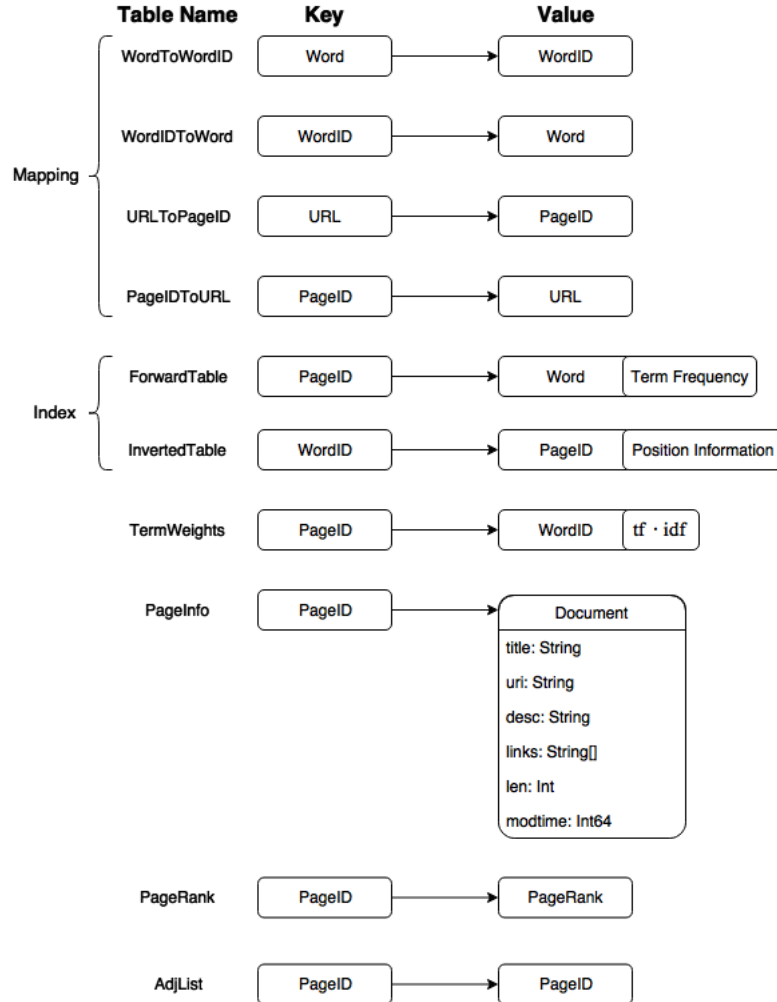


# Database Schema

Adrian Prawira Susanto and Kevin Pratama and Randitya Setyawan Mohamad



The package used for our database manager is Bolt for Go, which is a pure key/value storage package for Google's Go Language. As for the schema itself, we partitioned the tables into three big groups:

1. Mappings

These tables contain convenient mapping from one value to another, to allow ease of use during implementation.

2. Indexes

These tables are one of the most integral part of our search engine. The indexes allow for both forward lookup (From DocumentID to the words) and backward lookup (From Word to DocumentID), commonly known as Forward and Inverted index respectively.

3. Miscellaneous

These tables contain extra information which are used ubiquitously in the program. TermWeights, for example, store each term weights ( $tf \cdot idf$ ), which were precomputed to allow faster retrieval during runtime. PageInfo stores each document information which then are processed to the indexes, or used during crawling. AdjList, as can be referred from its name, stores the adjacency list of the indexed pages. From this AdjList, PageRank of each pages can be obtained and precomputed, which are then stored in the table PageRank.