# License Keyword Search

## Problem

We need to build a system that determines how often common terms appear in various software license agreements.

We have provided a database that includes a set of common license terms, as well as a list of references to licenses that we would like to parse and search. The referenced licenses may exist directly in the database, have a reference to a local resource, or reference a remote URL. The data for these tables, and their structure, is visible in the CSV/SQL files in the *src/main/resources* folder (terms.csv/terms.sql and licenses.csv/licenses.sql).

A skeleton project is provided that includes database described above and some code stubs that can be used to start your work.

## Solution

Our solution will be measured on

* Correctness
* Design and maintainability
* Speed and scalability (consider possible bottlenecks and multithreading)
* Documentation

Initial Unit Tests are provided and can be run through Maven (‘*mvn clean test’*). These can be used as a starting point for writing your own tests. Additionally the TermCountServiceTest will be used as an integration test to validate the completeness of your solution.

The surefire-report plugin for Maven is configured to generate reports for all tests, including the TermCountServiceTest acceptance test. To run all tests and generate an HTML report of status, simply run ‘*mvn clean surefire-report:report’* or ‘*mvn clean site’*, the HTML output will be generated to

*target/site/surefire-report.html*