D6.1.1 First Content Collection

ECP-2008-DILI-528001

EuropeanaConnect

D6.1.1 – First Content Collection

Deliverable number/name	*Deliverable D6.1.1 – First Content Collection*
Dissemination level	*PU*
Delivery date	*October 20 2010*
Status	
Author(s)	*Martin Gordon, RBB*

eContent*plus*

Österreichische
Nationalbibliothek

EuropeanaConnect is coordinated by the Austrian National Library

Distribution

Version	Date of sending	Addressee	Role in project
0.1	20.10.2010	ÖNB	Coordinator
0.2	21.10.2010	David Haskiya, Daniel Teruggi	EF Liaison Manager, INA
0.3	27.10.2010	ÖNB	Coordiantor
0.4	03.11.2010	Max Kaiser	Coordinator
1.0	14.01.2011	EU, Liferay, web page	

Approval

Version	Date of approval	Name	Role in project
0.2	21.10.2010	David Haskiya	EF Liaison Manager
0.2	25.10.2010	Daniel Teruggi	INA, expert
0.4	14.01.2011	Max Kaiser	Coordinator

Revisions

Version	Status	Author	Date	Changes
0.1	Draft	Martin Gordon	21.10.2010	Initial version
0.3	Draft	Martin Gordon	27.10.2010	Comments by David Haskiya Daniel Teruggi incorporated
0.4	Draft	VPZ	03.11.2010	Minor layout changes
1.0	Final	VPZ	14.01.2011	Minor changes

Executive Summary

This Deliverable (EuropeanaConnect Deliverable D6.1.1 – First Content Collection) delivers over 100,000 digital music audio objects (plus a small number of additional related items including video, images and text) to Europeana for immediate access.

Table of Contents

1. Introduction

During the first 18 months of the project, WP6 made considerable efforts to aggregate new content providers, as specified in the Description of Work. At the time of writing, 69 content providers (see appendix 1) have submitted their audio data to the EuropeanaConnect Audio Aggregation Platform (AAP) http://www.dismarc.org/ for subsequent harvesting by Europeana.

Content owners, including sound archives, broadcasters, institutions such as museums and libraries, record companies, etc.) were sent the regular AAP Newsletter, which was followed up by individual invitations to deliver content to Europeana. In almost all cases, the recipients were aware of the DISMARC/EuropeanaConnect project. Discussions led, in most cases, to a fruitful conclusion, with delivery of content.

In a minority of instances content owners preferred to deliver their material either themselves or via another content aggregator. A very small number of content owners were still at the beginning of the process of creating a digital archive, and will be followed up when appropriate.

The current total of digital items which have been harvested from the AAP and which are now available via the Europeana platform is 110,730 (data retrieved 11/10/2010).

Statistical details of content delivered by the AAP can be found at the following Europeana URL: http://www.europeana.eu/portal/briefdoc.html?query=europeana_collectionName:%2008901* .

2. Details of content

2.1 Formats:

As noted above, the current total of digital items which have been harvested from the AAP and which are now available via the Europeana platform is **110,730** (data retrieved 11/10/2010).

As well as **101,228** music audio items, there are an additional **6,692** text items, **2,798** images and **12** videos. These additional items are all related to accompanying music items; they include album covers, manuscripts, demonstrations of folk dances, etc.

2.2 IPR:

Content delivered to Europeana is either in the public domain (approximately 20%) or is protected by various IPR restrictions. Where the material is in the public domain, it will (from the end of October 2010) be marked by the Public Domain Mark http://creativecommons.org/publicdomain/mark/1.0/. Where IPR restrictions are applicable to the content, content owners have agreed to deliver extracts from their original audio files, usually at reduced quality (mp3 196 kbs). Material which is IPR-protected will be flagged by the relevant CC license (from the end of October 2010). In this manner, contemporary audio which is protected by IPR is nonetheless available via Europeana. Users have the option to follow links to access full-length, full-bandwidth versions. Extracts usually consist of between

30 seconds and 1 minute 30 seconds of audio, and the duration of the extracts is visible in the Europeana search results.

2.3 Hosting:

Audio extracts are either hosted by the content provider or, where content providers have no audio server, on the DISMARC/EuropeanaConnect Audio Aggregation Platform itself, at no cost to the content owner.

2.4 Europeana Content Provider Agreement:

After discussion with the Europeana office, it became clear that all audio aggregated under the EuropeanaConnect Description Of Work is covered by the DoW. We are advised that it is therefore unnecessary for content providers who are providing their content via EuropeanaConnect to sign an agreement with Europeana while the project is still running.

3. Workflow

Following an archive/content provider's agreement to deliver content to the AAP for harvesting by Europeana, a test sample of data is sent to the WP6 technical partner AIT – Applied Information Technology ltd. AIT creates a mapping in order to conform the data to the ESE schema ESEv3.2XMLSchema and to identify any data which may be missing.

When a test mapping has been created, the content owner is invited to give his/her response to data presented via the AAP test server.

Following the content owner's approval, the data are then uploaded and checked on the Europeana Content checker EuropeanaContentChecker.

Following a successful completion of this process, Europeana is notified that an AAP harvesting is requested.

Finally the content owners and the AAP agree on an update procedure, and the content owners are provided with tools to enable future automatic uploads/updates of their content (using FTP/OAI-PMH).

4. Future

During the remaining year of the EuropeanaConnect project, aggregation will continue. We anticipate that there will be no difficulty in reaching the projected figure of 200,000 music audio items by the conclusion of the project.

Appendix 1 – List of content providers

Content providers who have delivered music audio to Europeana via the DISMARC/EuropeanaConnect Audio Aggregation Platform as at 12/10/2010 are as follows:

Short name	Name	Type
AIMP	The Archives Internationales de Musique Populaire	Archive
AVX	Alia Vox	Record label
AYL	Ayler Records	Record label
AZM	Azzurra Music	Record label
BHM	Boheme	Record label
BLA	Bela Records	Record label
BPL	Bergen Public Library	Library
BRR	Brambus Records	Record label
CHR	Charly Records	Record label
CJS	Complete Jazz Series	Record label
CLA	Claves Records	Record label
CLB	Calibrated	Record label
CLM	Columna Música	Record label
CRG	Copyright Group	Record label
CVM	Cavendish Music	Record label
CYG	Cygnus	Record label
DBM	Doublemoon Records	Record label
DCR	Documents	Record label
DPR	Delphian Records	Record label
ELK	Elkar	Record label
EMEM	Ethnomusicology, Media-Technology and Berlin Phonogramm-Archiv	Archive
ETMC	Estonian Traditional Music Center	Archive
F2R	Fuel 2000	Record label
FMR	FM Records	Record label
FNM	Fania	Record label
FNML	Future Noise Music Ltd.	Record label
FVR	Fantastic Voyage	Record label
GGN	Gega New	Record label
GHT	Gesellschaft für Historische Tonträger	Archive
GLN	Gael Linn	Record label
GLR	Green Linnet	Record label
GMB	Gambit Records	Record label
GMC	Global Music Centre	Record label
GNI	Glasbenonarodopisni inštitut	Archive
GOT	Gotlandstoner	Archive
ISPAN	Instytut Sztuki PAN	Archive
LLTI	LLTI tautosakos archyvas	Archive
MMR	Music Mecca	Record label
MSS	Musikstrasse	Record label

MTD	Music Today	Record label
MYT	Myto Records S.a.s. / Qualiton Imports, Ltd.	Record label
NCA	Nuova Canaria	Record label
ORC	The Orchard	Distributor
ORM	Oriente Musik	Record label
OSU	Osutaunu	Record label
OVC	Ovação	Record label
PCP	Picap	Record label
PKP	Park Records	Record label
PLV	Plenum Vox	Record label
PMC	Plaza Mayor Company, Ltd.	Record label
PNM	Pneuma	Record label
PRE	Preiser Records	Record label
RGS	RGS Music	Record label
RKOU	RKO/Unique	Record label
SGR	Signum Records	Record label
SPP	Signum Records	Record label
SRC	Simplyred.com	Record label
SRG	Saregama	Record label
SSR	Silver Streak Records Limited (UK)	Record label
SYD	Seydoni (Record Label)	Record label
TCM	TCB - The Montreux Jazz Label™	Record label
TER	Talking Elephant Records	Record label
VDE	VDE-GALLO	Record label
WED	The EuropeanaConnect Audio Exhibition	Virtual Exhibition
WHL	Wigmore Hall Live	Record label
WOMEX	WOMEX	Organisation
XGL	Xenophile/Green Linnet	Record label
YLE	YLE Radio archives	Broadcaster
ZTI	The Hungarian Academy of Sciences Institute for Musicology	Archive