# Cross-Language Information Retrieval in QUAERO-MSSE Project

Malek Boualem*        Jean-Philippe Cabanal*        Nabil Bouzerna**

(*) France Telecom Orange Labs        (**) Bertin Technologies

| List of partners (MSSE project has only French partners) |
| --- |
| Orange Labs, France (coordinator) |
| Bertin Technologies, France |
| IRCAM, France |
| LTU Technologies, France |
| Vecsys, France |
| Institut National de l'Audiovisuel, France |

**Project duration (5 years): May 2008 - April 2013**

**Summary**

MSSE (Multimedia Search Services for European portals) is one of the application projects of the QUAERO collaborative program. This project aims to develop a searching and navigation service prototype allowing to access to various audiovisual contents. The major innovations lie in the use of advanced content analysis technologies which enrich automatically their description, and thus improve the relevance of the results and the user experience. MSSE addresses video contents in several languages: French, English, German, Spanish and Arabic. To allow a French user (for instance) to access contents in other languages, we have integrated a Cross-Language Information Retrieval function based on machine translation of metadata (video titles and summaries). Indexation of metadata is operated in French language in order to meet French queries. To maximize the efficiency of machine translation, we use two machine translation systems: the Bertin statistical MT system and the Systran rule-based (hybrid) one. The Bertin MT system is based on the RWTH (Rheinisch-Westfälische Technische Hochschule Aachen) MT technology. The back office of machine translation uses XLIFF XML Localisation Interchange File Format. Next steps of the CLIR functionally aim to operate machine translation on the queries instead of the metadata of the whole contents. Also some experiments will be done on translating texts resulting from speech to text transcriptions. A demonstration of the MSSE prototype is available.