4.3 Architecture

The design of our architecture ad hoc, meet our requirements. This is the creation of tool ready-to-use and of easy access. In this regard, as a user we want a web application where any user may be capable to build a corpus given a query, the same query a user may do through GitHub. This design was inspired by the fact that technical barrier sometimes remove the attention on the main objective, which for us is the discovering of requirements-related information. It is worth to mention, that exists a tool form dumping information from GitHub on demand [17] that certainly covers more information than just readmes, but several steps are needed to get it. The following figures (Figure 7, Figure 9, Figure 10) show the tool architecture with an implementation vision.

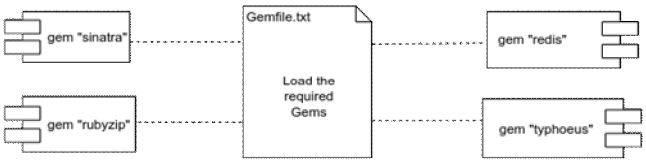

Figure 7: Libraries used

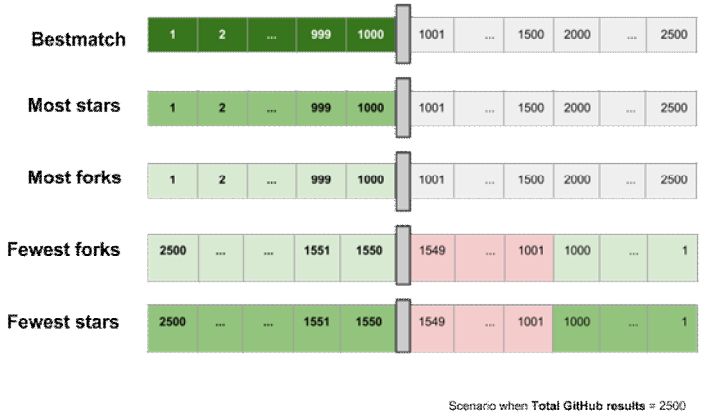

Figure 8: Retrieval of readmes combining sorting options

Those libraries permits the extraction and construction of a corpus in a web browser, taking into account that the quantity of readmes to be requested sometimes may be high, and this kind of load could affect the browser performance.

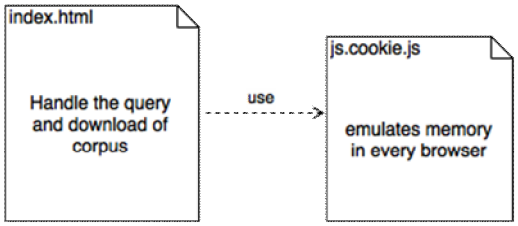

Figure 9: The view layer

As JavaScript does not possess an event indicating when the server ended a request, unless it is an AJAX library (Asynchronous JavaScript and XML), we had to emulate this behavior with the js.cookie library which provides a friendly API. Next, we describe to what is useful the gems required in our approach.

4.3.1 Sinatra Library

Is a Domain Specific Language (DSL), which allowed us the quickly creation of a web application

4.3.2 RubyZip Library

Permits the reading and writing in the creation of a zipped corpus of readmes.

4.3.3 Typhoeus Library

Allow us a faster processing of request. It supports parallel requesting wrapping curl library, which is used for transferring data.

4.3.4 Redis Library

Creates a data structure in memory, used as a database for the sets of information transferred.

4.4 Disclosure

We made available the whole code in the following link of GitHub: *omitted for blind review* and the project github-proxy: *omitted for blind review*, which is a complement to support the performance of the application.

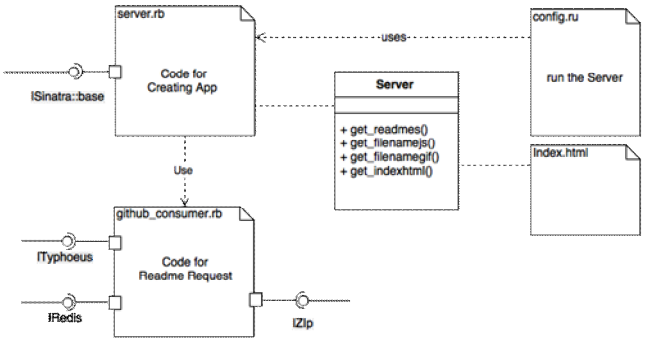

Figure 10: Implementation vision of the architecture