Benchmarks for Keyword-based Search

• Benchmark
  – Use cases
    • Dataset
    • Query
  – Expected Result
  – Metrics

• Fact: Too many benchmark efforts
Benchmarks for Keyword-based Search

• Repository for benchmarks
  – Different scenarios benchmarks
  – Services
    • Analysis of results by other researchers
    • Quality Information / Fit of use for specific tasks
      – reputation of the datasets
      – Coverage / Dimensions
      – Metrics
  – Have Query Functionality
  – References to uses/applications of the benchmarks
  – Benchmark Spaces

• Benchmark as a service
  – “We run it for you” ..
    • Or at least we partially do the job
Benchmarks for Keyword-based Search

• Dynamic Datasets
  – Streams
  – Datasets cannot be old
  – Coordinating Support Actions
    • For research and Industry
      – Different requirements
    • Standardized procedures for benchmark specification

• Repeatability
Benchmarks for Keyword-based Search

• Global and Open Standards
  – Extendable
    • For tasks that have no benchmarks
  – Aggregation of Existing Standards
    • e.g. Berlin Benchmark, INEX, TREC, etc.
  – Crowdsourcing + Experts
  – Format of the data
    • Expert & Industry requirements
    • Linked-Open-Data Formats
  – User testing to set the base lines
Benchmarks for Keyword-based Search

- Leverage Existing Initiatives
- Multilingualism
- Real user behavior
- The value of the queries in a dataset in terms of information needs of the user (same level of details)
Benchmarks for Keyword-based Search

• Visualization
• Ability to create Metadata/Annotations
• Coarse and Fine Grained datasets for inexact matching
• Complex Scenarios from commercial application
• Interoperability among benchmarks