



ECP-2005-CULT-038202
www.DISMARC.org

Annual Report

Period: September 1 2006 - August 30 2007

Deliverable number/name	D Report 2 (Annual)
Dissemination level	Public
Delivery date	September 24 2007
Status	Final
Author(s)	Martin Gordon/Johannes Theurer



eContentplus

This project is funded under the *eContentplus* programme¹, a multiannual Community programme to make digital content in Europe more accessible, usable and exploitable.

¹ OJ L 79, 24.3.2005, p. 1.

<i>page number</i>	
Table of contents	2
Project Objectives	3
Consortium	4
Project Results	5
Target Users	9
Underlying Content	10
Summary of Activities	12
Impact & Sustainability	14

2 Project Objectives

The DISMARC project addresses the need to create wider access to the contents of European audio archives, museums, broadcasters and private collections. Access to audio and audio-related material in European archives has historically been limited, for technical/technological reasons. The enormous variety of archiving and database formats has meant that content catalogues are specific to individual archives, and interoperability, both between archives themselves and between archives and the general public, has until now been unachievable.

The DISMARC project addresses the 'invisibility' of this important material by harvesting the data collections of participating institutes and individuals, mapping them to a DISMARC-created universal format and storing them in a 'metastore' for public access. DISMARC aims to create easy access for the general public, the scientific community, the media, the educational community and other audio businesses. Innovations introduced by the DISMARC project are already making themselves felt in the archive world, as eastern European archives become part of a wider European digital library community; established archives are taking note of the emerging DISMARC strategy and the opportunities presented for increased exposure and discovery of their content.

Contribution to programme objectives, to EC policies and Community added value

Benefits of the work include increased discovery and use of valuable European audio and audio-related content collections; simultaneous searching of distributed archives, with benefits for comparative study; the addition of a trans-European aspect to national cultural initiatives by aggregating content across European borders and allowing users to discover content in many locations simultaneously.

DISMARC addresses the European Digital Libraries i2010 Flagship Initiatives by allowing easier access to more content through the DISMARC browser and data collection; we also address 'Aggregation Across Borders' by contributing to stimulation of cross-border public sector information supply/re-use by helping to establish the same basic market conditions for players in the European information market. DISMARC's technical partner AIT will ensure that DISMARC solutions are compatible with varied systems of storage and access across the EU. By providing multiple-language versions of the browser which accesses multilingual core information, DISMARC will contribute to development of multilingual and multicultural access to digital content.

Contribution to economic development and social objectives

DISMARC contributes to the goal of improved discovery of culturally valuable material, leading to enhanced educational possibilities, including e-learning, life-long learning and long-distance learning, as well contributing to a general improvement of life-quality and improvement of cultural understanding by enhanced access to cultural information. Thus DISMARC contributes to the diffusion of knowledge and dissemination of information.

DISMARC also stimulates cross-border use of cultural / educational / scientific content by contributing to the establishment of similar basic market conditions for all players in the European information market, as referred to in Public Sector Information Directive 2003/98/EC. DISMARC contributes to a general pool of knowledge of / experience with digital rights management systems in the educational and scientific sectors, where the lack of effective DRM/IPR knowledge and management is most pronounced, relative to the commercial sector. DISMARC contributes to a wider recognition of the importance of public sector information, its commercial value and the associated societal implications of its use both in an educational /scientific context as well as in the general context of daily life. DISMARC will specifically contribute to the i2010:DIGITAL LIBRARIES initiative, especially to its objective of improving the "online accessibility" of cultural content. DISMARC's platform is fully compliant with and complementary to the TEL (The European Library) platform, and DISMARC content will be searchable via TEL. DISMARC will support audio business where relevant.

3 Consortium

The partners in this project are drawn from Europe's leading museums, archives, broadcasters, institutes and technology specialists. Their digital audio collections comprise some of the most extensive and specialised collections of music, other audio and audio-related objects in Europe.

Consortium members are:

Rundfunk Berlin Brandenburg - (RBB) Germany - coordinator/broadcaster / www.rbb-online.de

RBB is the regional public service broadcaster for the federal states of Berlin and Brandenburg. With bases in both Berlin and the media city of Potsdam-Babelsberg, RBB employs 1,700 staff, runs one regional television channel and seven radio stations, and provides a wide range of online and videotext services. RBB also holds a sound archive.

Berlin Phonogramm-Archiv (Ethnological Museum/Stiftung Preußischer Kulturbesitz) - (EMEM), Germany - archive/research / www.smb.spk-berlin.de/smb/de

Since its foundation the Berlin Phonogramm-Archive has played an important role in the field of ethnomusicological research and has gained international prestige. With exhibitions, concerts and the publication of books as well as sound recordings the Ethnomusicological Department has reached a wide audience outside the specialist ethnomusicological worlds and contributes to an increasing respect for the music of other people. The department also collects audiovisual recordings, musical instruments and other documents representing music. The department's main focus is on non-European music and traditional music.

Institute of Art of Polish Academy of Sciences - (ISPAN), Poland - archive/research www.ispan.pl/eng/instituteofart.htm

Instytut Sztuki Polskiej Akademii Nauk (Institute of Arts of the Polish Academy of Sciences) is a research institution concentrated on interdisciplinary studies on music, film, theatre and fine arts. The Institute's task is the research into and documentation of Polish art and artistic culture, fine arts and architecture, music, theatre and film, in their historic and modern manifestations, and viewed in European and world contexts.

Angewandte Informationstechnik Forschungsgesellschaft mbH - (AIT), Austria - technical www.ait.co.at/

Leading Austrian technical company: research work is done primarily in the field of applied information management (e.g.: distributed databases, collection management, knowledge engineering). This is carried out within the EC (e.g. IST, Ten-Telecom etc.) action lines or on national / regional level.

Grieg Musical Education - (GME) Norway - educational/technical www.griegmusic.com

Grieg Music Education as (GME) is a Norwegian company developing and distributing internet based education systems and education content. Customers are primary and secondary schools, colleges and universities. Together with the University of Stord/Haugesund, GME offers internet based education at university level. Grieg Music Education cooperates with the Fraunhofer Institute in Germany, developing specialised applications for internet-based music education.

Hochschule für Music und Theater Hannover - (HMTH), Germany - educational/archive www.hmt-hannover.de/

Leading German artistic / educational / research academy with approximately 1,100 students. The renowned Federal Music Academy and the Hannover Acting School were united in 1950 to become what is today the Hochschule für Musik und Theater Hannover (HMTH). Among the 20 subjects of HMTH are ethnomusicology and music pedagogics. The genre focus ranges from classical music to jazz and pop. HMTH offers masters degrees, doctorates and habilitation. (Affiliate institutes under the HMTH umbrella include the European Centre for Jewish Music, the World Music Study Centre, Institute for Journalistic and Communications Research and others).

WOMEX AG - (WOMEX), Germany - music network www.womex.com

Organiser of Europe's most successful 'world music' event, the annual WOMEX trade fair which includes concerts, seminars, TV and film broadcasts. Founded in 1994, WOMEX is the world's premier networking point for professionals exclusively dedicated to world, roots, folk, ethnic, traditional, local and diaspora music of all kinds. Eleven WOMEX editions in Berlin (1994), Brussels (1995), Marseilles (1997), Stockholm (1998), Berlin (1999, 2000), Rotterdam (2001), Essen (2002, 2004), Seville (2003) and Newcastle Gateshead (2005) have affirmed the value and potential of networking across borders, whether these be musical, political, cultural or commercial.

Yleisradio Oy - (YLE) <http://www.yle.fi>

YLE is the public broadcaster in Finland and is responsible for 33 radio and 5 TV channels. Radio broadcasts are made in Finnish, Swedish and Sami. The company is committed to introducing new broadband services in the near future. The YLE radio archives are currently involved in digitising the entire contents of the historical radio program archive, consisting of about 300,000 radio programs made since 1935. YLE radio archives have also participated in several projects with the national library and several public libraries, in order to make YLE's catalogue of commercial sound recordings available online.

Svenskt Visarkiv - (SVA) <http://www.visarkiv.se/en/index/htm>

SVA was founded in 1951, on the initiative of Ulf Peder Olrog, scholar and song-writer with an interest in old song traditions. Financial support was given by shipowner Sven Salén. Initially a foundation, the archive's main task was to index and systematize Swedish traditional song material. During the 1960's economic and personnel resources increased through state subsidies. In 1968 the SVA started to do field recordings. Since then the archive collects both vocal and instrumental folk music. In 1970 SVA became a state institution and in 1999 became part of the Swedish National Collections of Music.

School of Oriental and African Studies - (SOAS), UK - [educational/archive/research](http://www.soas.ac.uk/educational/archive/research)

The School of Oriental and African Studies (SOAS) is a college of the University of London and the only Higher Education institution in the UK specialising in the study of Asia, Africa and the Near and Middle East. The department has additional links of diverse kinds with the University of California, Kathmandu University, the National Conservatoire of Mali and the University of Rome (la Sapienza). SOAS also operate an in-house radio station (Open Air Radio www.openair.org.uk), creating a unique hybrid of research, academia and media under the same roof.

4 Project Results/Achievements

Data

The DISMARC metadata architecture relies on the transformation of data in local formats to a common data format. Such a transformation can be resolved electronically by a computer program. Therefore all data units of the source format need to be linked to the data units of the target format. Such a list of correspondences, often referred to as 'mapping' or 'crosswalk', relies on the knowledge of both formats. Its complexity varies with the complexity of the involved data structures. Such a program can be run periodically to transform the data automatically from the local archive to the DISMARC format.

The development and definition of this DISMARC format required substantial work that could be successfully accomplished in due time. EMEM was responsible for this part and carried out the work in close co-operation with all DISMARC archives and AIT, in consideration of different archives' needs and potential users' expectations. All relevant decisions, definitions, descriptions and inheritances were documented in the DISMARC Metadata Structures Manual. The manual was set up for different purposes: information and documentation of metadata work processes within the consortium, as a template for the prototype, as template for local database-to-DISMARC-crosswalks and as an instruction tool and guideline for aggregated archives in the future.

The developed DISMARC metadata architecture is based on the wide-spread and well-documented Dublin Core framework and seeks compliance with The European Digital Library (TEL/EDL). It reflects the heterogeneity of the content partners - from broadcasting archives to specialized

scientific archives - by allowing a relative high level of complexity in the metadata without forcing archives to use this complexity for all of their records.

In order to ensure not only a more detailed level of metadata description for archives but also specific functionalities in the DISMARC portal for users, like browse-ability and refined searches, the consortium agreed on the inclusion and development of simple and low-hierarchical controlled vocabularies and thesauri. A thesaurus working group had been set up to tackle this problem and to draw up DISMARC specific word lists. Next to common standards that can be used within DISMARC (e.g. TGN, LCSH, UDC, DDC, ET, DCMIType), vocabularies have been compiled and defined for the following thematic fields: contributor, geography, genre, performance context, performance setting, format, non-ethnic groupings, ethnic grouping and nationalities (IBSS cultures), era, historical geography and language. Substantial work on the compilation of those lists has been undertaken by EMEM, HMTH, SVA and ISPAN.

All results were fixed in the manual and eventually delivered to AIT for further implementation. At an early stage all content partners were asked to provide sample data of their local databases for initial mappings and testing routines. After the successful harvesting of a first quantity of partner data into the DISMARC metastore, further adjustments, revisions and specifications of their metadata mappings were constantly provided and have been subsequently implemented to the prototype.

More specific results and achievements by respective partners are mentioned in the following:

One side effect of WP4.2 Downloads & CDs was the updating of a large amount of material in the RBB archive with additional metadata that conformed to DISMARC requirements, enabling access to details (duration, composer name, track title etc.) that were previously unavailable to users. This activity was not part of the DISMARC project but resulted from the intention to make this information searchable via the DISMARC portal.

EMEM restructured its entire database by clearing fields and data entries while realizing data export facilities for DISMARC. They further plan to adopt DISMARC word lists in their local database to improve data quality and to make data management easier.

ISPAN's original metadata template was modified after analysis of other DISMARC partners' data, and after consultations with ISPAN archive's most frequent users, mostly MA and PhD students in musicology, as well as musicians of the Warsaw "Dance House" association. In particular, due to homogenous character of the collection, ISPAN database did not contain genre-description. To make the collection searchable together with other partners' data, genre identification was added. Other missing fields like Citation, Bibliography were added and the information will be completed in the future.

HMTH adopted partially the DISMARC metadata template for their local database and confirmed the underlying GBV/ PICA data structure to their needs. The possibility of cross-checking their own mapping results with other partners reinforced the positive effect.

GME has focused on the DISMARC data set that will be used in the educational tool. This data set will be necessary to make the search for audio files within the DISMARC project possible.

SVA plans the introduction of controlled vocabularies in their local system, which is stimulated by DISMARC.

YLE states positive impression due to the fact that the conversion to DISMARC has proven quite successful and their data is fully readable and searchable. While comparing with data of other DISMARC archives (EMEM, ISPAN, SVA) missing information could be detected. From their perspective the database is now ready for demonstration to potential partners and selected user groups.

Status of IPR and user service in DISMARC partner archives

ISPAN has prepared, circulated and summarized results of an IPR questionnaire. Detailed descriptions of the procedure and results have been fixed in a report.

As regards ISPAN collection, all the content is free in terms of authors' IPR, as it belongs to anonymous tradition. However, the IPR of folk performers must be respected. Therefore ISPAN has successfully negotiated with Stowarzyszenie Twórców Ludowych (Association of Folk Artists), the official collective representative of Polish folk performers, their consent to make selected recordings available for the DISMARC users in the internet. The permission allows 30 sec. per tune, which seems a satisfactory solution, regarding the fact that many recorded items are very short in original.

Copyright Legislation on Databases

Discussions about copyright protection and violation, theft and piracy have centred around musical content in recent years. While this is a matter of great importance to DISMARC with regard to the streaming, download and licensing of music on the portal, copyright legislation is relevant for DISMARC already at a more fundamental level: Databases are protected by laws, too, and in the process of creating the DISMARC platform, two kinds of databases have to be dealt with legally: On the one hand, the databases of the individual archives having joined or wanting to join DISMARC, on the other hand, the DISMARC database which is being created through the database content of the individual archives.

For DISMARC to operate on legally safe grounds in its dealings with databases, HMTB did extensive research on database legislation within the EC, cooperating with the The European Library/European Digital Library (TEL/EDL), the Rights Commission of the German Library Association (DBV) and The North German Library Association (GBV).

The EC already in 1996 undertook steps to harmonise copyright laws on databases within its member states. It adopted a legally unique approach by differentiating between creative and accumulative databases and devised an intellectual property right for the first and a *sui generis* right for the second. The "Directive on the legal protection of databases" (96/7/EC) was implemented in the legislation of the then member states by 2003. This means that database legislation within the EC is fairly standardized with some national variations, whereas database legislation outside the EC is considerably less consistent. This will become relevant, as soon as DISMARC includes archives from beyond the EC.

The findings of the survey on database legislation were put down in a report, which includes an overview of the general legal setting as well as information concerning the concrete DISMARC case. On the whole, the results of this research are not only relevant to the DISMARC project, but can be used as a legal reference work for other EC projects which aim at establishing networks and align databases from different member states and beyond.

IPR Management Tool

WOMEX created a flowchart to deal with copyright clearance for DISMARC content and different potential licensing cases. It was intended for the development of simple licensing solutions and template contracts and eventually should lead to adequate software solutions. Due to the high complexity of the copyright situation in many archives the flowchart stayed rough in its current form, however it has served as a basis for further investigations and work already taken up by the task partners.

RBB cooperated with YLE in preparation of the report 'Copyright for Archives'. The issues were discussed prior to YLE's work, and were subsequently jointly discussed and amended before the report was submitted to RBB's legal department. Further amendments resulted in the final document, for distribution to partner archives and, equally importantly, to joining archives.

It was discovered that contracts between rights owners and the host institution vary widely with all partners, and that in some cases there are variations even within partners' institutions, whereby different types of material is handled in various ways. This reality will be addressed in the next

phase of the project, either by incorporating various possibilities in the IPR management tool, by encouraging best practice or by a combination of both.

YLE has created a copyright guide, which summarises some of the IPR problems. It suggests that IPR issues may prove more difficult than originally thought, for instance archives often do not know whether their materials are in public domain or not. YLE has produced a memorandum on copyright in sound archives. There are also plans to provide a multilingual tool for DRM management. A draft of this exists, but further work on this awaits clarification of IPR issues which have arisen. The work will continue during the autumn.

Records on server

We anticipated that by month 12 (September 2007), there would be 50,000 records (i.e. individual items) in the database - in fact, as at August 23 2007, there were 224,000 items in the database, which represents a significant improvement on the original estimate. This is due to a more efficient upload process than predicted, and to greater availability of content than previously calculated. Partners were asked to send as much data as possible in order to make early performance testing and in order to have content selected for the CD project available in the metastore. Thus the data-gathering part of the project is well ahead of schedule.

Educational Tool

GME designed and developed an education tool within the DISMARC project. GME created a special designed portal for children and students which gives easy and direct access to music contained within the DISMARC archive collections.

The DISMARC education browser is now incorporated in the GME Virtual Learning Environment 'Music Delta'. GME also designed a new search tool / browser which searches DISMARC's audio files. The resource is developed as a world map, where kids and students can search for music using the criteria of country, genre, archive and time. The audio results are playable via a special designed mp3 player. The World Map is in development, and can be found at <http://www.musicdelta.com/map/dismap.html>.

Technical Requirements specification

AIT produced a comprehensive Requirements document which detailed all possible/envisaged use situations relevant to the DISMARC Portal and its underlying functionalities. This document was been circulated to partners for comment. Following the delivery of this document, a team of designers began the visual design and implementation of the Portal.

Multilinguality

Searches will be possible in 32 languages, and the underlying data can also be searched in these 32 languages. The actual data itself will not be translated from its original language, but using keywords and a translation table, keywords entered in one language will produce results across all the languages of the underlying data. In other words, if users search for 'drinking songs' in English, they will also find details of all other 'drinking songs' regardless of the language of the original data.

Approximately 3000 terms for translation have been selected and inserted into a special online translation tool. These terms are the initial input which will be the foundation of the multilingual vocabularies. The terms that have been chosen are first of all most of the controlled vocabularies, but also other relevant terms that are used in the individual archives' metadata. This means that even free text search will be possible, generating results in several languages at the same time. Also, a strategy for how to integrate the translations into the portal has been established. This includes technical as well as visual aspects. More than one scenario has been taken into consideration, since it has not been decided exactly how everything will work and look in the end.

RBB was responsible for finding suitable internal/external translators to carry out the multilingual tasks. Languages represented by partners were translated by those partners at no extra cost to the project. For other 'non-partner' languages, external translators were sought. RBB worked on the translation of the original 'scientific English' in the DISMARC dictionary into colloquial English.

Current position is as follows:

Language	Quantity of terms	Percentage translated to date (Sep 7 2007)
Bulgarian	2,767	99.93%
Chinese	2,766	99.89%
Croatian	227	8.20%
Czech	2,768	99.96%
English	2,756	99.53%
Estonian	828	29.90%
German	592	21.38%
Italian	2	0.07%
Japanese	514	18.56%
Korean	2,766	99.89%
Luxembourgish	2,763	99.78%
Polish	182	6.57%
Slovenian	2,768	99.96%
Spanish	183	6.61%
Swedish	946	34.16%

5 Target Users & their Needs

User groups were established and coordinated by RBB. The first user group consisted of partners only, and the group was asked to indicate the accuracy of the upload of their data to the DISMARC metastore. Partners reported general satisfaction with the presentation of their data via the Demo Portal, although various improvements were suggested regarding the underlying data structure in order to present a more comprehensive and comprehensible search result, and these suggestions were acted upon. It was noticed that, in the original datasets, various data elements were entered into 'general' fields, and partners were encouraged to rectify this matter, with the assistance of 'data-washing' technique developed by AIT. A second user group was asked to provide feedback about the Portal design, and this group was drawn from professional colleagues of partners and the general public. The partners were asked to respond to a variety of design possibilities and provided useful feedback. Another user group was established by partner YLE to determine the needs of broadcasters, and the conditions under which DISMARC would meet their requirements. Other user groups are planned to provide feedback on the Portal when the final design is available, drawn from the wider public.

Positive aspects of DISMARC from the users' perspective are identified as:

- no need to replace or change existing databases
- opportunity to present metadata online (where such a possibility does not preexist)
- standardisation of audio-related metadata on a European level, also functioning as guidance/support for those beginning data enrichment

Questions from users:

- how will DISMARC solve rights questions, especially on a pan-European level?
- will a compromise work?
- can metadata in a foreign language be useful?
- will this really be a free service, even after the EU-funded phase?

Overview of potential users identified:

USER	DEFINITION
General public	All with corresponding interests

Experts	Journalists
	Music specialists from media
	Scientific experts (ethnomusicology, comparative musicology, anthropology)
Producers	Audio-visual productions (TV, film, advertising)
	Audio productions (TV, film, advertising, record companies, broadcasters)
Institutions	Archives, NGOs, educational users, networks, libraries, research organisations
Students (music, musicology, ethnomusicology, anthropology, other)	Undergraduate students
	Master/performance students
	Research students
	Other disciplines
Scientific	Musicologists
	Other disciplines

6 Underlying Content

RBB:

The audio archive of RBB reflects the history of Germany and its capital Berlin - audio from the former GDR is hosted by RBB in the city of Potsdam/Berlin. The archive of former Berlin broadcasters SFB and ORB was transferred to RBB in 2003, when those organisations merged to form RBB. The seven RBB radio stations also hold archives for their various broadcasting needs. Further usage following the original broadcast of this material is normally not considered, despite these original recordings being considered 'audio treasures'. It is hoped that DISMARC will provide a major breakthrough on this point, as potential re-users from outside RBB will be able to find and request these recordings. RBB has also provided a number of thematic contributions to the Virtual CD, based upon material in the archive, as described in WP 4.2, including the infamous Fela Kuti 'Disaster Concert' from 1978. Currently there are 4,541 records from the RBB Music database (Musad) on the DISMARC server. The quality of the data is high since especially detailed information about contributors is available.

EMEM:

The EMEM archive contains recordings of music representing many global cultures (excluding western art and pop music) and is stored on various sound-carriers including Edison wax cylinders, analogue and digital tapes, 78 rpm shellac discs, LPs and CDs. "This is a world-wide unique sound collection with universal significance and intercultural dimensions" - UNESCO Memory of the World Register. The archive comprises the following audio collections: Berlin Phonogram-Archiv, Original Archive, and Published Media. EMEM has delivered 21,618 metadata records from all three collections to DISMARC. This was the total quantity of available metadata in May 2007. Due to permanent digitization at EMEM more data can be delivered in future. The quality of provided data corresponds to that of the original data though it is not as high as those of other archives like, e.g. YLE. Encouraged by DISMARC EMEM will pursue data enrichment in the near future.

ISPAN:

The Sound Archive of ISPAN in Warsaw is the oldest, the largest and the most comprehensive collection of folk music recordings in Poland, representing different regions, genres, vocal techniques, musical instruments and bands, etc. Although the Institute was officially founded in 1949, the oldest recordings date back to 1904 when Roman Zawilinski, a researcher from Cracow, made the first Polish recordings on wax cylinders, a number of which have been preserved in the ISPAN Archive.

ISPAN has delivered a quantity of 85,797 records, of which 3,692 are a sub-collection of recordings on decelith plates, made in 1946-1949 in Western Poland, which will be a subject of ISPAN's pilot (WP4.7). All the metadata represent Polish folk music, vocal and instrumental, from different regions of Poland, as well as examples of speech (mostly recorded interviews with folk artists and

wedding speeches). Apart from the decelith-recordings, the rest of the collection has been registered on reel-tapes. Recordings differ in length - from single stanzas till long ballads. Further, ISPAN has delivered a selection of 30 MIDI-files as experimental sound examples of archive's content free of IPR licence. Of this set, 10 examples have been delivered in double form (as mp3 and MIDI).

HMTH:

The centerpiece of the Wolfgang Laade collection consists of roughly 44.000 records (LPs, EPs, singles, shellac), the majority of which is of non-European provenance and production, reaching back as far as to the 1920s, and is not accessible and available in other European libraries and archives. In addition to its commercial recordings (records, tapes and CDs) the archive contains Wolfgang Laade's own field recordings from Morocco, Corsica, India, Sri Lanka, Australia and Oceania, covering the whole collection period. Concerning the data content, HMTH is preparing the delivery of data to DISMARC. HMTH has begun to catalogue the Collection Wolfgang Laade Music of Man Archive. By now, around 550 LPs have been indexed and 100 of them are already catalogued. As HMTH wants to assure the compatibility of its data with the German standard library system of the GBV, their data has been sent to the GBV for testing purposes. As soon as the GBV will have terminated the testing, the data will be delivered to DISMARC.

SVA:

SVA has contributed with more than 60,000 documents to the metastore. Much of the content consists of unique or rare recordings, of great interest to both researchers and the general public. This does not mean that SVA holds 60,000 recordings. Sometimes the different tracks on a recording are catalogued separately. SVA decided to contribute almost all of their audio/video related metadata. The data that is not included is mostly data about recordings that SVA doesn't have the rights to let just anyone listen to, for various reasons.

The exact amount of current data in the DISMARC metastore is 63,219. Due to different cataloguing levels (from object description to micro-description) the quality of SVA data varies, however some records are very detailed. One of the problems has been a lack of consistency: too many different people have been cataloguing over the years with different concepts. Another problem - the lack of personnel - leads to prioritizing certain records instead of cataloguing everything.

YLE:

So far YLE has delivered to DISMARC 48,604 datasets. They represent:

- 1) all documents which were in their TEHOSTE database in May 2007 (c. 16 000 sound effects). At YLE, the database grows daily, and the next update will include at least 2000 more documents. Eventually YLE's complete collection of sound effects (c. 60,000 documents) will be included, as soon as the cataloguing is completed, but this will take a few years.
- 2) all documents relating to YLE's own music productions in the FONO database at the same date, c. 32,000 documents. Not included are Fono documents relating to commercially published records in our record library (everything from Caruso to Roger Cicero). These are more than a million. All these documents relate to existing sound carriers (tapes, discs) AND equivalent digital sound files at YLE.

The small number of audio files delivered so far represent just a preliminary sample of YLE's collections. The basic criterion for inclusion was IPR: all the files are in public domain, or fully owned by YLE. Most of their sound recordings have a more complex rights history - the quality of catalogue data represent a full documentation of recorded works, performers, dates etc, but details of rights are not included. The data which has been converted for inclusion in DISMARC generally represents the same quality as original, although there are still some minor details to be cleaned up (words broken into two, etc).

SOAS:

The Department of Music at SOAS has an international reputation for research, teaching and performance, offering courses in, amongst others, African, North African, Middle Eastern, Chinese, Central Asian, Japanese, Korean and South Asian music. The majority of its audio material is held in SOAS Library's extensive collection, in a variety of formats including CDs, DVDs, VHS, cassettes and LPs. Currently DISMARC contains 1,222 records derived from SOAS library.

7 Summary of Activities

RBB maintained an overview of all partner activities, established a regular communications procedure between partners, was responsible for maintaining timely delivery of work, and established a quality management regime. RBB is also responsible for the design of the DISMARC portal, which is currently in progress, and managed the translations for the portal dictionary (which enable multilingual data searches) and portal languages (which enable multilingual portal operations). RBB cooperated with AIT and the metadata working group to establish metadata requirements. RBB also coordinated user group activities, provided specialist legal advice, worked with WOMEX in the development of the DVD guide to DISMARC, is in contact with a number of archives who are interested in joining the project, and participated in various dissemination activities, including production of multilingual material and attendance at/participation in conferences and seminars. One side effect of WP4.2 Downloads & CDs was the updating of a large amount of material in the RBB archive with additional metadata that conformed to DISMARC requirements, enabling access to details (duration, composer name, track title etc.) that were previously unavailable to users. This activity was not part of the DISMARC project but resulted from the need to make this information searchable via the DISMARC Portal. User group activities were developed to produce input into design and technical aspects of the project, and a potential-user profile was developed.

EMEM organized, conducted and elaborated the metadata work, established a thesaurus working group, and finally put all results down in a metadata structures manual. EMEM collected and sifted sample data of partner archives and created a first mapping table together with AIT. EMEM collected from each partner an initial amount of 100 proposed key search words for testing the translation tool. EMEM took part in the test phase of the multilingual work and was responsible for the German translation. EMEM held intensive discussions with all taskleaders of WP3 related tasks and maintained close contact to RBB concerning organizational issues and to AIT concerning technical issues. EMEM prepared successfully data export facilities at their database and selected suitable data for provision to DISMARC. A small amount of audio content has been prepared for diverse purposes: testing the prototype, provision for DVD and deliverance of traditunes for the portal. EMEM discussed with WOMEX WP4 topics, like university seminars, DVD and virtual CD concepts and delivered herefor relevant material and proposals. EMEM identified different user groups of their archive, established and fostered relationships to potential new archives and undertook various dissemination activities at national and international conferences and lectures in the museum, archive, and scientific world. EMEM established a cooperation with TEL/EDL on metadata work and has recently joined the EDLnet project as official Thematic Network Partner Expert in order to help building the European Digital Library.

ISPAN actively participated in the thesaurus working group, whereas they especially provided different proposals on the genre-vocabulary and worked out the ethnic list. They were involved in all translation activities concerning the Polish language and manufactured a questionnaire and report on copyright issues and work flows within the partner archives.

AIT gave feedback to the finalization of the DISMARC metadata scheme on object and collection level. AIT has provided a tool for the translation of the multilingual wordlists and the DISMARC controlled vocabularies, and was in close contact with SVA on the implementation issues for the multilingual search within DISMARC. AIT has provided the partners RBB and YLE with a template for the use case description for the IPR tool.

HMTH took part in the discussion on the DISMARC metadata template and in the creation and development of controlled vocabularies. The introduction of standardized cataloguing procedures is an important part of the work of HMTH in order to ensure the user friendliness of the DISMARC portal. HMTH has given impetus to introduce international standards from the social sciences to DISMARC. During the search for appropriate standards and the discussion of DISMARC categories with colleagues from the social sciences, they noted that potential new archives might have strong interests only in internationally standardized projects. HMTH participated in the creation of a multilingual DISMARC portal and fosters the dissemination of the project inside and outside the academic world in a global perspective. Concerning multilinguality, various potential partners and members of user groups have recommended to HMTH the inclusion of further languages for the

user portal, as there have particularly been mentioned Portuguese and Russian. Further HMTH examined the problem of database legislation within DISMARC and created a report.

SVA has worked closely with EMEM, ISPAN, AIT and HMTH in developing and completing both the Metadata Structures Manual and the Controlled Vocabularies/Encoding schemes. Also, SVA has worked with organizing translations and fed the translation tool/dictionary tool with almost 3000 terms. These terms will help users to find metadata in languages other than their own and consist of both terms from the mentioned Controlled Vocabularies and proposed words from the content providers. SVA supported translators by creating a Powerpoint and other material.

YLE participated at initial testing of the mapping routine and delivered almost 50,000 documents of both databases - FONON (own music recordings) and TEHOSTE (sound effects) - to the DISMARC metastore. They have worked at further specifications of their data mapping to make their data more compatible with materials from other archives. YLE has been responsible for the Finnish translations during the testing phase of the translation tool and in the subsequent work phase. YLE has created a copyright guide, which summarises some of the IPR problems, and produced a memorandum on copyright issues in sound archives. A draft for a multilingual tool for DRM management exists, but further work on this awaits clarification of IPR issues which have arisen.

List of events

2007

65

WOMEX 2007
Seville, Spain
October 25 - 28
RBB, ISPAN

64

DISMARC Board second meeting
Seville, Spain
October 23 - 24
All

63

ESEM : "Music Archives in the 21st Century"
Lisbon, Portugal
October
ISPAN

62

DISMARC to be presented at the GFM (Society for Music Research) annual conference
Cologne, Germany
September 27-29.
HTMH, EMEM

61

Austrian Society for Documentation and Information (ÖGDI)
Graz, Austria
September 17 - 18
AIT

60

Presentation of DISMARC to the Study Group on School Music (Arbeitskreis für Schulmusik) and the Association of German Music Teachers (Verband Deutscher Schulmusiker)
Stuttgart, Germany
30th September
HTMH

59
IASA-BAAC Conference
Riga, Latvia
September 15 - 20
YLE, SVA, ISPAN

58
Conference: New Technologies in the Humanities
Tallin, Estonia
September 3
ISPAN

57
RBB MultiKulti interview
Berlin, Germany
August 30
RBB

56
Nordic-Baltic Music Library meeting at Sibelius Academy
Helsinki, Finland.
Aug 30th
YLE

55
Polish Radio 2 interview
Warsaw, Poland
August 14
ISPAN

54
ICTM World Conference
Vienna, Austria
July
SVA, EMEM

53
Zenit e-Contentplus Workshop
Berlin, Germany
July 6
RBB

52
Presentation to Bayerischer Staatsbibliothek
Berlin, Germany
July 4
RBB

51
Attendance at European Science Foundation conference
Amsterdam, the Netherlands
June 20
ISPAN

50
Seminar Museums-Insel - 'Management of cultural Heritage Assets'
Berlin, Germany
June 16
EMEM

49

Attendance at Humanities in the European Research Area conference
The Hague, the Netherlands
June 11
ISPAN

48
Lecture in Millstatt
Carinthia, Austria
June 9
EMEM

47
DMB Conference
Frankfurt, Germany
June 6
EMEM

46
Lecture at Kulturwissenschaftliches Institute
Essen, Germany
June 6
EMEM

45
EBU workshop
Warsaw, Poland
May 11 - 12
RBB

44
Sound Archives: "Yesterday, Today and Tomorrow" conference
Warsaw, Poland
May 11 - 12
ISPAN

43
Lecture at Musicological Department of Karls-University
Prague, Czech republic
May 10
EMEM

42
BIMUC
Bergen, Norway
April 25 - 27
GME, RBB, WOMEX

41
Second DISMARC Board meeting
London, England
March 29 & 30
All

40
TAPE (Teaching Audiovisual Preservation in Europe)
Rome, Italy
March 8 - 10
EMEM

39
Presentation to SOAS Open Air Radio / students

London, England
February 27
SOAS

38
Presentation to EUDICO
Berlin, Germany
February 21 + 22
EMEM

37
Presentation to Wissenschaftskolleg zu Berlin - Ways of Knowledge
Berlin, Germany
February 17
EMEM

36
Presentation to Foundation Lower Saxonia
Germany
January 2007
HMTH

35
Presentation to Volkswagen Foundation
Hannover, Germany
January 31
HMTH

34
Presentation to presidential board of University of Hildesheim
Hildesheim, Germany
January 19
HMTH

33
Information about DISMARC was included in the European Seminar in Ethnomusicology bulletin
(ESEM point 41)
(Print) International
January 11
ISPAN

2007

32
Presentation to Archiv für die Musik Afrikas, Institut für Ethnologie und Afrikastudien, Johannes
Gutenberg-Universität Mainz
Mainz, Germany
December 13
HMTH

31
Presentation on DISMARC and OpenAir radio at the World Music and Dance Centre, during the
conference 'Cultural Diversity in Music Education'
Rotterdam, the Netherlands
December 16
SOAS

30
Presentation to Warsaw Learned Society, Institute of Musicology of Warsaw University
Warsaw, Poland
December 9

ISPAN

29

DISMARC report in E-News, online magazine of the Berlin Senat
Berlin, Germany
December 6
WOMEX

28

Presentation to PARMA (Polish Association of Research Managers and Administrators)
Warsaw, Poland
December 1
ISPAN

27

DISMARC outline published in Mitteilungen der Deutschen Gesellschaft für Völkerkunde e.V (DGV)

(Print) Germany
December issue
EMEM

26

Presentation to Stiftung Preussischer Kulture (SPK) Daten Redaktions-Konferenz
Berlin, Germany
November 30
EMEM

25

Presentation to European Digital Library conference
Vienna, Austria
November 27 - 28
RBB

24

Presentation to Meeting of Council of Museum of Folk Instruments
Szydlowiec, Poland
November 24
ISPAN

23

Presentation to Nyköping Conference of Swedish Folk Music Archives
Nyköping, Sweden
November 22 - 23
SVA

22

Presentation to Music Academy
Warsaw, Poland
November 20
ISPAN

21

Participation in Society of Ethnomusicology Annual Conference
Manoa, Hawai'i
November 16 - 19
EMEM

20

Participation in Electronic Imaging and the Visual Arts (EVA)
Berlin, Germany
November 8 - 10

E MEM

19

Presentation to Aga Khan Music Initiative in Central Asia
Cologne, Germany
November 11
HMTH

18

Presentation at Archivi di Etnomusicologia dell'Accademia Nazionale di Santa Cecilia
Rome, Italy
November 15
E MEM

17

Presentation to International Association of Sound Archives (IASA)
Wiesbaden, Germany
November 3 - 4
E MEM

16

Presentation to Ethnomusicological Department, Institute of Musicology, University of Cologne
Cologne, Germany
October 30
HMTH

15

Participation in Museumskonferenz
Berlin, Germany
October 30
E MEM

14

Presentation to European Broadcast Union (EBU) Workshop
Seville, Spain
October 25 - 29
RBB

13

Presentation to WOMEX
Seville, Spain
October 25 - 29
RBB/HTMH

12

Presentation to doctoral students at ISPAN,
Warsaw, Poland
October 23 & 30
ISPAN

11

Presentation to Norwegian Council for Folk Music and Dance, organised by Norwegian University of Science and Technology
Trondheim, Norway
October 20
ISPAN

10

Presentation during Polish Days
Trondheim, Norway
October 18-19

ISPAN

9

Participation in Study Group on Documentation of the German Museum Council
Berlin, Germany
October 16 - 18
EMEM

8

Presentation during Polish Days
Uppsala, Sweden
October 16 -17
ISPAN

7

DISMARC Kick-off meeting at the Deutsches Musikarchiv
Berlin, Germany
September 28 - 29
All

6

Participation in Third International symposium on Traditional Polyphony
Tbilisi, Georgia
September 29
EMEM

5

Presentation to Department of Music Collections at the Polish National Library
Warsaw, Poland
September 25
ISPAN

4

Presentation to the board of CHARM (Centre for the History and Archiving of Recorded Music)
London, UK
December 12
SOAS

3

Presentation to ESEM (European Seminar in Ethnomusicology),
Stockholm, Sweden
September 6 - 10
ISPAN, EMEM, SVA

2

Presentation to the organisational committee of the ODOK 07
Graz, Austria
September
AIT

1

Presentation of DISMARC to the Austrian BAM (Libraries, Archives, Museums) committee
AIT
2006

8 Impact & Sustainability

DISMARC will contribute to the Commission's i2010 European Digital Libraries initiative objective to improve online accessibility of digital content. For this purpose, the results of the project will remain accessible online for 5 years after the end of the duration of the project. A link to the DISMARC project will be provided on The European Library web site (or a follow-up web site).

Both the EU and national governments are committed to continued improvement of access to cultural, scientific and educational material. Therefore, in the context of this political environment and given the relatively low cost of maintaining/updating the DISMARC database, it will be important that the efforts of the DISMARC consortium continue to be supported by the participating institutes and broadcasters themselves subsequent to the EU-funded phase of the project. Partners' future plans for DISMARC are detailed below, and all partners state that they will continue to provide access to the DISMARC Portal for users of their local networks and will continue to operate the automated metadata updating procedure provided as part of the DISMARC technology. New archives will continue to be aggregated as contributors of content during the five years following the EU-funded period.

- RBB will host the server and assume all server-related costs following the funded phase; each participating institution will continue to support the DISMARC collection with continued regular metadata updates and continued access to the Browser. RBB commits itself to hosting the DISMARC infrastructure for at least five years after the EU-funded phase.
- The opportunities arising within the DISMARC project will be of the greatest importance for the ongoing development of GME's products. The service will be incorporated in the programs offered to educational institutions. It will enhance the access of information about music to such a degree that the DISMARC-tool could have the potential of becoming a permanent part of what GME can offer to its costumers. GME is committed to continue supporting and incorporating the DISMARC tools in their products for at least as long as the RBB is hosting the server.
- SOAS is committed to supporting the DISMARC project beyond the EU funded phase, and will provide support by providing access to the DISMARC Browser for users of the local network, continuing to operate the automated metadata update procedure and continuing to generate new digital audio content in-house through the OpenAir Radio project and the SOASIS label. Annual short term FM broadcasts through OpenAir will generate roughly 100 hours of audio content per year, with further programmes being broadcast over the Internet. SOASIS will continue to produce new albums with traditional artists from around the world, with a minimum of 3 releases (2 CD, 1 DVD) per year.
- WOMEX will inform all WOMEX delegates in continuous newsletters about the state and update of DISMARC; set up a webpage on the WOMEX-site to inform about DISMARC and to give contacts for the WOMEX delegates for possible contribution to DISMARC; research the possibility of including DISMARC database information in the advanced search of the WOMEX-site and virtual WOMEX; promote and publish the final launching of DISMARC in the WOMEX guide; give DISMARC members the opportunity to hold a press conference at WOMEX to disseminate information; stimulate DISMARC members and WOMEX delegates to hold a WOMEX conference about the aims, progress and further researches to DISMARC. An aim should be to set up an annual DISMARC meeting / conference at WOMEX of archives, research and science users; organize a submission to produce 4 virtual CDs partly using content of the DISMARC project. These CDs will also be distributed digitally worldwide on digital platforms.