

Information Retrieval — Example Exam Questions

Norbert Fuhr

SS 2016

We urgently recommend that you practice to answer these questions loudly - ideally to a colleague (but a rubber duck or a teddy bear will also do ;-)).

Remark: These are only examples for queries that might be asked during the oral exam - this is not a complete list of exam questions!

- Introductory questions: please talk 3 minutes about one of the following topics. Mention the issues and how these problems can be solved:
 - IR and databases
 - Ontologies and OWL
 - Document classification
 - Hierarchic clustering
 - Learning to Rank
 - Inverted lists
 - Interactive retrieval
 - Web Search
 - Social Media Retrieval
- Models based on predicate logic
 - What is the difference to models based on propositional logic, and what are the advantages?
 - What is the difference between extensional and intentional semantics? How can the latter be implemented using event keys and event expressions?
 - Which aspects of database query languages have not been fully supported in IR-based approaches so far?
 - What are the advantages and possibilities of XML Retrieval?
 - What are the major elements of an ontology? How can it be used in IR?
- Learning and classification
 - Explain the difference between classification and clustering!
 - Explain the core ideas underlying naive Bayes (NB)
 - What are the theoretical weaknesses of NB? 'Why does NB work in spite of these problems?'
 - What is 'learning to rank' (LTR)? Explain the basic method!
 - What is the advantage of LTR in comparison to classic retrieval models?
- Clustering
 - What are possible applications of clustering? Which types of clustering methods can be distinguished?
 - Describe the k-means algorithm!

- 'What is the computational complexity of k-mean in comparison to agglomerative hierarchic clustering?
- What are the two main problems with k-means and how can they be addressed?
- Which types of hierarchic methods (agglomerative/divisive) do you know?
- What are the best agglomerative methods?
- How is optimum clustering defined? How can it be achieved?
- What does the optimum clustering framework say about exiting clustering methods?
- Indexing and Searching
 - Describe the structure of inverted lists, and how they can be used for efficient retrieval!
 - How can inverted files be efficiently constructed?
- User Interfaces
 - What are the assumptions underlying the interactive PRP, and what does it say about interactive IR? How can it be used for improving interactive IR?
 - Describe Ellis' seeking model and its possible consequences for the design of interactive IR systems!
 - Describe the Select/Organize/Project model
 - What is the benefit of having a taxonomy for Web search queries?
 - 'What are the different levels of search activities? Give examples from Web search engines!
 - Describe the different degrees of system support! Give examples from Web search engines!
 - Which functions of search interfaces are essential for supporting search sessions?
- Web Search
 - Describe the major components of a Web search engine! How can it be made more efficient via cluster-based architecture and caching?
 - What are the specific requirements for ranking in Web search engines?
 - Describe the basic concepts underlying PageRank and HITS!
- Social Media Retrieval (SMR)
 - What is the major difference between SMR and classic document or Web retrieval?
 - Explain the richer representation with the example of a popular social media web site!
 - How can the richer representation be used for retrieval?
 - What is the role of User Context, how can it be captured and used for IR?
- Multimedia IR
 - What are the major differences between text and multimedia retrieval? Explain the 'semantic gap'!
 - Which image properties should not have an effect on the retrieval method?
 - Explain the basic methods for colour- and texture-based retrieval!
 - What are 'salient points' in image retrieval?
 - How does 'fingerprinting' for music retrieval work?
 - Explain the basic methods for generating video abstracts!
- Multilingual IR (MLIR)
 - What is the purpose of an MLIR system, and what are possible applications?
 - Describe the architecture of an MLIR system
 - What are possible approaches for query translation in MLIR?
 - What are the challenges for interactive MLIR?