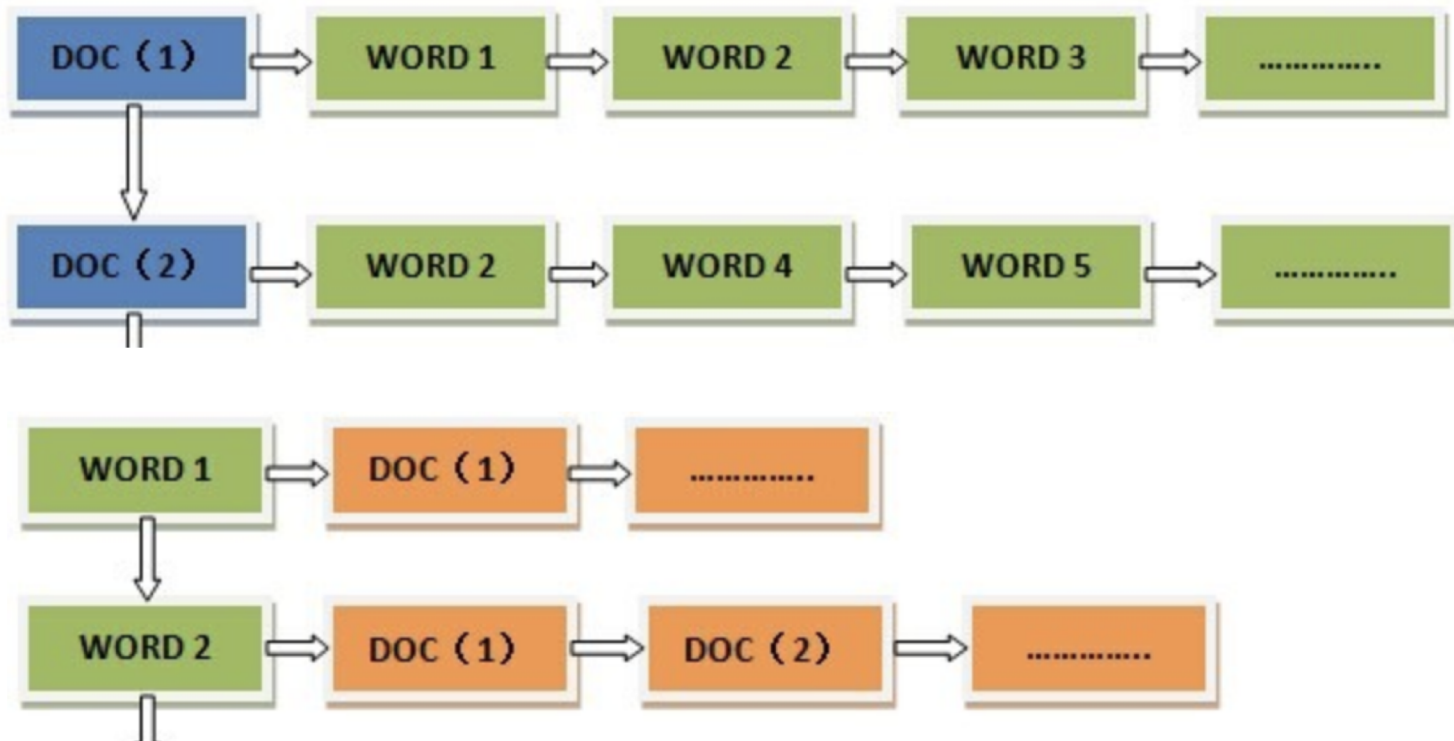


Twitter Keyword Search

Zhibin Huang
Mingdao Che

Introduction: indexing



Introduction: forward

DOCID	DESCRIBE
NO.1	[Tom] [like] [apple] [banana] [tony] [like] [apple]
NO.2	[Tom] [like] [apple] [orange]

Introduction: inverted

WORD	DESCRIBE (DOCID[FREQUENCY],POSITION)
TOM	NO.1[1] NO.2[1],1 <u>1</u>
like	NO.1[2] NO.2[1],2 8 2
apple	NO.1[2] NO.2[1],4 9 3
banana	NO.1[1],6
orange	NO.2[1],5

http://blog.csdn.net/chunlei_zhang

Introduction: Apache Lucene

- Text Analysis and Compress
- Scalable, High-Performance Indexing
- IndexReader/ Searcher