Deliverable 4.5

Report of the 3rd CHORUS Conference

Deliverable Type *	:	PU
Nature of Deliverable **	:	R
Version	:	Released
Created	:	June 22nd
Contributing Work packages	:	WP 4
Editor	:	Jussi Karlgren**,** Markus Kauber
Contributors/Author(s)	:	Joost Geurts, Nozha Boujemaa, Nicu Sebe, Robert Ortgies, Christoph Dosch, Ramon Compañó, Joachim Koehler
File		

* *Deliverable type: PU = Public, RE = Restricted to a group of the specified Consortium, PP = Restricted to other program participants (including Commission Services), CO= Confidential, only for members of the CHORUS Consortium (including the Commission Services)*
** *Nature of Deliverable: P= Prototype, R= Report, S= Specification, T= Tool, O = Other.*
Version: Preliminary, Draft 1, Draft 2,…, Released

Abstract:
The third and last CHORUS conference on MMSE took place from the 26th to the 27th of May 2009 in Brussels, Belgium. About 100 participants from 15 European countries, the US, Japan and Australia learned about the latest developments in the domain. An exhibition of 13 stands presented 16 research projects currently ongoing around the world.

Keyword List: Multi-media access, Multi-media retrieval, Cross-media retrieval, Mobile informatics, Geopositioning, Microsearch, Language disconnect, Search vocabulary, Semantic analysis, Temporal retrieval, User-contributed data, Retrieval architecture, Personalisation, Evaluation

The **CHORUS Project Consortium** groups the following Organizations:

JCP-Consult	JCP	F
Institut National de Recherche en Informatique et Automatique	INRIA	F
Institut fûr Rundfunktechnik GmbH	IRT GmbH	D
Swedish Institute of Computer Science AB	SICS	SE
Joint Research Centre	JRC	B
Universiteit van Amsterdam	UVA	NL
Centre for Research and Technology - Hellas	CERTH	GR
Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V.	FHG/IAIS	D
Thomson R&D France	THO	F
France Telecom	FT	F
Circom Regional	CR	B
Exalead S. A.	Exalead	F
Fast Search & Transfer ASA	FAST	NO
Philips Electronics Nederland B.V.	PHILIPS	NL

Information Society
Technologies

Contents

1. INTRODUCTION TO THE 3$^{\text{RD}}$ CHORUS CONFERENCE

1.1 Venue

The third and final conference of the CHORUS project took place on May 26 and 27, 2009 in the Bedford Hotel & Congress Centre in the heart of Brussels, Belgium, under the chairmanship of Drs. Nozha Boujemaa, Christoph Dosch, and Ramón Compañó.

This conference, in contrast with the previous CHORUS conferences, was not colocated with another event but entirely independent: the location was chosen by the consortium after visiting several alternative venues. Brussels and the Bedford Hotel turned out to fit the conference requirements quite well: Brussels is conveniently located for the participants arriving from all over Europe and the world, and the conference room provided enough space to allow people to move freely and an additional connected hall hosted the demonstrations and exhibition stands, registration and coffee.

1.2 Participants

The conference was attended by 100 people from all over Europe (15 countries), Japan, Australia and the USA. The range of professional background covered private and public research institutes (various universities, INRIA, IFAAR, IDIAP, CERTH, Fraunhofer, ...), industry (Nokia, Yahoo, Exalead, Thomson, ...), public bodies (EC, German Ministry of Economy) and national initiatives such as THESEUS and QUAERO.

In addition, the conference was accompanied and animated by a stand exhibition of 11 research projects participating in the CHORUS cluster and two external research projects (all in the domain of MMSE), namely the Grand Voyage project from Japan and the BRONCHOVID, GAMA, INDECT and COTENT projects from Poland (one stand). The complete list of projects can be found at the end of this report or on the Conference Web site at "http://www.ist-chorus.org/conference_downloads.asp", where the presentations by the speakers also can be found. The exhibition is described in more detail in D0.3 "Rising Public Participation and Awareness".

1.3 Conference agenda

Conference Chairs (Day 1 and Day 2):
Nozha Boujemaa (INRIA); Christoph Dosch (IRT); Ramon Compañó (JRC-IPTS)

08.45: Arrival of Participants: Registration and Coffee

09.15 – 10.00: Morning Session 1: Opening of the Conference – Jean-Charles Point (JCP-Consult)

- 09.15: Opening of the Conference and statements by EC representative "Multimedia Search in the perspective of Future Internet Services" - Joao Da Silva (EC)

- 09.35: "CHORUS in the landscape of European effort in MMSE" – Luis R. Rosello (EC)

- 09.50: Introduction of Conference Objectives by Nozha Boujemaa (France)

10.00: Coffee Break & Poster and Demo Session

10.30 – 12.30: Morning Session 2: CHORUS Roadmap and International Perspectives – Nozha Boujemaa (INRIA)

- 10.30: Prospective in MMSE Research - Prof. Ramesh Jain (USA)

- 11.00: Overview of Information Grand Voyage project - Multimedia Search and Interface for Next Generation - Prof.Miki Haseyama (Japan)

- 11.20: THESEUS – Advanced Services for MMSE – Regine Gernert (Germany)

- 11.40: A review of first preliminary results from Quaero - Pieter Van Der Linden – (France)

- 12.00: CHORUS Roadmap and recommendations- Nozha Boujemaa – (France)

12.30: Lunch

13.30 – 14.50: Afternoon Session 1: CHORUS Vision and Industrial Trends – Christoph Dosch (IRT)

- 13.30: Industrial trends for Internet search - Ricardo Baeza-Yates (Yahoo! Research)

- 13.50: "Enterprise search trends and challenges" – Gregory Grefenstette (Exalead)

- 14.10: MOVIMOS - a scalable, distributed multimedia search engine for mobile applications - Thomas Breuel (DFKI)

- 14.30: "CHORUS Vision: Outcome of the Think-Tank" – Henri Gouraud (INRIA), Pieter Van Der Linden (THOMSON)

14.50 – 16.10: Afternoon Session 2: Highlights from MMSE Research Projects – Joachim Koehler (Fraunhofer)

- 14.50: Search engine and result optimization: Internet Score Cards - Christoph Glauser (IFAAR)

- 15.10: CHORUS: Status and Challenges of MMSE Technology – Joachim Köhler (Fraunhofer IAIS)

- 15.30: "3D Scene Structure Analysis for Semantic Annotation and Retrieval of Unedited Video" Oliver Schreer (FH HHI)

- 15.50: ERC project SECO - Stefano Ceri (DEI, Politecnico di Milano)

16.10: Coffee Break & Poster and Demo Session

- 16.30 – 17.30: Panel closing session: Disruptive Technologies and Services in the near Future – Nozha Boujemaa (INRIA)/ Christoph Dosch (IRT)

- 17.30: Closure of Day1 – Loretta Annania (European Commission)

Day 2

09.00 – 10.50: Morning Session 1: FP6 and FP7 Project Presentations – Nicu Sebe (UvA)

- 09.00: FP7 project PETAMEDIA - Alan Hanjalic (TU Delft)

- 09.15: FP6 project VITALAS - Arjen P. de Vries (CWI)

- 09.30: FP6 project PHAROS - Francesco Nucci (Engineering)

- 09.45: FP7 project LIVINGKNOWLEDGE - Fausto Giunchiglia (Univ. Trento)

- 10.00: The EU AMI and AMIDA projects: Recognition and Understanding of Meetings and Lectures - Hervé Bourlard (IDIAP)

- 10.15: Towards Web-scale content search: the SAPIR approach - Fausto Rabitti (CNR) and Pavel Zezula (Masaryk University)

- 10.30: Round table discussions and questions

10.50: Coffee Break & Poster and Demo Session

11.10 – 12.30: Morning Session 2: Challenges beyond Technology - Ramon Compañó (JRC-IPTS)

- 11.10: CHORUS Findings - Ramón Compañó (JRC-IPTS)

- 11.20: The Future of Web Search - Hendrik Speck (Univ. Kaiserslautern)

- 11.40: Search for Mobile 2.0 - Ajit Jaokar (OpenGardens & Futuretext)

- 12.00: Round-table discussion of Options - Ramón Compañó (JRC-IPTS)

12.30: Lunch

14.00 – 15.50: Afternoon session: "Use Cases - Bridging Benchmarking and Validation" – Jussi Karlgren (SICS)

- 14.00: Introduction, and brief overview of CHORUS findings and recommendation regarding user-centric SE design – Jussi Karlgren (SICS)

- 14.10: CHORUS studies outcome on use-case typology dimensions - Paul King (ITI CERTH)

- 14:30: VideoCLEF: evaluation of moving image retrieval" - Martha Larson (Univ. Amsterdam) and Gareth Jones (DCU)

- 14.50: User scenarios and user requirements from media professionals - Xavier Vives Surroca (Catalunya TV)

- 15.10: ImageCLEF LS-VCDT: Evaluation of multilabel image annotation incorporating domain knowledge and concept subjectivity - Stefanie Nowak (Fraunhofer)

- 15.30: Round-table discussion and questions

15.50: Closure: Loretta Anania (European Commission)

16.00 Coffee Break & End of Conference

2.   CONFERENCE SESSIONS AND PRESENTATIONS

2.1 Session 1 - Opening of the Conference and presentation from the Commission

This session was dedicated to the organization of the conference, and views expressed by the commission. It started by the opening of the conference by Jean-Charles Point (JCP-Consult, CHORUS project coordinator). Jean-Charles introduced the conference in the more global CHORUS perspective, and reminded the different outcome and deliverables of the project. Then he introduced the conference agenda and the general objectives.
Dr. Nozha Boujemaa (Scientific Coordinator) presented the different sessions, each structured around CHORUS findings in the different areas of MMSE. A session is dedicated on the presentation of FP6/FP7/National projects results.

Dr. Joao Da Silva (Director of the Network and Communication in Directorate of DG-INFSO), presented search in the perspective of the Future Internet. He first presented current trends: computing power and data amount becomes increasingly high, the servers are being concentrated now in the hands of a small number of search players like Microsoft, Google, Yahoo, He then reminded main challenges in multimedia like information representation, extraction (both in real and non real time) and summarizing. He illustrated by some examples the increasing size of the Internet, especially with the number of video and pictures introduced each day via Flickr; You Tube, etc.
The Internet is dominated by 10 languages, and there is a large unbalance in the usage of Internet between continents to the profit of Europe, America and Asia. Less than 100 sites reach more than 90% of the audience. In summary, the Web is currently mostly dominated by a few US based organizations. Future trends will be the rapid increase of video traffic, and mobile use of the Internet.
He presented then the vision of a future internet where information finding will play a vital role. He reminded the Commission's ambition, especially to improve R&D - industry collaboration and facilitate Trans-European collaboration. In addition to the classical tools associated to the Work program, Public-Private partnerships were presented as a complementary tool. As a conclusion he invited the MM search community to define the positioning of MM search as a key component in the Future Internet and to be active in the Future Internet Actions.

Dr. Luis Rodriguez Rosello (Head of EC Networked Media Unit) presented the context and landscape of EU efforts on MMSE. After a brief introduction of FP7 areas, and general current challenges of Networked Media, he presented a vision towards user generated content and services by comparing current gaps and possible future trends. These trends emphasize the problem to archive, categorize, and search the enormous amount of data generated, and transform them into personal / community media services. Dr Rosello recalled the objectives related to MMSE in the 2099-2010 Work Program, and detailed the context in which the different Integrated projects and STREPs interact and are coordinated by CHORUS. He detailed the progress made so far, like in 3D and P2P AV search engine, use cases, adaptive search by content / context, methods for information retrieval of diverse content types. He concluded that MM search will be a key enabler for the Future Internet, and that academia/industry collaboration should lead to real impacts in the domain.

2.2 Session 2: CHORUS Roadmap and International Perspectives

The CHORUS roadmap presents the CHORUS vision to synthesize and focus research within Europe in order to bridge the gap between ongoing research efforts and socio-economic needs

and novel services. To this end, CHORUS identified six prototypical areas that are used to characterize relevant dimensions regarding technical and socio-economic challenges. Notably Web search, Personalised TV, Enterprise Content search (variety, business), Library search (value, old), Personal Content Search (variety, local) and Monitoring, Detection & Alert (flux vs base). After conducting gap analysis studies, CHROUS identified several directions that deserve European effort towards more efficient search engines in order to reveal the implicit knowledge and makes it reachable in fair and attractive ways to the user.

From cross-disciplinary viewpoints, CHORUS recommends to:

1. Achieve more efficient indexing techniques for multimedia content enrichment and automatic meta-data creation. Socially-enriched automated indexing will empower the robustness of the indexing techniques.

2. Develop new multimedia search paradigms based on content/context/event, to go beyond current retrieval systems that are merely keyword-based or query-by-example-based. Event structures are expected to be the main driver for media contextualization.

3. Model efficiently both the implicit and explicit feedback to improve personalization and recommendation abilities of a search engine (including collaborative tags filtering, user preference detection ...).

4. Develop more informative user interfaces (UI) for future applications (too little overlap between networked media technology providers and UI designers today): toward smart visualization of media delivery and enhanced user quality of experience (QoE).

5. Break complexity and afford scalability: besides the amount of input data and generated features, complexity need to be managed for other growing quantities such as the number of users, the number of information sources and the number of data attributes / features dimension

6. Develop interoperable meta-data standards: open ended content, association with object and preservation through its life,

7. Make available and develop open multimedia corpora which is a key enabler for MSE scientific and commercial success,

8. Address privacy concerns to afford guarantees to the users: minimum regulation is necessary for consumer protection, privacy protection or unfair competition. Current EU regulation does not cover adequately or are not applicable to search engines,

9. Address security, integrity and trust issues related to search and networked storage (International cooperation needed): will foster user participation on a bidirectional media scene while preserving the trust models afforded by editorial material

10. Support Pan-European privacy certification of IT products or IT-based services compliance with European data protection regulation

The coordination effort through CHORUS allows positioning various EU efforts among the technological landscape dimensions. The mapping reveals that Europe is rich with very sharp expertise in many separated topics in the field of search engines. It was pointed out that innovative commercial services are feasible today using exiting pieces of research results for some niche markets (mainly for business search market). Regarding the consumer search market, Europe is lacking today an integrated program that gathers all needed expertise for building competitive real life search engines.

CHORUS recommends to:

1. Empower aggregation and orchestration of such expertise into an operational end-to-end search organizational structure (idea such creating a "Center for New Generation European Search Engine" needs to be investigated).

2. Foster user-centric design (market-pull) requirements for EU funded projects against the technology-driven design (techno-push).

2.3 Session 3: CHORUS Vision and Industrial Trends

The goal of the session was to point out the CHORUS Vision in relation to the industrial trends regarding audio visual search. The CHORUS Vision was created by the CHORUS think tank members and presented by Thomson and JCP. The industrial trend were presented by Yahoo focusing on Web search, from Exalead focusing on enterprise search and from DFKI & U. Kaiserslautern presenting a concrete implementation of a multimedia search engine with the name MOVIMOS.

The mission of Yahoo! Research is to develop the world-class science that will deliver the next generation of businesses to their company. They are interested to establish an open culture of collaboration with peers from academic and research institutions. The vice president of Yahoo Research for Europe, Middle East, Africa and Latin America, Ricardo Baeza-Yates, presented "A possible future for the future Web search". Search is no longer only about document retrieval. Future Internet services will be focused on user needs in whole. Services will have Web-mediated goals. Users basically like to get tasks solved by a service, instead of wasting time by formulating a query and choosing the right source from a search result list. Therefore services will support the user by synthesising information generated by different content sources for solving tasks.

For this, services will need to handle user generated content coming from social networks and real-time applications like Twitter in large quantities and different qualities. Quality may increase due to quantity by advancing the wisdom of crowds by using correlations: for the content itself, for tagging, ranking and building "folksonomies" (method of collaboratively creating and managing tags). But three impacts need to be considered for the content: the fragmentation of content ownership, the fragmentation of content access and the fragmentation of rights for accessing the content.

The new breed of search experiences of Yahoo is focused on the ecosystem (e.g. Searchmonkey), on the usage (e.g. SearchPad), on the tagging (e.g. Faceted Image Search) and on the content (e.g. Correlator).

Enterprise search represents a huge market opportunity for search technology according to Gregory Grefenstette, Chief Science Officer from Exalead. He presented the "Enterprise search trends and challenges". Enterprise search is not just a reduced Web search. Enterprise users are identified, meaning access rights have to be managed, but also that user profiles can be built. Enterprise search can tailor its interface to the user group, implement group specific semantics. Input is much more varied than Web pages. Trends are to consolidating information from different sources, including structured sources, providing interfaces and data treatment tailored to the enterprise data models. Enterprise search can advantageously replace many database functions, providing wider access to company data. The principal challenges are extracting wider varieties of typed semantic information from raw and structured data.

Thomas M. Breuel from DFKI & U. Kaiserslautern presented the group, the research and the results from the project MOVIMOS with its primary target on mobile search by using mobile phones with embedded cameras for retrieval. For this, the mobile phone captures an image and sends it to a multimedia search engine. MOVIMOS is the driver for many more applications: mobile augmented reality, image/video search engines, digital forensics, automated pornography filtering, book and OCR search. MOVIMOS looked on already existing retrieval system and setup technical goals like full multimedia support, easy extensibility, dynamic database updates, scalability, tagging and categorization, semi-supervised learning, context dependent search and personalized search. They succeeded to build a new, flexible and distributed platform by taking into consideration the state-of-the-art technologies and using open standards. They will release the open source code for this research platform in autumn 2009.

Henri Gouraud from JCP Consult and Pieter van der Linden from Thomson R&D presented the "Outcome of the Think Tank" as published in the "CHORUS Vision Document" D3.4 (distributed during the conference). As part of the Think Tank activity, a functional description

of Search Engines has been developed, resulting in a unified and shared vision. This functional description highlights the central role played by meta data both in executing the search itself (the Match component), but also in preparing the query and presenting the results. This second aspect stresses the central role played by the user whose efficiency should be augmented (and not replaced) by the system.

2.4 Session 4: Highlights from MMSE Research Projects

The session "Highlights from MMSE Research Projects" covered a variety of technologies and solutions in this area. In the first presentation from Christoph Glauser from IFAAR (Institute for Applied Argumentation Research) with the title "Search engine and result optimization: Internet Scores "a methodology to analyse and optimize the matching between queries and the content of a Web site. The service offered by IFAAR analyses the content clusters of search results and visualizes the analytics in a graphical score cards. This presentation shows the importance to consider the content and indexing with the search and navigation part of a search engine. As defined in the generic search engine architecture defined in the CHORUS project the optimization of indexing and search functionalities is a key issued by designing advanced search engines. This is the case for text based search engines as well as for MMSE

In the second talk, Joachim Köhler (Fraunhofer IAIS, Germany) presented the main results from WP 2 of the CHORUS project. It is a key issue to improve the indexing methods for the different kinds of media types (speech, music, image, video, 3D). Here the exploitation of large annotated data sets (e.g. Flickr) will provide a promising direction to improve the quality of automatically generated metadata. The second key issue addresses the scalability of the indexing and search technologies. It is expected that the amount of multimedia content in the Internet will grow exponentially. Therefore powerful algorithms and tools must be applied to enable an automatic indexing and to create the required metadata.

A more detailed description of video and 3D indexing was given by Oliver Schreer (Fraunhofer HHI, Germany). Here a 3D video processing framework was presented which extracts low level information (2D feature points) from a video. The low level information is used to extract medium level information (3D camera information) by using a-priori knowledge about 3D objects. The technology developed in the EU-IST project RUSHES is a representative example how indexing technology moves from low level to mid level analysis.

In the final talk given by Stefano Ceri from the Politecnico Di Milano the project "Search Computing - www.search-computing.org " was presented. The main objectives of the project are to investigate semantic technologies to allow natural language based questions and to develop an intelligent search architecture based on Web services. The results of this project will be also an important contribution for MMSE to provide more semantic based search capabilities.

All four presentations gave an impression of the diversification of the MMSE landscape in Europe. It will be an important political and strategic issue to synchronize all these research efforts in this field

2.5 Panel closing session: Disruptive Technologies and Services in the near Future

At the end of the day speakers were invited to participate in a panel to discuss about disruptive technologies and services in the near future. Panelists agreed that search will be a pervasive and ever present function fully integrated into applications and often not explicitly visible to the user.

2.6 Session 5: FP6 and FP7 Project Presentations

The goal of the session was to give a glimpse into the existing FP6 and FP7 projects by allowing representative projects to present their envisioned outcomes, the use cases considered, scenarios,

functionalities, target audience, etc. The selected projects addressed explicitly the large scale media search problem and they were selected with the intention of having a good coverage of the different types of projects (IPs, e.g., VITALAS, PHAROS, AMI/AMIDA; STREP, e.g., SAPIR; NoE, e.g., PETAMEDIA; FET-IP, e.g., LivingKnowledge).

Alan Hanjalic from TU-Delft, presented the PETAMEDIA project and insisted on the role of this project to bring together the research of four national networks in the Netherlands, Switzerland, UK and Germany in the area of multimedia content analysis (MCA) and social peer-to-peer networks and their efforts to eventually to establish a European virtual centre of excellence. The collective research of the project is directed towards the synergetic combination of user-based collaborative tagging, peer-to-peer networks and multimedia content analysis, and towards identification and exploration of potentials and limitations of combined tagging, MCA and SP2P concepts. Addressing some of these aspects, Arjan de Vries from CWI, presented the project VITALAS. The project is relying on three core activities which include cross-media indexing and retrieval methods, interactivity and context adapting, and search scalability issues. The functionalities of their system is specified and validated by major European multimedia content providers, in conjunction with the academic and the industrial partners of the project.

An integrated search platform paradigm has been presented by Francesco Nucci from the project PHAROS. This platform is meant to be built on an innovative, open, and distributed architecture that enables consumers, businesses and organizations to unlock the values found in audiovisual content. The PHAROS search platform aims at creating a new infrastructure for managing and enabling access to information sources of all types, supporting advanced audiovisual processing, content handling, and management that will enhance control, creation, and sharing of multimedia for all users in the value chain. An important issue is the scalability problem and the need for processing huge amount of images.

These aspects have been addressed by Pavel Zezula in his presentation of SAPIR project. This project aims at developing a large-scale, distributed P2P architecture that makes it possible to search audio-visual content using the query-by-example paradigm. The vision of the project is to conduct innovative research that will lead to a technology where end-users are peers that can produce audio-visual content from their mobile devices.

Two different perspectives on search were presented by Fausto Giunchiglia (LivingKnowledge project) and Herve Bourlard (AMI/AMIDA projects). The vision inspiring LivingKnowledge is to consider diversity an asset and to make it traceable, understandable and exploitable, with the goal to improve navigation and search in very large multimodal datasets (e.g., the Web itself). To achieve these challenging goals the project employs interdisciplinary competences from, e.g., philosophy of science, cognitive science, library science and semiotics. The proposed solution is based on the foundational notions of context and its ability to localize meaning, and the notion of facet, as from library science, and its ability to organize knowledge as a set of interoperable components (i.e., facets). The AMI and AMIDA projects address search from a different perspective: search in automatic constructed repositories that summarize meetings and lectures. AMI technologies can be integrated into existing solutions, or as software meeting assistants running in parallel to other meeting technologies. A meeting assistant is any system that supports its user to get more value out of meetings. It enables a user to navigate an archive of meetings, viewing and accessing the full multimodal content, based on automatic annotation, structuring and indexing of the information streams. For example, navigation may be enabled using automatic annotations such as speech transcription and identification of participants. This presents a more focus view of the search problem and uses heavily the available context information that is present in a meeting room (e.g., limited number of people, controlled conditions, etc.).

Overall, the session stirred extensive discussions and questions from the audience. The conclusion is that large scale search covers a wide variety of applications and perspectives.

2.7 Session 6: Challenges beyond Technology

The purpose of this session was to discuss a couple of important non-technological trends and challenges. Given the breath of the domain only three topics were selected namely (i) the challenges arising from the current techno-economic model of Web search, (ii) the potential and prospect of mobile search and (iii) the role of search for privacy. Afterwards the views of the three presenters were discussed with the conference audience, with the view also to derive lines of activities for different stakeholders.

Hendrik Speck from the University Kaiserslautern gave an overview of Issues arising indicating that the current market situation is a quasi monopolistic one. Google is the dominant player in Europe; holding (close to and over) 90% of the search engine advertising market in the most relevant EU Member States. He pinpointed to the potential from the current techno-economic model of Web search. He provided figures adverse effects and risks of having a single player holding such a dominant position.

The prevailing search engine business model still relies (almost exclusively) on advertising. Non-advertising business options, like subscription-fee service, enterprise search, pay for inclusion, consulting, etc. a very minor role. One potential exception in the mid-term might evolve around technology licensing. In 2008 licensing accounted to close to three per cent of the business, up from 1-2% in the previous years. Although this is a net increase, it is too early to judge if this increase is the consequence of a clear trend in business models.

With regard to advertising, Google follows two major routes, either through their own Web sites or through their network sites. While in 2003 revenues where (nearly) equally split, by 2008 advertising in Google Web sites accounted already for double the revenues of the Google network sites. The effect that most of the advertising revenues are generated on the search engines own (or affiliated) sites is positive for Google, as the margins are higher as those where the benefits are shared.

One major reason that Google was able shift the source of advertising, is an increased stickiness of the audience to Google sites by providing a diversification of products. Google has followed a deliberate strategic objective to offer a range of attractive tools and services. Advanced search tools, for Web search, news, products, blogs or movies, have been complemented by personalization features (iGoogle, Toolbars, RSS reader), browsers (Chrome), mapping services (for different platforms including mobiles, sms, and offering local features), etc. The objective is to assure the business model by creating fidelity of an audience. Speck pinpointed to the risk that this dominance may result in a high entry barrier for competitors. Another concern is that is that Google has the opportunity to collect massive data about people (profiles through search histories, interests, habits, shopping patterns, geographical information, etc) and that there may be the risk of misuse.

Ajit Jaokar presented the trends and opportunities with respect to mobile search in which he gave this view the market dynamics and their trends to discuss the drivers, barriers and disruptive factors for mobile search.

As for the market dynamics, Jaokar presented some market indicators driving mobile search, by emphasising the importance of meta data enabled content, the increasing quantity of information indexed, the likely emergence of social, real-time, the growing share of user generated content (the rise of appstores and long tail content – which increase the ability to monetise consumption intent) even in the mobile sector, the relevance of reputation based dynamics and the potential of the mobile as a unique device that can automatically add semantic data to content captured on the phone (meta data enabled content captured from mobile devices at the point of inspiration)

Business drivers were also presented, one of them becoming the growing relevance of privacy related issues, the progressive leaning away of advertising from mass media, the emergence of different forms of content discovery (platform and service specific search forms), the increasing recognition of the key role of recommendation engines, the awareness that in general, openness drives better search since more content is accessible and that content has better links (i.e. is

referenced better), the recognition of the need of deeper integration between devices, networks and services, the emergence of new search models like reverse search (where the source is found from the content – ex: Shazam (for music) and Tineye for images) or discovery based models which tend to 'searchless' search and reverts the concept of search to an 'agent' which fetches information based on a set of parameters.

The above mentioned issues indicate the way the mobile sector will evolve and will have to be matched with the following dynamics affecting the mobile device (such as emergence of new devices, a greater role for Internet of things through the evolution and the dominance of the browsers, the appearance of new interfaces like touch screens, 3D etc and the capability of devices to capture). In addition, the range of factors are leading to an innovation and evolution of the ecosystem, including Google Gears/offline browsing, open source ecosystems, Widgets, Javascript enhancements(Chrome, JS libraries), Location including Cell id databases, SIM/Smart card Web server, APIs (GSMA, OpenAjax, Bondi, Gears, RCS), Browser plugins (MS Silverlight, Mozilla,W3C), Social network APIs, Local Web, Near Field Communications.

Ramón Compañó started his talk by highlighting the dilemma that improved search capabilities have enabled unprecedented added-valued services but at the same time has raised concerns on misuse. As both are different sides of the same coin and need to be discussed together. Technology allows for ever increasing storage capabilities, increase of processing power higher performance in search. The abundance of (personal) data combined with the increasing power to process it, generates a tension between adding value and invading citizens' privacy. Access to user data is useful to improve the relevance of search results (personalization), to improve the targeting of ads (revenue) and to identify phishers and malware distributors. This might create, however, a tension between the search engine providers and the user.

Compañó explained some of the European Commission activities to preserve citizens rights to privacy while keeping attention not to fall into the error to design regulations that would be detrimental for technological innovation. He introduced the "Article 29 Working Party" opinion on search engines, whose motivation included to promote both industry competition and citizen enforcement. Here, the obligations with regard to fair, lawful and transparent processing of personal data, is not limited to retention, but also to access, correction, deletion, information rights of data subjects. The access right to these data and action is key. The OECD privacy principles are still valid for use. The Opinion outlines some obligations with regard to fair, lawful and transparent processing of personal data by search engine providers and received some criticisms form industry, particularly on the applicability of the proposed measures, or the proposed data retention periods.

Future challenges for data protection offices include three major issues. The first one is to get closer to viable online-consent measures that specify the purposes of use. Second, to monitor and assess the impact, and take appropriate measures of the convergence of social networking and search engines, most notably people search including whereabouts (presence). Third, how to deal with behavioural profiling/targeting on the content of the profiles and/or traffic data of non-registered users.

2.8 Session 7: Use Cases - Bridging Benchmarking and Validation

This session addressed the question of evaluation of multimedia search engines. After some initial remarks from Jussi Karlgren of SICS, Stockholm, on the starting points of the CHORUS approach to evaluation, the CHORUS studies and findings were presented by Paul King of the Information Technology Institute, Thessaloniki. The CHORUS use case typology is presented in the project deliverables as well as this presentation. On the topic of evaluating information access technology for multimedia, Martha Larson of University of Amsterdam and Gareth Jones of Dublin City University presented activities from CLEF and the VideoCLEF evaluation track; Stefanie Nowak of Fraunhofer Institute for Digital Media Technology, Ilmenau, presented activities and resources from the imageCLEF track. These efforts will give system benchmarking

a large set of sustainable test materials to continue working with. Xavier Vives Surroca from Catalunya TV, Barcelona, and participant in the SEMEDIA project, presented how professional users in a media production setting can contribute to system design and validation - he showed several examples of how prototypes can inform system development through rapid reaction, if the stakeholder partner in a project is involved early and throughout the system development process.

The CHORUS approach is characterized by the notion of use cases. Benchmarking, the activity of testing system performance, can be done independently of user studies, but needs to be informed by studies of usage to establish relevant parameter settings and target notions. The presentations by Martha Larson, Gareth Jones, and Stefanie Nowak underscore the level of effort put into benchmarking. User studies on the other hand, are by their nature always centered on the specific case, and risk becoming irrelevant and their results lost for future efforts unless they are related to more general and abstract problem formulations: the efforts made in the SEMEDIA project, presented by Xavier Vives Surroca demonstrate the effort made by informed user studies. The CHORUS project has advanced the notion of use cases as a conduit between engineering and benchmarking on the one hand and user studies and validation on the other. Use cases can be established with little effort in a development project. User studies, if performed properly incur considerable effort but if used to validate use cases can provide sustainable results for future projects, which, if they accept the use cases in question as relevant for their development efforts, then will have less need for user studies of their own.

The discussion centered on some concrete issues of

- How can the quality of prototypes be ensured to gain acceptance of test users? And how can behavioral results, such as provided e.g. by the SEMEDIA project, be taken and generalized into use by other projects? Suggestions were made both to suggest that each development project needs to make their own user studies and share methodology rather than results; and to suggest that suitably formulated use cases might be a way to package the results of use studies for future use.

- How can test sets built in various ways by different groups of researchers and developers be shared in view of different knowledge models? As a specific case of the interoperability problem, the consensus of the audience appeared to be that best practice in constructing test sets would address these questions to some extent and that the world and human information processing being the way it is they might never be solved conclusively anyway.

- How to leverage efforts made by benchmarking exercises such as CLEF to larger sets and to live data streams as suggested in earlier sessions to be more relevant for the success of future projects, rather than working and testing on archival data sets. This question was left unanswered, but suggestions of learning technologies for this purpose were made. A related discussion was that of reproducibility of test results - if research is performed using live data streams rather than archival data sets, how can the results be guaranteed to be comparable? Physical and chemical research often is reproducible across data sets, where the method is documented in detail rather than having experimental data shared across laboratories.

The final question, left for further debate, is if use cases could be used to steer the research activities within a public call for research proposals in some desired direction?

3.   CONCLUSIONS

The two days of the CHORUS conference were separated into 7 sessions of different topics such as "International Perspectives", "Industrial Trends" and "Challenges beyond Technology". The consortium received an overall positive feed-back from the participants and the stand exhibitors, for example through the participation on the CHORUS mailing list, or through the publication of news and events on the CHORUS Web site. An article regarding the conference was published on the search engine blog Pandia.com.

All the seven sessions were well frequented, even at the end of the second day when people normally tend to catch their flights. The stand exhibition was visited by most of the participants during the coffee breaks, but also during sessions it was possible to encounter lively discussions at the booths.

The presentations in the sessions were more or less technically detailed, depending largely on the topic and the background of the speaker. It could be said that the technical discussions slightly exceeded the socio-economical ones, although the latter one tends to be "overseen" by users and professionals (discussion on legal and social aspects is running behind technical development), and therefore would have needed an equal emphasis. In the introductory session the Commission concluded that search engines in the multi-media domain will play the role of key-enablers for the Future Internet, which was again confirmed by the penal session at the end of the first day: search functionalities will be pervasive, always present, although not necessarily directly visible as a standalone service or application. Especially the latter aspect makes it difficult for the normal end user to understand the technology and thus understand the discussion on legal and privacy problems related. As discussed in session 6, especially data protection offices need to play a stronger role in view of the future challenges. The level of diversification in research and development was impressively demonstrated in the 4th and 5th session, where various projects presented their current status and future goals, also demanding at the same time an increased effort to synchronize research in this domain.

The main and most general conclusions out of the CHORUS conference could thus be:

- Search Engines will concern (and affect) the end user('s behavior) in almost every step he makes in the Future Internet

- Research and development needs to synchronize its efforts on a European and world wide level

- The user needs to be informed (and willing to learn) about potential legal and social consequences of his activities in the Internet

- The discussion on legal and socio-economic aspects needs to catch up with the technical development

In addition to this report the conference is further described in D0.3 "Report on Rising Public Participation and Awareness" and D0.2 "Dissemination Plan". Other documents related to the conference are the "CHORUS Executive Summary" that was published during the conference, D3.4 "CHORUS Vision Document" and D2.2 "CHORUS Gap Analysis".

The project presentations are downloadable from the Conference Web site:

http://www.ist-chorus.org/conference_downloads.asp

The conference videos can be watched at VideoLectures.net:
http://videolectures.net/chorusfc09_brussels

The list of conference participants is available from JCP Consult

- End of document -