Automated Subject Indexing in Libraries

Jan Simane
Scope of indexing literature

1. Retrieval
2. Classification/Systematization
3. Description
4. Selection
(intellectual) subject indexing under discussion

- High investment of time
- Benefit?
- Redundancy!
- Alternatives
on the other hand ....

- most searches performed are 'about' („aboutness“)
- aboutness data enhance discovery
- controlled vocabularies strongly support data linkage in the semantic web
Automated, machine-based indexing?

- linguistic analysis (subject headings)
- machine learning (classification)
- evaluation of links to authorities
- accumulation („Digital Assistant“)
P & C

- Fast processing of Big Data
- Constant refinement of analysis algorithms
- Rudimentary subject information is better than no subject information
- In line with the LOD paradigm

- Complexity of language(s)
- Misinterpretations (synonyms, homonyms)
- Limited relevancy recognition
- General loss of quality and exactness
Questions

1. When *Automated Indexing* is acknowledged as an appropriate bibliographic method, does it mean that the intellectual input in traditional subject indexing is dispensable?

2. Will the inferior quality and precision of *Automated Indexing* have an (negative) effect on retrieval processes?

3. Do advantages of *Automated Indexing* outweigh its shortcomings?