

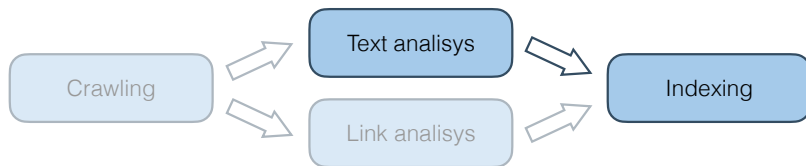
# Information Retrieval 1

## Query Processing

**Ilya Markov**  
i.markov@uva.nl

University of Amsterdam

# Recap IR0



# IR1

Evaluation

Document  
representation  
& matching

Conversational  
search

Learning to rank

IR—user  
interaction

Recommender  
systems

# Evaluation

Evaluation

Document  
representation  
& matching

Conversational  
search

Learning to rank

IR—user  
interaction

Recommender  
systems

# Document representation and matching

Evaluation

Document  
representation  
& matching

Conversational  
search

Learning to rank

IR—user  
interaction

Recommender  
systems

# Outline

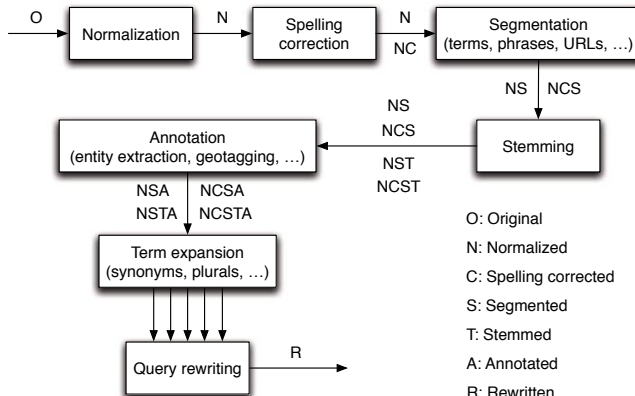
- 1 Query analysis
- 2 Query processing

# Query analysis pipeline

- Normalization
- Spelling correction
- Segmentation
- Stemming
- Term expansion
- Etc.

Should be the same as  
the document processing pipeline

# Query analysis pipeline



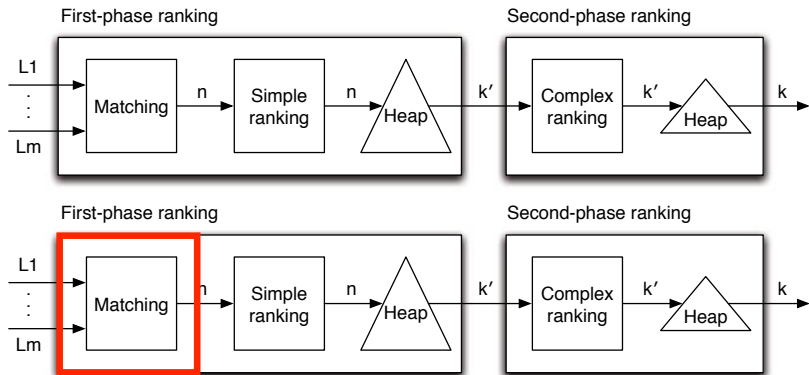
B. Cambazoglu and R. Baeza-Yates, "Scalability Challenges in Web Search Engines"



# Outline

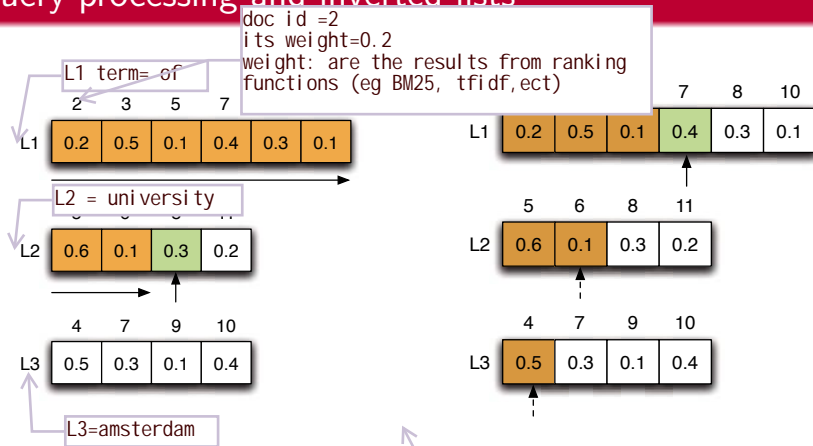
- 1 Query analysis
- 2 Query processing

# Query processing



B. Cambazoglu and R. Baeza-Yates, "Scalability Challenges in Web Search Engines"

# Query processing and inverted lists

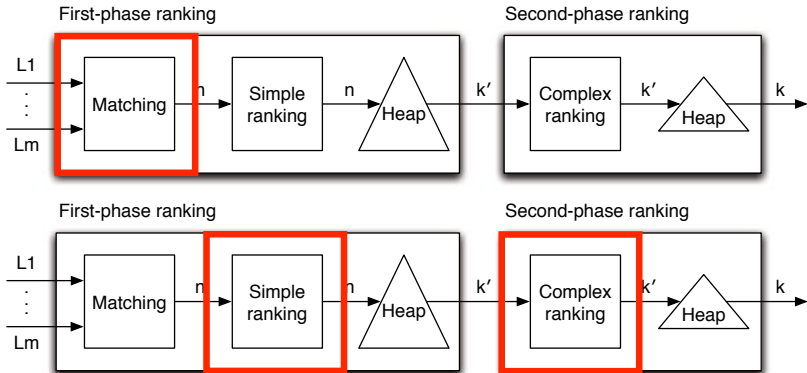


we have a query, and lots of docs  
goal: select one of the words in the query. look for which doc contains this term?  
left=right inverted list

# Practical considerations

- Conjunctive mode (AND)
- Document-at-a-time
- A score is usually computed as a linear combination of query-dependent and query-independent scores

# Next topics



B. Cambazoglu and R. Baeza-Yates, "Scalability Challenges in Web Search Engines"

# Materials

- Croft et al., Chapter 5.7
- B. Barla Cambazoglu and Ricardo Baeza-Yates  
**Scalability Challenges in Web Search Engines**  
Morgan & Claypool Publishers, 2017