



Multilinguality in Information Access to Digital Libraries

User Needs and Evaluation of multilingual resources use

Workshop at ICSD 2009
Trento, Italy 9 September 2009

Motives & objectives for the workshop

- Questions:
 - What do we know about multilingual access to DLs?
 - Which steps should be taken on the way to a truly multilingual DL system?
 - What do users really want wrt multilinguality?
 - How do we evaluate the quality of the multilingual access features of a DL?



Workshop presentations

- Multiplicity of languages & media
- Multiple languages: work – metadata – subject access – NE / proper names – user contributions / social data
- Inter-document linking
- Named entities & proper names
- Semantic enrichment / classification
- Domain-specific dictionaries
- Input language selection
- Interaction with the system (write your own translation)
- Context-dependent retrieval / NER
- Semantic / translation ambiguities



Workshop presentations

- User behavior to enhance translation (log analysis)
- Results presentation (relevance, language)
- Query expansion (geographic expansion)
- Mapping / crosswalks between subject vocabularies
- Association / aggregation of „the same“ documents („work“)
- Transliteration / data formats
- Interoperability / compatibility of resources & tools
- Evaluation – new user-determined measures
- From project to practice – implementation
- Sustainability of development / resources
- Re-use of project results / resources



Goals for road map

- Identifying themes of MLIA issues
- Priorities for implementation
- User needs in each topic area
- Approaches / problems in each topic area
- Assessment methods
- Evaluation of quality

- Best practices
- State-of-the-art overview



Survey Questions

- User profile incl. native & other languages and DL use
- Multilingual content interaction
- Multilingual user interface
- Information access and retrieval
- Multilingual information retrieval
- Multilingual query formulation & expansion
- Multilingual results presentation



Themes

- Metadata organization / structure / modelling
- Language Resources
- Encoding
- NLP / Linguistic data processing / indexing
- Controlled vocabularies translation / mapping
- NER
- User interfaces for search
- User interfaces for browsing
- Query formulation
- Language detection
- Query translation
- Query expansion / feedback
- Result presentation (order)
- Query refinement
- Result translation
- User interaction features
- Personalization features
- Log analysis
- Evaluation
- → Architectural infrastructure
- → basic building blocks



Issues

- financial sustainability for maintenance / updating (transfer of responsibilities from researchers → business)
- scalability of system (partners, documents, users) → efficiency issues
- transparency of retrieval process
- Interaction capabilities for users
- language-dependent Ranking?
- “why-button” for ml results set
- user Expectations of how search works
- strategies for query refinement



Issues

- concept-based retrieval = language-independent
- language detection by locale
- default language for DL?
- parallel documents / information object / document alignment (ml versions of the same document)
- non-translation (when do we not want /are unable to translate certain phrases?) – shades of meaning



Essential questions

- What do users want / expect?
- Which are the minimal requirements for a multilingual DL system?
- What are the most important features of a multilingual DL system?
- How do check / test / evaluate the functionality and quality of a multilingual DL system?



Topics of Discussion

- User & usability indications for multilingual access within digital libraries
- Analysis of the impact of multilingual / multicultural differences on interface / search design
- Identification of focus areas for multilingual access: media types, object models, retrieval algorithms
- Preferences for presentation options for multilingual content: result translation, ranking options for different object types depending on multilinguality
- Architectures and services for multilingual information access in digital libraries
- User behaviour and log analysis for multilinguality
- Contextual information retrieval and multilinguality
- Best practices for multilingual retrieval: manual, automatic or interactive functionalities

