



CHORUS

Coordination Action on Multimedia Search Engines

<http://www.ist-chorus.org/>

Executive Summary

Objectives:

CHORUS set up an information exchange platform between the EU projects, national initiatives and key players in the domain of multimedia search engines (MSE). CHORUS activities aim to bridge the gap between *researchers view* (academia and industry) and the *new services* within a market *prospective* in the MSE area. Through its different structures and events: working groups, Think-Tank, "A-V Search" cluster and workshops, CHORUS tackles identifying and deriving *critical issues* through technological aspects together with socio-economic and legal aspects.

Search evolution and prototypical areas of business:

Search will be a pervasive and ever present function fully integrated into all applications and oftentimes invisible to the user - users will not realize they are posing queries.

Six prototypical areas of business have been identified¹ with regards to their different salient characteristics and attributes: content management, content ownership, access rights and the revenue model: **(1)** Web search, **(2)** Personalised TV, **(3)** Enterprise content search, **(4)** Library search, **(5)** Personal content search, **(6)** Monitoring, Detection & Alert.

Each of these areas gives rise to number of challenges both at the technical and the socio-economic and legal aspects.

Cross-disciplinary challenges and recommendations:

CHORUS elaborated and agreed to a common **functional breakdown** model of a generic search engine² regardless of the application domain or business sector. It presents the benefit of shared projects' description and vocabulary across industry and academia. Search is about *making best use of available meta-data*³ to provide the user with *meaningful information* in spite of the fact that the user's request is possibly poorly formulated and typically *unanticipated*. Keeping the "user in the loop", maximize its efficiency.

After conducting gap analysis studies, CHORUS identifies several directions that deserve European effort toward 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 recommend 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 ...).

¹ See "Vision" deliverable for more details

² See D2.2 deliverable on "Gap Analysis" for more details

³ Metadata are defined as being all information besides the raw content that make the content searchable: manual annotations, automatically generated low-level signatures, automatically generated semantic labels, device generated media context ...

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 number of users, number of information sources, 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.

Organisational Recommendations:

Coordination effort through CHORUS allows positioning various EU efforts among the technological landscape dimensions. The mapping reveals that Europe is **rich with very sharp expertises** 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 consumer search market, Europe is lacking today an integrated program that gathers all needed expertises for building competitive real life search engines. CHORUS recommends to:

1. **Empower aggregation and orchestration of such expertises into an operational end-to-end search organisational 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).

How should the Future Internet be in 2020 from the search perspective?

Media will be plentiful all over the Future Internet (FI), often in distributed form. Making all media **searchable and accessible** (metadata generation and structuring) is the major challenge for the FI. Users will be accustomed to retrieving what they need delivered just-in-time. Some users will choose to deal with data, applications, storage from the "cloud" as a service infrastructure.

Search is at the **application** level: the network should be neutral (agnostic) with respect to applications. One exception can occur for the media delivery optimization stage (network layer) which impacts the user QoE.

Search engines will be omnipresent, often embedded. Hence, they will not have significant differentiating impact on FI: either **unique** or **multiple Internets**: Internet of Things, Internet of Services, and Internet of Content. Search operates on **metadata of searchable entities**.

The derived knowledge from the **growing power of social networking** needs to be highlighted. It may represent an alternative to information retrieval through trusted recommendation mechanism.

Interoperability, in this context, appears as a major concern. It will need to be carefully addressed in the coming years beyond the Internet of Media reaching all networked entities including "Services" and "Things".

Contacts:

Jean-Charles Point: pointjc@jcp-consult.com, Nozha Boujemaa: Nozha.Boujemaa@inria.fr