

Data Structures for Inverted Index

- Dictionary: modest size
 - Needs fast random access
 - Preferred to be in memory
 - Hash table, B-tree, trie, ...
- Postings: huge
 - Sequential access is expected
 - Can stay on disk
 - May contain docID, term freq., term pos, etc
 - Compression is desirable