

Summary

- Web search is one of the most important applications of text retrieval
 - New challenges: scalability, efficiency, quality of information
 - New opportunities: rich link information, layout, etc
- Crawler is an essential component of Web search applications
 - Initial crawling: complete vs. focused
 - Incremental crawling: resource optimization