



Evaluation

Nicola Ferro

Information Management Systems (IMS) Research Group

Department of Information Engineering (DEI)

University of Padua, Italy

- Definitions
- Use Cases
- Data sets
 - differences in using open data
- Solutions / Approaches

- Evaluation

- evaluation vs benchmarking

- evaluation is for a purpose, benchmarking is generic

- evaluation is not only part of design but as a continuous quality assurance / monitoring

- efficiency-oriented

- effectiveness-orientes

- user-oriented evaluation

- holistic system evaluation

- focus groups

- information seeking behavior

- Keyword-based search

- semantic keyword is not something that is in the data but something is meant in the data

- topical search

- question answering

- Clear use cases (RepLab) grounded in reality, may be domain specific
- Examples
 - travel planning
 - good weather conditions for surfing (domain knowledge, contextual, ...)
 - real estate market: I want to buy a property in a foreign country plus smart city data
 - catalog for complex products / product configurator

- Quality / Trust in the data set
- Data set management life-cycle (provenance can inform search for patterns in the logs)
 - curation of the experimental data (explicitly attach the experimental setup)
- Open data sets
 - they may provide a real scenario
 - multisource
- Examples
 - Wikipedia, Eurovoc (resources), Transition Memory
 - System logs
 - ...

- real estate market:
 - structure is important, coordinate system framework (depth, dispersion)
 - session-based
 - diversity
- product catalog
 - marketing people, buying guys, engineers
 - CAD drawings, beneficial for the environments, money
 - diversity
 - final goal: number of purchases
- Response time, user feedback
- Usual measures: precision, recall, accuracy, ...
- Ground-truth creation (domain experts, crowd-sourcing and gamification)

Thank You

