

## Appendix F: Database Structure

Table	1 SELECT * FROM 'BM25resultsIR_Arabic';		
Arabic_Queries			
BM25resultsIR_Arabic			
BM25resultsIR_Czech			
BM25resultsIR_English			
BM25resultsIR_French	queryNum	item_id	relevanceScore
BM25resultsIR_German	clef2015.test.1	hair3061_12_000117	1
BM25resultsIR_Italian	clef2015.test.1	first1181_12_000830	0.9420099973866756
BM25resultsIR_Persian	clef2015.test.1	first1181_12_000404	0.921631659816425
BM25resultsIR_Portuguese	clef2015.test.1	nhslo3844_12_006773	0.9185305809924444
Czech_Queries	clef2015.test.1	nhslo3844_12_006778	0.9182256781089743
DSresultsIR_Arabic	clef2015.test.1	first1181_12_000035	0.9168663344548287
DSresultsIR_Czech	clef2015.test.1	women1681_12_000705	0.9162059220258686
DSresultsIR_English			

**Figure 1: Database**

Note that there are 3 columns for each different table in the database.

- queryNum: This is the number of the query.
- Item\_id: This is the document that is being evaluated for the given query.
- relevanceScore: This is the scoring of the document against the query.

There are tables in the database for the 3 different algorithms used, as well as a table for each of the different languages used.

- BM25resultsIR\_\*: is the naming convention for BM25.
- VSMresultsIR\_\*: is the naming convention for Vector Space Model.
- DSresultsIR\_\*: is the naming convention for Dirichlet Smoothing.
- \*\_Queries: is the naming convention for the normal queries when translated.
  - Note: '\*' is the different languages used.