

Stage 1:	Stage 2:	Stage 3:
Lexicon	Forward Index	Inverted Index
Input: dataset (tokens)	Input: dataset + Lexicon	Input: Forward Index
Output: lexicon (token->id, df)	Output: forward index segments	Output: inverted index segments

|

v

Stage 4:
Ranking Precompute
Input: Forward Index + Lexicon
Output: dl.bin, avgdl.bin, idf.bin

|

v

Stage 5:
Query Engine
Input: Lexicon, Inverted Index, Stage 4 binaries
Output: Top-K search results

|

v

Stage 6:
Barrels + Cache
Input: Inverted Index segments
Output: Sharded barrels, LRU cache, manifest

|

v

REST API / Search Service
Input: Queries
Output: JSON results

