**Technology**

Database

It’s not clear to me what indexes there are.

A single character seems inadequate for part of speech in terms of distinguishing adverbs from adjectives and pronouns from prepositions; we might also want to designate word forms. It gets worse if you read the [Wikipedia article](https://en.wikipedia.org/wiki/Part_of_speech).

Coding

What is the technology stack?

**Functionality**

User Interface

* *I was relieved to hear Chris’s comment that he also favors a descriptive, English-based UI, rather than one with cryptic symbols like Qat. Given that, perhaps the requirements document should be tweaked to remove the examples which invent and use cryptic strings. That would also improve the readability of the document.*
* *What do you think of the general approach of the UI on alfwords?*
* *For multi-word searches, I like the equation solver as a fallback for unanticipated queries, but I think we should come up with a more intuitive UI for more common types of things. One possibility to consider is to have multiple instances of the single-word UI with a scheme to indicate how they relate; another is to allow saving result sets as word lists, then the UI can be a lot simpler as we only need to support selection of lists and the joining part.*

Search capabilities

* *Find two words where the second one differs from the first by repeated string replacement. For example, meter / motor, sleep / sloop, and bellyweed / Bollywood. More generally, any kind of regular expression replacement.*
* *Search based on letter weights. For example, any five-letter word which scores 60 points or more as an opening Scrabble play: hafiz, quick. Alternatively, use letter position in the alphabet or custom values.*
* *Search based on content of web page which sources the term. For example, for Wikipedia article titles, one might look for text on the page (either literally or as a regular expression), the page size, the creation or modification date, or the category (and possibly subcategory). There is also structured information in templates which we might want to access.*

Result display

* Provide a link to the original page (e.g., the Wikipedia article)
* Optionally show the canonical form, the alphagram, the bank, the length, the weight (see second search proposal), matching characters if a regular expression was used, or context if a search on the original source was done.
* Allow sorting on any of the above, or display the results as a sortable table.
* When a result is a word that is part of a phrase (or multi-word term), provide a link to list all matching phrases.
* Display result in either natural or canonical form
* Force display in upper or lower case and/or a fixed-width font
* Optionally show multiple hits per line

Word Lists

* *Definitions should specify the language (even if it will always be English for now)*
* *Add quality/frequency somewhere*
* *The requirements doc (I think) says that “Wikipedia has done a lot of work on scoring Popularity” and references the inclusion criteria. This is true, but it is only a two-level score—either a term is in the Encyclopedia or it is not. If we want a graded scale, we need to tease out the information in other ways, such as Google counts, the number of languages for which Wikipedia has an article, the number of edits to the article, and whether the article is covered by a Wikiproject and, if so, what its importance rating is.*

**Process**