

Summary of System Implementation

- Inverted index and its construction
 - Preprocess data as much as we can
 - Compression when appropriate
- Fast search using inverted index
 - Exploit inverted index to accumulate scores for documents matching a query term
 - Exploit Zipf's law to avoid touching many documents not matching any query term
 - Can support a wide range of ranking algorithms
- Great potential for further scaling up using distributed file system, parallel processing, and caching