Project no. 033104

MultiMatch

Technology-enhanced Learning and Access to Cultural Heritage
Instrument: Specific Targeted Research Project
FP6-2005-IST-5

Deliverable 8.3.2
Organization of workshops

Start Date of Project: 01 May 2006
Duration: 30 Months

Organisation Name of Lead Contractor for this Deliverable: Alinari

Version: Final

_____

1 Document Information

Deliverable number:	D8.3.2
Deliverable title:	Organisation of workshops
Due date of deliverable:	October 2008
Actual date of deliverable:	November 2008
Author(s):	Francesca Tavanti, Sam H. Minelli, Giovanna Naldi, Ilaria Del Secco, Andrea de Polo (Alinari)
Participant(s):	ISTI-CNR, BandG, BVMC, WIND, Alinari
Workpackage:	WP8
Workpackage title:	Dissemination and Exploitation Planning
Workpackage leader:	Alinari
Dissemination Level:	PU (Public)
Version:	Final
Keywords:	workshop, user group, dissemination, exploitation

History of Versions

Version	Date	Status	Author (Partner)	Description/Approval Level
0.1	25th November 2008	first draft	Francesca Tavanti (Alinari), Sam H. Minelli (Alinari), Giovanna Naldi (Alinari), Martha Larson (UVA): contribution for SSCS and IACH workshop	Table of contents
v.prefinal	03rd December 2008	Pre-final	Francesca Tavanti	Formal revisions
final	04th December 2008	Final	Pasquale Savino	Formal revisions

Abstract

This document updates the previous deliverable D.8.3.1, reporting MultiMatch workshops held during 2008. This deliverable describes the features and reached results of the workshops organised to promote and exploit the MultiMatch project results towards the target user groups and scientific communities. The Workshops represented the occasion to gain feedback on the MultiMatch prototypes and establish contacts for clustering with other European projects and exploitation opportunities.

Table of Contents

2   Executive Summary

Since last deliverable [1] (reporting MultiMatch workshops held during the first project period) additional workshops have been organised by the project Consortium.

This document describes the workshops organised by MultiMatch Consortium during 2008 with the scope of promoting and exploiting the project results. This document was scheduled to be produced at month 30 of the project in coincidence with the end of the project and after the final MultiMatch workshop.

In the following sections we are going to detail features, goals, results and activities employed for the organisation of SSCS workshop held at Singapore (section 4), IACH workshop held at ECDL (section 5) and the final MultiMatch workshop held at VSMM (section 6). These workshops aimed to present the final project results to the target audience gaining feedback on the MultiMatch prototypes and creating contacts for the exploitation of the project and clustering activities.

3   Introduction

In the previous deliverable we reported the first four MultiMatch workshops organized by the project Consortium, which took place in the occasion of DELOS 2007-Italy, EVA/MINERVA 2007-Jerusalem , Online Educa Berlin 2007-Berlin, 2nd DELOS 2007-Italy.

Our strategy in organizing the Workshops is influenced by the need to address different target communities represented by:

- The MultiMatch user group made up of educational users, professional users (cultural heritage), general users (tourism)
- The Scientific community
- The European digital Library community.

SSCS Workshop was a sectional event focused on content retrieval which represented the occasion to reach the scientific community in that field.

To target the Digital Library Community and Cultural heritage research community a suitable opportunity was represented by ECDL 2008 held in September 2008 in Aarhus, Denmark. MultiMatch Consortium decided to organize and sponsor a workshop entitled "Information Access to Cultural Heritage" (IACH) at ECDL.

The final MultiMatch workshop was supposed to be organised after the release of the second prototype to coincide with the end of the project. The MultiMatch consortium monitored all the events taking place at month 30 of the project (October 2008). A list of possible events was drafted: Festival Creatività[1]- Italy, Workshop Information Retrieval 2008 (WIR 2008)[2]-Germany, ICT event[3]-Lyon, ECEL[4]-Cyprus, EVA BERLIN[5]- Berlin, ACM 2008[6]-Canada, Creactivity[7]-Italy, VSMM[8]-

---

[1] http://www.festivaldellacreativita.it/index.jsp?language=it

[2] http://lwa08.informatik.uni-wuerzburg.de/Wiki.jsp?page=LWA08

[3] http://ec.europa.eu/information_society/events/ict/2008/index_en.htm

[4] http://www.academic-conferences.org/ecel/ecel2008/ecel08-home.htm

[5] http://www.gfai.de/pinboard/eva/e_index.html

[6] http://www.mcrlab.uottawa.ca/acmmm2008/

Cyprus. Some of them were planned to be held in November 2008, others addressed sectional topics, other were limited to a national circuit. Finally, after several considerations, the Consortium decision fell on VSMM event: this was the ideal context to attract a large target audience (CH users, Multimedia digital Library, scientific community) in an international Cultural heritage environment.

All the workshops were organised by partners among the MultiMatch Consortium. The workshops were addressing MultiMatch topics, some of them were organised within a sectional context, others were expressly dedicated to MultiMatch workshop. We believe that all the mentioned workshops are relevant to the MultiMatch scopes as they represented the occasion to disseminate the project results and set up fruitful discussion with different user groups.

4 SSCS 2008 Workshop, Singapore

The second workshop on *Searching Spontaneous Conversational Speech*[9] (SSCS 2008) was held in conjunction with SIGIR 2008 on 24 July in Singapore.

SIGIR Conference[10]
The 31st Annual International ACM SIGIR Conference was held from 20-24 July 2008 in Singapore. SIGIR is the major international forum for the presentation of new research results and for the demonstration of new systems and techniques in the broad field of information retrieval (IR).

Workshop objectives
The goal of the workshop was to bring the speech community and the information retrieval community together. The forum was designed to be conducive to the close interaction and the intense discussion necessary to promote fusion of these fields into a single discipline with a concerted vision of spoken content retrieval.

Target Audience
The workshop target audience was represented by the scientific community in the field of IR

SSCS Workshop
The workshop was not directly addressing MultiMatch, but thematically was closely linked to the project. Both the audio and the video collections that were the subject of investigation in MultiMatch contained spontaneous conversational speech. In particular, interviews and studio discussions that occur in the podcasts (the audio collection) and the BandG video collection are examples of cases in which issues involved with searching spontaneous conversational speech needed to be addressed.

---

[7] http://www.progettocreactivity.com/

[8] http://www.vsmm2008.org/

[9] http://ilps.science.uva.nl/SSCS2008/

[10] http://www.sigir2008.org/

One of the SSCS 2008 organizers was from a MultiMatch consortium member (Larson from UvA). Additionally, MultiMatch members submitted two papers: a full paper and a poster paper; both were accepted for the workshop.

Workshop outcomes

At the workshop, talks and posters were presented covering a wide range of topics including vocabulary independent search, spoken term detection, combination of models/indexes, use of speech recognition lattices for search, segmentation, temporal analysis, benchmarking, exploitation of prosody, speech surrogates for user interfaces and multi-language collections.

Demonstrations of speech-based retrieval systems from a variety of application domains introduced a strong practical emphasis into the workshop program.

The workshop concluded with a panel discussion, whose goal was to identify future research directions for speech retrieval: a MultiMatch member (Jones from DCU) accepted an invitation to serve on the panel.

A point that emerged in the panel discussion that is particularly relevant for the CH domain is that users and not researchers are the one that gives rise to the research problems that are addressed by the area of searching spontaneous conversational speech. The curators of multimedia collections must be motivated to take an active part in the design of speech-based multimedia access systems. Furthermore, user enrichment of spoken content plays an important role in system training and development. Participation of content owners and of content users is particularly critical in cases in which multimedia collections are small or specific to specialized needs or interests. The challenge of involving content owners in the system design process is counter-balanced by the trend that users have a difficult time imagining new technology before they see it being demonstrated. Part of the challenge of searching spontaneous speech is to make the technology accessible enough to users so that users start generating their own ideas about how the technology should be used.

Workshop Material
- The full proceedings of the workshop are available at
  http://ilps.science.uva.nl/SSCS2008/Proceedings/sscs08_proceedings.pdf
- All of the oral presentations made at the workshop were recorded with a video camera and processed by the UvA team. The videos were indexed using text and audio indexing techniques developed at UvA within the MultiMatch project. The indexing approach combines features derived from the title and the abstract of the presentation with text appearing in the slides and also speech recognition transcripts of the spoken content of the talk.
- The talks recorded at SSCS 2008 are accessible at http://ilps.science.uva.nl/SSCS2008 and can also be searched using the UvA search engine at http://zookma.science.uva.nl/islatv/

A detailed report on the SSCS 2008 workshop[11] was submitted for publication in the SIGIR Forum.

---

[11] http://ilps.science.uva.nl/SSCS2008/Proceedings/SSCS2008_report.pdf

5 IACH Workshop: ECDL, Aarhus, Denmark

"*Information access to Cultural heritage*"[12](IACH) workshop was held 18 September 2008 in Aarhus, Denmark. It was organized in conjunction with the 12th European Conference On Research And Advanced Technology For Digital Libraries (ECDL 2008)[13].

ECDL conference

ECDL2008, the 12th conference in the series of European Digital Library conferences, was held in from September 14 to September 19, 2008.

ECDL has become the major European conference on digital libraries and associated technical, practical, and social issues, bringing together researchers, developers, content providers, and users in the field.

The content of ECDL 2008 focused on research and development supporting information access and exploration from a technology perspective as well as in different application domains such as science, e-government, cultural heritage.

IACH Workshop: target group

The workshop represented the ideal context to reach both the Digital Library and Cultural Heritage research communities to disseminating and exploiting the main project outcomes.

Workshop objectives

The goal of IACH 2008 was to bring together practitioners from the field of cultural heritage and researchers in the area of information access.The workshop's aim was to promote exchange of ideas concerning creation, storage, retrieval and use of cultural heritage information. In particular, the workshop focused on the question of how keepers of cultural heritage resources can improve information access by applying information technology and principles currently being developed in the area of information access. Following current trends, users are actively engaging with cultural heritage on the web, prompting heritage institutions to stimulate communities to grow around them. The workshop placed an additional emphasis on active user participation, which creates new modes of

---

[12] http://ilps.science.uva.nl/IACH2008/schedule/schedule.html

[13] http://www.ecdl2008.org/

access and contextualisation. These are challenging developments for cultural heritage organisations, due both to the variety and richness of cultural heritage resources and to the diverse range of users with varying information needs and search behaviours.

Organization of the Workshop

The IACH 2008 organizing committee was composed of a member from UvA (Larson), BandG (Oomen), UNED (Cigarran), plus Kate Fernie, a consultant from the UK, who shared the vision and volunteered her time to serve on the organizing committee. USFD and CNR played a critical role in writing the workshop proposal.

Contributions for the workshop were solicited via a call for papers and were reviewed by a programming committee of experts. The workshop organizers drew on expertise both within and beyond the MultiMatch consortium in order to compose the programming committee. From the submissions, 5 were accepted as full papers and were presented as oral presentations and 5 were accepted as posters. The scope of the program was further extended by a keynote, two invited speakers, and two invited posters.

Workshop outcomes

The workshop was attended by experts from all over the world.

The workshop was characterized by lively discussion and energetic participation by the attendees. The poster session was an especially dynamic part of the workshop, with demonstrations being presented in parallel with the posters. The MultiMatch system was one of the demonstrations presented. The workshop concluded with a panel, which generated a productive discussion. The panellists were Marco de Niet (Digital Heritage Netherlands), Joyce Ray (Institute of Museum and Library Services, USA), Richard Wright (BBC Archives, UK) and Maarten de Rijke (University of Amsterdam) (see Appendix A). Issues addressed were technologies, open standards, information overload, user generated content, open source software and services, authoritativeness and quality and rights management.

Workshop material

- Details on the workshop can be found in the workshop report to be published: Larson, M., Fernie, K., Oomen, J. and Cigarran, J. "Information Access to Cultural Heritage Workshop Report" D-Lib Magazine Vol 14. No. 11/12

- A more detailed version of the workshop report was also submitted to SIGIR Forum.

- The proceedings, pictures, videos, keynote and oral presentations of the workshop are available on-line at the workshop website [14]. The videos were indexed using indexing techniques developed at UvA within MultiMatch, which combines features derived from the title and the abstract of the talk with text appearing in the speaker slides and also speech recognition transcripts of the spoken content of the talk.

- The talks recorded at IACH 2008 can be searched using the UvA search engine at http://zookma.science.uva.nl/islatv/

- Currently, the workshop organizers are in discussion regarding a call for a special issue of a journal based on the IACH 2008 workshop.

.

---

[14] http://ilps.science.uva.nl/IACH2008/

6   Final MultiMatch Workshop: VSMM, Limassol, Cyprus

The Final MultiMatch workshop: *Multilingual and Multimedia Access to Cultural Heritage Digital Objects: the MultiMatch solution*[15] was held 23 October 2008 in conjunction with VSMM 2008[16] (Limmasol, Cyprus).

VSMM 2008
VSMM 2008 is the conference on Visual Systems and MultiMedia Dedicated to Digital Heritage.
This event, organized in Cyprus from 20 to 26 October 2008, aims to exchange and share know-how in the areas of Cultural Heritage (CH) and Information & Communication Technology (ICT). In particular it focuses on Multimedia Technologies, Computer Graphics and digital documentation/preservation in Cultural Heritage (Multilingua, Multimedia Digital Library).

Workshop objectives

In accordance with the project timetable this final MultiMatch workshop took place at month 30 of the project. The main workshop's goal was to present the final results of the MultiMatch project emphasizing the solutions adopted to address specific user needs, involving the target user groups and all contacts collected during the project lifetime.
The workshop aimed to:
- discuss the approach followed to support interoperability between the different sources of Cultural Heritage information, and the use of the MultiMatch system in different applicative settings
- demonstrate the final version of the MultiMatch system prototype and show the functionality provided to support multilingual search and browsing of CH objects and multimedia retrieval.
- discuss future trends in the field of multilingual and multimedia access to cultural heritage digital objects.
- Establish clustering activities with other European projects

---

[15] http://www.cs.ucy.ac.cy/conferences/vsmm2008/downloads/MultiMatch-Workshop-Agenda.pdf

[16] http://www.vsmm2008.org/

Workshop dissemination and target group

The workshop was disseminated through several channels:

- Alinari mailing list (from its CH network)
- Advertising of the event in Alinari corporate website,
- Advertising of the event in Alinari educational website
- European Commission mailing list (DG Information Society and Media Unit E3: Cultural Heritage and Technology Enhanced Learning)
- Project partners' contacts

The invited participants were members of the MultiMatch user groups involved during the project lifetime, members of the Advisory Board, Project reviewers, other representatives from CH or educational institutions and Research Communities.

Poster exhibition

At the event, emphasis has been given to the presentation of short papers at the poster exhibition, with direct interaction between authors and audience.
Two MultiMatch posters have been shown during the exhibition: the posters provided an overview of the project work and results empathizing with visual diagrams and pictures the Multimatch system.

Workshop outcomes

The workshop represented the occasion to:

- Disseminate of the project results (1st prototype and 2ndprtototype) to the target audience
- Gain feedback from the participants
- Involve the participants in the MultiMatch user group
- Involve the participants in the evaluation activity of PT2
- Establish clustering with European projects

The workshop chairman was Pasquale Savino (ISTI-CNR), Coordinator of the MultiMatch project. The first part of the workshop was dedicated to the MultiMatch project presentation : an introduction by Pasquale Savino, a presentation on user requirements by Sam H. Minelli(ALINARI), a presentation on Metadata by Johan Oomen (BandG). The second part was a presentation of the MultiMatch system and a Demo session of the second MultiMatch prototype. Participants were involved in the evaluation of PT2: after the demo presentation a questionnaire collected users feedbacks and some of the participants gave their availability to participate in the MultiMatch field trials.
The third part was dedicated to MultiMatch users: Rafael Carrasco (UA) presented users feedback on the first MultiMatch prototype; afterwards a panel session involved invited speakers: representatives from IMAGINATION (Antonella Fresa), EUROPENANA Local (Rob Davis) and MICHAEL (Kate Fernie). The speakers presented a quick overview of the concerned projects; a brief description on how multiligualism had been taken into account in the project, and which technological solutions had been applied; a reflection on how the projects could benefit from MultiMatch results and possible future interaction envisaged. This session represented an occasion the start hypothesis of clustering activities and investigate joint exploitation opportunities

Organization of the Workshop

This workshop was organized by Alinari with the support of University of Alicante, ISTI-CNR, BandG.

Workshop material

- <u>Speakers' presentations</u> are available at MultiMatch website[17]

- VSMM conference has found a very good worldwide acceptance in the scientific area as well as in the area of academia and publishers of the results of research and good praxis. For that reason the international accredited UK publisher Multi-Science Publishing Co Ltd has offered to publish, in a <u>special book edition</u>, all the papers from the parallel VSMM2008 workshops. The book will be published at the end of this year and not later than January 2009. Some of the speakers of the MultiMatch workshop are working to provide their papers for the special book edition.

7   Conclusions

The organisation of the MultiMatch workshops has been in line with the expected timetable of the project. By month 24 of the project (second milestone) four MultiMatch workshop have been held[18]. By month 30 (third milestone) other three workshops organised by the MultiMatch Consortium have been taken place.

The workshops, which addressed different target communities (educational/CH/general user group, European digital library community, scientific community), represented the occasion to disseminate and exploit the final project results, collect feedback on the MultiMatch final prototype, create collaborations and clustering with related projects, and investigate exploitation opportunities.

---

[17] http://project.alinari.it/diss-publish/mm_vsmm.php

[18] see D.8.1.3

8 References

[1] MultiMatch Deliverable 8.3.1 Organisation of Workshops

9 Appendix A: Workshop programs

9.1 SSCS Program

Opening Session 8.30 -10.00
Welcome

- **Keynote Address** Query-by-Example Spoken Document Retrieval *H. Li*
- **Oral Presentation** Cluster-based Model Fusion for Spontaneous Speech Retrieval *M. Alzghool, D. Inkpen*
- **Oral Presentation** Using Term Clouds to Represent Segment-Level Semantic Content of Podcasts *M. Fuller, M. Tsagkias, E. Newman, J. Besser, M. Larson, G. Jones, M. de Rijke*

Tea Break 10.00 - 10.30

Morning Session 10.30 -12.30

- **Oral Presentation** Combination of Multiple Speech Transcription Methods for Vocabulary Independent Search *J. Mamou, Y. Mass, B. Ramabhadran, B. Sznajder*
- **Oral Presentation** Fast Approximate Spoken Term Detection from Sequence of Phonemes *J. Pinto, I. Szoke, S.R.M. Prasana, H. Hermansky*
- **Oral Presentation** Towards Large Scale Vocabulary Independent Spoken Term Detection: Advances in the Fraunhofer IAIS Audiomining System *D. Schneider, J. Schon, S. Eickeler*
- **Oral Presentation** Hybrid Word-Subword Decoding for Spoken Term Detection *I. Szoke, M. Fapso, L. Burget, J. Cernocky*

Lunch Break 12.30 - 13.30

Afternoon Session 13.30 -15.30

- **Invited Speaker** Speech-based methods in the video search mix
  T. Davis
- **Oral Presentation** Audio stream segmentation and classification based on Jump Function Kolmogorov analysis in wavelet domain *H.D. Tran, H. Li*
- **Oral Presentation** Word-Lattice Based Spoken-Document Indexing with Standard Text Indexers *R. Peng Yu, K. Thambiratnam, F. Seide*
- **Poster Boasters**

Tea Break 15.30 - 16.00

Posters and Demonstrations 16.00 -17.00

- **Poster** From Audio Content Analysis to Conversational Speech Detection and Characterizationv *B. Bigot, I. Ferrane*
- **Poster** Enhanced Search and Navigation on Conversational Speech *F. Cailliau, A. Giraudel*
- **Poster** Classification of Dual Language Audio-Visual Content: Introduction to the VideoCLEF 2008 Pilot Benchmark Evaluation Task *M. Larson, E. Newman, G. Jones*
- **Poster** A Korean Spoken Document Retrieval System for Lecture Search *D. Lee, G. Geunbae Lee*
- **Poster** Using Prosodic Features to Prioritize Voice Messages *T. Polzehl, F. Metze*
- **Poster** Subword-based Indexing for a Minimal False Positive Rate *L. van der Werff, W. Heeren*

Panel on future research directions 17.00-18.00

9.2   IACH Workshop Program

9:00 Welcome and workshop opening
9:10 Keynote Chris Batt [19]"Building Bridges to Knowledge"

Session I: Digital Cultural Heritage: European initiatives
- 9:45 Invited Speaker: Stefan Gradmann (Humboldt-Universität zu Berlin) Europeana Infrastructure
- 10:15 Invited Speaker: Rob Davies (MDR Partners) EuropeanaLocal—Its Role in Improving Access to Europe's Cultural Heritage through the Europeana service of the 'European Digital Library'

10:45 Coffee Break
Session II: Cultural heritage exhibitions and collections
- 11:15 Enabling Semantic Access to Cultural Heritage: A Case Study of Tate Online. *P. Clough, J. Marlow, N. Ireson (University of Sheffield)*
- 11:45 Enigma Helvetia: Promoting an Exhibition through Multiple Channels. *N. Di Blas (Politecnico di Milano), E. Rubegni (University of Lugano), P. Paolini (Politecnico di Milano)*
- 12:15 Having Triplets - Holding Cultural Data as RDF. *K. Byrne (University of Edinburgh)*

12:45: Lunch/Demos

14:00 Poster Boasters

Session III: Access to cultural heritage multimedia
- 14:30 Access to the Music of Man Archive: The MARIMBRO System. *A. Schneidt (Logica), K. Klenke (Hanover University of Music and Drama) T. Mandl (University of Hildesheim), J. Medinvil, R. Vogels (Hanover University of Music and Drama), C. Womser-Hacker (University of Hildesheim)*
- 15:00 Towards Affordable Disclosure of Spoken Word Archives. *R. Ordelman, W. Heeren, M. Huijbregts, D. Hiemstra, F. de Jong (University of Twente)*

---

[19] http://ilps.science.uva.nl/IACH2008/schedule/batt/batt.html

15:30 Coffee Break

16:00
Session IV: Posters
- An Intelligent Service Orineted Approach for Improving Information Access in Cultural Heritage. *A. Baruzzo, P. Casoto, P. Challapalli, A. Dattolo (University of Udine)*
- Providing Linked-up Access to Cultural Heritage Data. *T. Blanke, M. Hedges (King's College London)*
- Multimedia Retrieval in MultiMatch: The Impact of Speech Transcript Errors on Search Behaviour. *J. Carmichael, P. Clough (University of Sheffield), G. Jones, E. Newman (Dublin City University)*
- A System Architecture for Conveying Historical Knowledge to Museum Visitors. *D. Dannells (University of Gothenburg)*
- Feasibility of a European Registry of Standards for Digital Heritage. *M. de Niet (DEN: Digitaal Erfgoed Nederland)*
- Marking Up Cultural Materials for Time and Geography. *F. Gey, R. Shaw, R. Larson, M. Buckland, B. Pateman, D. Melia (University of California Berkeley)*
- Searching for Linguistic Phenomena in Literary Digital Libraries. *F. Sánchez-Martinez, M. L. Forcada, R. C. Carrasco (Universitat d'Alacant)*

17:00 Panel of Future Directions

18:00 Closing

9.3   MultiMatch final Workshop Program

-**Chairman Pasquale Savino**-
First Part: MultiMatch presentation
- 15.15-15.30 *Registration*
- 15.30-15. 50 *Introduction* - Pasquale Savino, (ISTI-CNR, Italy)
- 15.50-16.15 *User Requirements in the Cultural Heritage Domain*- Sam Habibi Minelli (Alinari 24 Ore S.p.A, Italy)
- 16.15-16.45 *Metadata in the Cultural Heritage Domain*- Johan Oomen, (Netherlands Institute for Sound and Vision, The Netherlands)

Second part: MultiMatch DEMO
- *Searching Cultural Heritage Information with MultiMatch:*
  - » 16.45-17.10 *Presentation of the MultiMatch system* – Franca Debole (ISTI-CNR, Italy)
  - » 17.10-17.30 Franca Debole (ISTI-CNR, Italy)

17.30-17.45 Coffee break
Third part: MultiMatch users

- 17.45- 18.10 U*ser feedback*: Rafael Carrasco (Biblioteca Virtual Miguel de Cervantes, Spain)
- 18.10 -19.20 *Invited speakers* Invited speakers: Antonella Fresa (IMAGINATION), Kate Fernie (MICHAEL), Rob Davis (Europeana and EuropeanaLocal).
- 19.20-19.30 *Conclusion of the workshop*- Pasquale Savino (ISTI-CNR, Italy)