenhagen.

An interesting trial based on StreamOnTheFly was developed to proof-of-concept stage at the Vorarlberg University of Applied Sciences, where the node was used as a backend for a music publication site. In this case, the user interface was completely developed in Typo3, a PHP-based Content Management System. The interaction with the node was designed completely by using the XML-RPC API and the RSS distribution methods.

Radioswap.net uses this platform for content collection for Belgian community radio stations. At the moment, their node cannot join the network because of legal and copyright issues. They are also working on a desktop tool for easy upload of radio shows to StreamOnTheFly nodes.

The XBMF metadata schema as a derivative of Dublin Core is fitting well with the standard formats of

eLearning. IEEE's Learning Object Metadata and SCORM are easily connectable to XBMF. Developing an own eLecturing product called "AudioLectureLib"; Team Teichenberg broadens the spectrum of possible usages [12]. A bridge between content syndication in broadcast sector, eLearning and eGovernment was established. The system is now used in the scientific context (Austrian Academy of Science) as well as in eGovernment for bringing Lectures to desperate located development agents of rural areas (NÖ Dorf- und Stadterneuerung). But potential of the bridging of these various worlds goes far beyond. First trials to cover artists' works as well as develop further augmented services are planned and underway.

As a follow up of efforts already started within the StreamOnTheFly project, a transnational consortium (Slovenia, Slovakia, Austria, Germany, and Italy) is using the repository developed for creating a European archive of spoken word content. The "Acoustic Literature Map of Europe" takes existing modules of StreamOnTheFly, and integrates new user interface and taxonometries [13]. Connecting with Typo3 as CMS, we get full compatibility to existing nodes and can provide an adapted GUI to the literature community for their better usability.

StreamOnTheFly is also being used to network campus radio stations. Under the umbrella of the Competence Network for Media Design [14], educational institutions are joining forces to enable local content production under Creative Commons licenses. Meanwhile, the network, as being open to other partners in the design world, is also serving concepts and content to agencies, IT-companies or cultural organisations.

The StreamOnTheFly suite is open-source and licensed under the GPL. The StreamOnTheFly suite uses only established and open protocols and standards making the content accessible to all. The StreamOnTheFly suite is agnostic towards content right and licensing schemes thus freeing the operators of node networks to implement their own policies. No digital rights management is presupposed, and while watermarking and DRM technologies may be integrated, they need not be.

5. Related work

Cooperation in archival and programme exchange is an old demand of the radio community, and there are some services supporting this, for example the Open Meta Archive (http://oma.sourceforge.net) and the CBA (Cultural Broadcasting Archive, http://cba.fro.at). These services are all centralized and lacking the full

support of the radio "supply chain" from production to listening.

Personalisation of radio or television programme is getting more and more interest nowadays. An art project and experiment called Frequency Clock [18] provided a broadcasted TV channel where the programme could be modified during broadcast through a web server.

Last.fm is a personalised online radio station where users create their profiles by selecting songs to listen, and then they are getting recommendations of new songs. Another similar experiment was called flycasting [10], where online rating influenced the broadcast of a radio.

StreamOnTheFly creates a unified platform for these ideas of personalisation including rating, collaborative filtering, personalised radio programme, personalised queries and portals.

6. Future Plans

Authors also try to broaden the applicability of StreamOnTheFly software with two additional areas: spoken word and video.

Within the last years the needs and the potentials for media content like audio and video accrued to a permanent request from costumers and audiences. In the field of print-media this trend causes the need for spoken words content format in high quality. Especially, the audience of original print media in the web is searching for such features.

In a StreamOnTheFly trial, printed content is recorded in cooperation with audio-studios and professional speakers and the resulting audio is optimized in a streaming format. Since StreamOnTheFly uses features to re-encode audio-files, it can provide the content in different qualities for the special needs of customers.

In the trial system, users are able to walk through the original print media website, read and listen content, and see semantic relations between content items available as text or as spoken words. Features of playlists and search results enable users to select articles on the original website, and listen to the collected playlist as a result of cooperation with the StreamOnTheFly node server.

Such a seamless integration of the node server features into the original website allows enriching print media portals with the use of new media formats.

According to the wide spread of flat rate connections, streaming of video content is getting more and more affordable. Free TV stations or commercial video-producers could use the StreamOnTheFly to archive video content and meta information, and to

make it publicly available through the net. The large content files used by video are one of the main restrictions for this use case. The traditional workflow of StreamOnTheFly is to upload content to the node and execute final media encoding there. In case of video, these tasks are preferred to be done on the local infrastructure in order to discharge the server system. This is achieved by implementing user friendly stand-alone XBMF generators, which help users to encode their video content, supply it with metadata, and reliably move generated XBMF to the node.

The metadata schema of XBMF can be easily extended to store also video related information. Since that, basic functionality of the node, the included search engine, ontologies and other relations between content items are ready to use with video content as well.

7. Summary

StreamOnTheFly is an operating example for a distributed network of heterogeneous services. The core of the network combines peer-to-peer networking and open archives principles, while the rest of the services are realised as separate network components communicating through open APIs.

The network fulfills the initial goal to establish a common format and software support for the archival, exchange and reuse of radio programmes, and it also proceeds towards experimentation with latest networking paradigms and social self-organization.

Although the StreamOnTheFly project officially finished in June 2004, voluntary work keeps going on implementing various extensions and use cases. Since that time StreamOnTheFly has proved its flexibility and extensibility in several cases. The architectural decisions made in StreamOnTheFly justified themselves as an open and scaleable interface to further developments.

Focusing on connecting possibilities of StreamOnTheFly, RSS technology and podcasting proved to be the most favoured and efficient for content reuse. In the topic of content exchange, XBMF defined by the project holds promising possibilities.

With StreamOnTheFly software solutions, community radios now have the option to reuse their content either in an exchange or as an accessible archive over the internet. Through RSS as well as a special portal engine, audio content can be syndicated to websites worldwide.

8. References

[1] Society of Motion Picture and Television Engineers (SMPTE), http://www.smpte.org

[2] AAF (Advanced Authoring Format) Association, http://www.aafassociation.org/

[3] MPEG-7, http://www.mpeg-industry.com/

[4] Dublin Core Metadata Initiative, http://dublincore.org/

[5] EBU Core Metadata Set for Radio Archives (Tech3293), http://www.ebu.ch/tech_32/tech_t3293.pdf

[6] Shared Online Media Archive (SOMA) metadata format 1.0, http://soma-dev.sourceforge.net/

[7] András Micsik, Thomas Hassan, László Kovács: "Distributed Archive and Web Services for Radio Stations", 12th International World Wide Web Conference, 20-24 May 2003, Budapest, Hungary

[8] L. Kovács, A. Micsik, B. Pataki, I. Zsámboki: "AQUA (Advanced Query User Interface Architecture)". Research and Advanced Technology for Digital Libraries, proceedings of the 4th European Conference, ECDL 2000 (José Borbinha, Thomas Baker eds.), Lisbon, Portugal, September 2000, Springer LNCS 1923

[9] András Micsik, László Kovács, Robert Stachel: Collaboration of loosely coupled repositories using peer-to-peer paradigm. DELOS Workshop: Digital Library Architectures 2004: 85-92

[10] Hauver, D. B. and J. C. French. "Flycasting: Using Collaborative Filtering to Generate a Playlist for Online Radio." International Conference on Web Delivering of Music (WEDELMUSIC'2001).

[11] Rachel Heery, Pete Johnston, Csaba Fülöp, András Micsik: "Metadata schema registries in the partially Semantic Web: the CORES experience". 2003 Dublin Core Conference, DC-2003, 28 Sept - 2 Oct 2003, Seattle, Washington USA

[12] AudioLectureLib, http://www.teichenberg.at/all

[13] Acustic Literature Map of Europe http://locker.teichenberg.at/alme/node/www

[14] Competence Network for Media Design, http://www.media.coop

[15] PloneRadio, http://ploneradio.sourceforge.net

[16] RSS 2.0 Specification, http://blogs.law.harvard.edu/tech/rss

[17] Open Archives Initiative, http://www.openarchives.org/

[18] The Frequency Clock, http://www.frequencyclock.ne