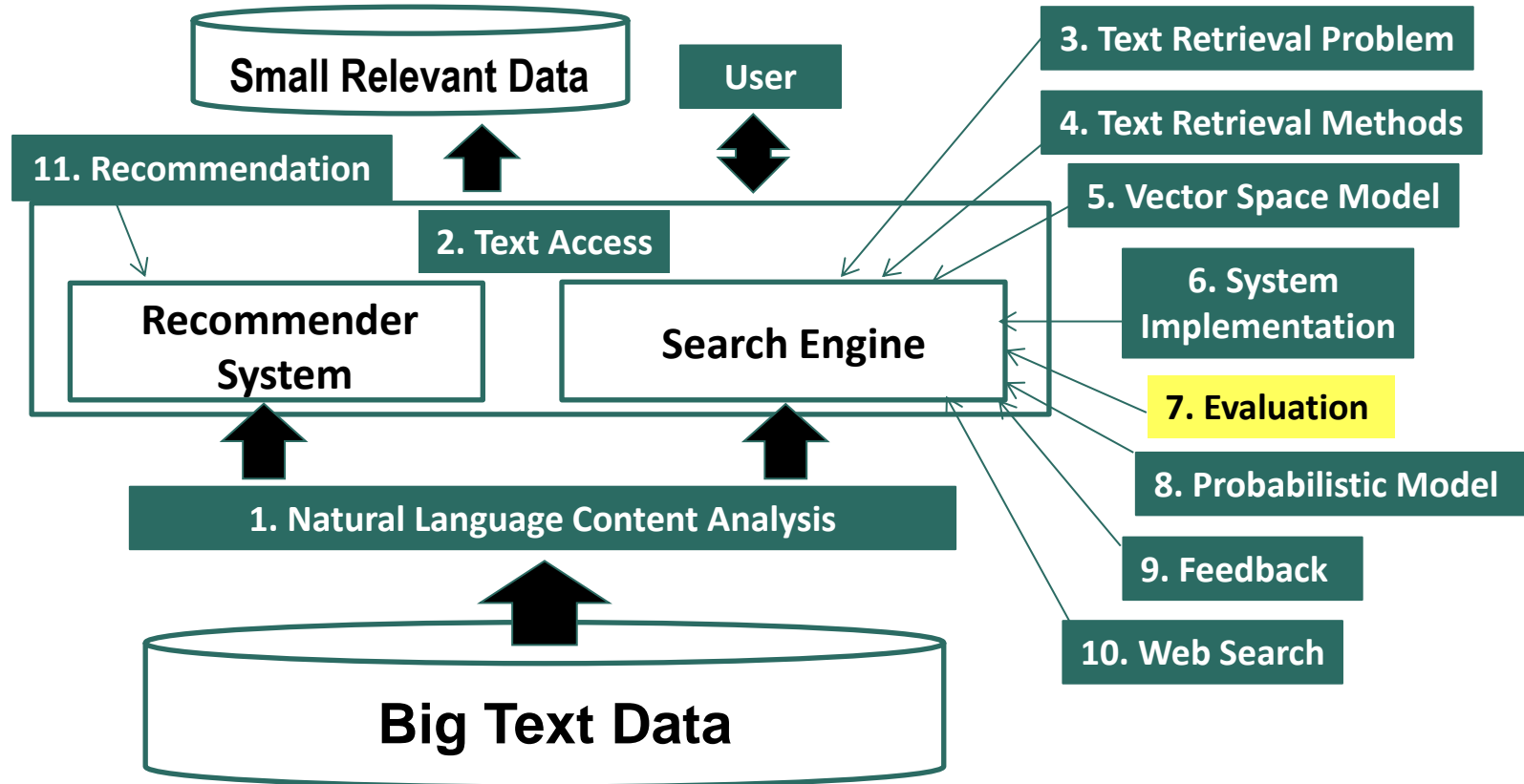


# Text Retrieval and Search Engines

Evaluation of TR Systems: Multi-Level Judgments

ChengXiang “Cheng” Zhai  
Department of Computer Science  
University of Illinois at Urbana-Champaign

# Evaluation of TR Systems: Multi-Level Judgments



# What If We Have Multi-level Relevance Judgments?

Relevance level:  $r=1$  (non-relevant) ,  $2$  (marginally relevant),  $3$  (very relevant)

	Gain	Cumulative Gain	Discounted Cumulative Gain		
D1	3	3	3	Normalized DCG=?	
D2	2	3+2	3+2/log 2		
D3	1	3+2+1	3+2/log 2+1/log 3		
D4	1	3+2+1+1		$\frac{DCG@10}{IdealDCG@10}$	
D5	3		...		
D6	1				
D7	1	...	DCG@10 = 3+2/log 2+1/log 3 +...+ 1/log 10		
D8	2		IdealDCG@10 = 3+3/log 2+3/log 3 +...+ 3/log 9+ 2/log 10		
D9	1				

D Assume: there are 9 documents rated “3” in total in the collection

# Normalized Discounted Cumulative Gain (nDCG)

- Applicable to multi-level judgments in a scale of  $[1, r]$ ,  $r > 2$
- Main idea of nDCG@k documents
  - Measure the total utility of the top k documents to a user
  - Utility of a lowly ranked document is discounted
  - Normalized to ensure comparability across queries