Query Methods

• Which methods are needed to query real time multi-lingual data streams?

• How can natural language understanding technologies and keyword-based query methods help non-expert users?

• Which Semantic Web technologies can support this task?
Outline

• Core technical and methodological challenges
• Critical technologies for each challenge dimension
Core Technical Challenges

- Comprehensive and efficient semantic matching (schema-agnostic/vocabulary-independent)
- Multi-linguality (semantic processing, translation)
- Search over semantically heterogeneous, large-schema/schema-less, decentralized data
- Contextualised search
- Dataset discovery
- Search over stream data
- Transportability/portability
- Queries over abstract, complex or non-symbolic datasets
Core Methodological Challenges

• Multi-disciplinary research (IR, NLP, Semantic Web, Databases, UX/HCI, Information Seeking Behaviour, Machine Learning)
  – Multi-disciplinary cross-domain evaluation

• Horizontal vs vertical case studies
  – Domain specific (e.g. Enterprise)
Notation

• Topic is central in the context of KEYSTONE
• Common central topic in information retrieval, databases, etc
• More mature/consolidated topic
Comprehensive and efficient semantic matching

- Hybrid matching models (logical, statistical, linguistic resources, graph-based metrics)
  - Statistical-based semantic matching methods (distributional semantics)
- Integrating large data resources into the search/query process (e.g. linked data sources, linguistic datasets)
- Context-based semantic matching
- Compositional semantic models
- Stream-based realtime semantic matching
- Semantic indexing models for structured data
- Entity resolution/disambiguation
Multi-linguality

- Coverage of under-resourced languages
- Generation of multi-lingual linguistic resources
- Ranking models for translation
- Multi-lingual indexing
- Integrating machine translation algorithms into data indexing and search
- Representation standards for multi-lingual/linguistic resources
  - e.g. Lemon, LMF, NIF (OLIA, MARL)
User interface/Experience/Info. Seeking

• Natural language interfaces (NLIs)
• Hybrid seamless keyword/NLIs/visual interaction paradigms
• User interaction (feedback/dialog/disambiguation)
• Navigation over the search space (generalizing specializing), exploratory search, facets
• Results visualization
• Personalization/user adaptation
• Search focused in impaired/special needs users
• Software agents as users
Contextualised search

• Automatic data source discovery
• Incremental knowledge refinement
• User profiling/personalization/user adaptation/cultural adaptation
  – Integration of Machine Learning into search
• Task representation
• User objective recognition
• Integration of spatio-temporal context
• Integration of social network data
Semantic query processing

• Semantic query processing/optimization approaches
  – Semantic query planning
• Federated search (keyword, NL, keyword)
• Integration with stream-based models
  – Model evolution
• Privacy/anonymization
• Post-processing
  – Measuring (and ranking by) Data quality
  – Resolving inconsistencies
  – Results conceptualization/clustering
• Spatio-temporal querying
• Query expansion/reduction
Topics and associated Resources/References (In Progress)

http://bit.ly/1wbXLZV
In progress

Relevance for KEYSTONE

Impact

Hybrid matching models
Search using large data resources
Context-based semantic matching