Adding Custom Dictionaries

Custom dictionaries can be added to Spell Checker Pro.

Configuring the Custom Dictionary Feature

Additional configuration to your `application.conf` file is required. (Don't forget to restart the Java application server after updating the configuration.

Adding the `ephox.spelling.custom-dictionaries-path` element activates the custom dictionary feature. It points to a directory on the server's file system that will contain custom dictionary files and should not contain anything else. It is a good idea to store these files where the `application.conf` file lives, i.e. if `application.conf` is in a directory called `/opt/ephox`, the dictionary files could live in a sub-directory `/opt/ephox/dictionaries`.

Example

```
ephox {
  spelling {
    custom-dictionaries-path = "/opt/ephox/dictionaries"
  }
}
```

Creating Custom Dictionary Files

One custom dictionary can be created for each language supported by the spell checker (see supported languages), as well as an additional "global" dictionary that contains words that are valid across all languages, such as trademarks.

A dictionary file for a particular language must be named with the language code of the language (see supported languages for language codes), plus the suffix `.txt`: E.g. `en.txt, en_gb.txt, fr.txt, de.txt` etc.

The *"global"* dictionary file for language-independent words must be called *"global.txt"*.

The server will scan the dictionary directory as per configuration above and pick up *"txt"*-files for each language and the global file as present.

Custom Dictionary File Format

A dictionary file must be a simple text file with:

* one word on each line,
* either Windows-style or Linux-style line endings (CR or CR+LF)
* no comments or blank lines, and
* saved in UTF-8 encoding, with or without BOM (byte-order mark).

The last point is important for files created or edited on non-Linux (Windows or Mac) systems, as these will usually encode text files differently. However, Windows or Mac editors such as Windows Notepad can optionally save files in UTF-8 if asked to do so. Please check your editor of choice for this functionality. Failure to chose the correct encoding will result in problems with non-English letters such as umlauts and accents.

NOTE for German and Finnish languages: Spell checking in German and Finnish will employ compound word spell checking. Compound words such as "Fußballtennis" will be assumed correct as long as the root words "Fußball" and "Tennis" are individually present in the dictionary. It is not necessary to add "Fußballtennis" separately.

Verifying Custom Dictionary Functionality

If successfully configured, the custom dictionary feature will report dictionaries found in the application server's log at service startup.

Example

```
2017-06-12 17:46:00 [main] INFO  com.ephox.ironbark.IronbarkBoot - Starting task (booting Ironbark)
2017-06-12 17:46:00 [main] INFO  com.ephox.ironbark.IronbarkBoot - using custom dictionary: [global] = 1 words
2017-06-12 17:46:00 [main] INFO  com.ephox.ironbark.IronbarkBoot - using custom dictionary: "en" = 3 words
2017-06-12 17:46:00 [main] INFO  com.ephox.ironbark.IronbarkBoot - using custom dictionary: "fr" = 2 words
2017-06-12 17:46:01 [main] INFO  com.ephox.ironbark.IronbarkBoot - Finished task (booting Ironbark)
```

The above log shows that 3 custom dictionaries were found, one "global", language-independent one and one each for English and French. They were found to contain 1, 3 and 2 words, respectively. Please check that this report matches your expectations.

Ongoing Dictionary Maintenance

Future additions/changes to dictionaries after the initial deployment will require a restart of the spell check service each time.