User Manual for the InifiniTag Document Tagging Service

Barnhill, Hanif, Vinokour, Wich, Sakib

June 2020

Contents

1 Introduction **2**

2 Document Overview / Start Page **2**

3 Adding Documents **2**

4 Maintaining Keywords **2**

5 Further Documentation **4**

6 Utilities **4**

7 License **4**

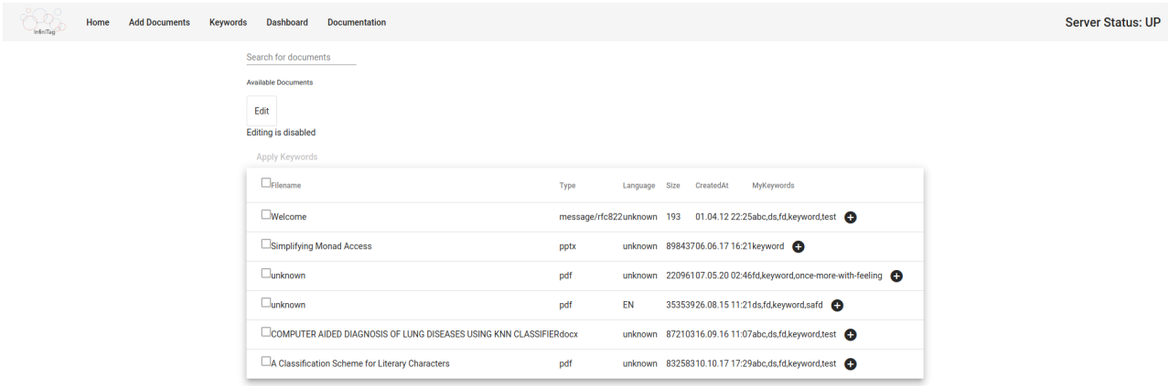

Figure 1: The Start page

1 Introduction

This documentation is aimed at the user attempting to use the InfiniTag software. The prerequesites for this documentation are:

1. The Solr database has already been setup
2. The back end has been installed with the requirements
3. The front end has been setup and is running.

For instructions on getting the front and back ends set up, see the README in the repository.

For help getting the Solr database up and running. See the various files in our Solr documentation.

2 Document Overview / Start Page

Once the front end is up and running, navigate to the start page to see a table with a list of documents as seen in Figure 1.

Notice in the top right corner is a status badge indicating the connection to the back end. If this is 'DOWN', then you will not be able to send or receive information from the server and will not be able to continue.

The search bar on this page enables you to filter documents by any criteria. Any word entered into the search box will be used to check for matching words in the name, tags, language, etc.

The + icon in the keywords column allows you to add any existing tag to a document.

Clicking on the edit button will enable the checkboxes to the right of the document name. By doing so you can apply one keyword to multiple documents.

3 Adding Documents

Navigating to `\documents` in the browser or clicking on the `Add Documents` link in the menu will take the user to a form where they can select one or more documents to be uploaded. This can be seen in Figure 2.

Upon selecting one or more documents all of the documents will be sent to the server for further processing. At the moment the documents will be stored in the database and the meta-information such as name, language, type, etc. will be extracted and stored along with the document content.

4 Maintaining Keywords

Navigating to `keywords` or clicking on the `Keywords` link will take the user to a page where they can add dimensions and keywords to a keyword model as seen in Figure 3. These keywords are then available to be assigned to the documents on the start / document overview page.

Here one can add dimensions as well as keywords, which will enable the construction of hierarchical keyword models.

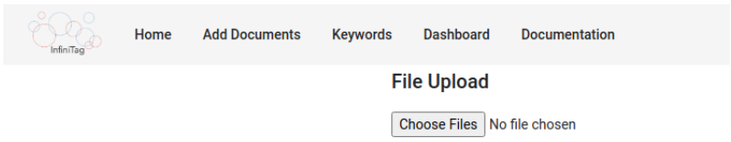

Figure 2: Uploading Documents

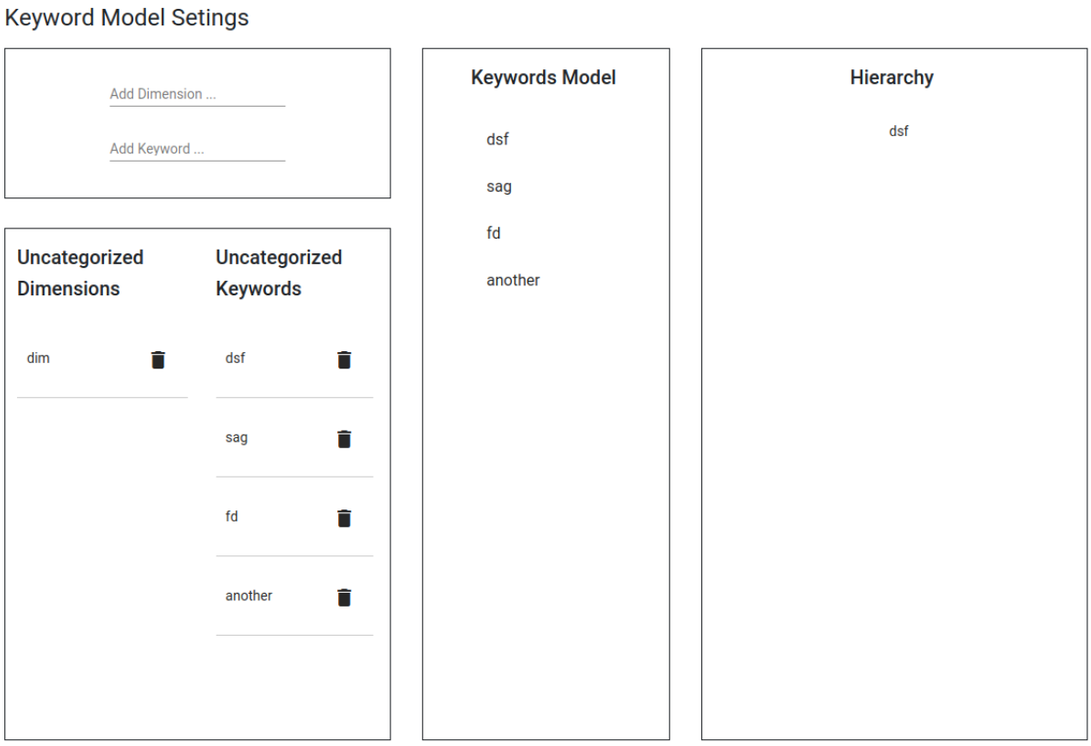

Figure 3: Maintaining Keywords

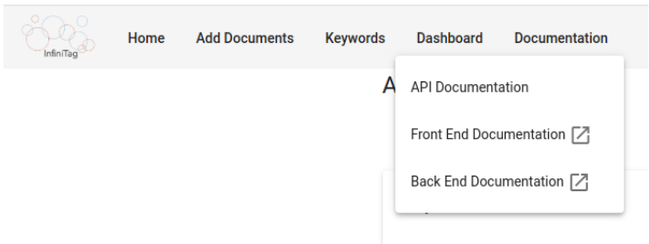

Figure 4: Documentation Drop Down

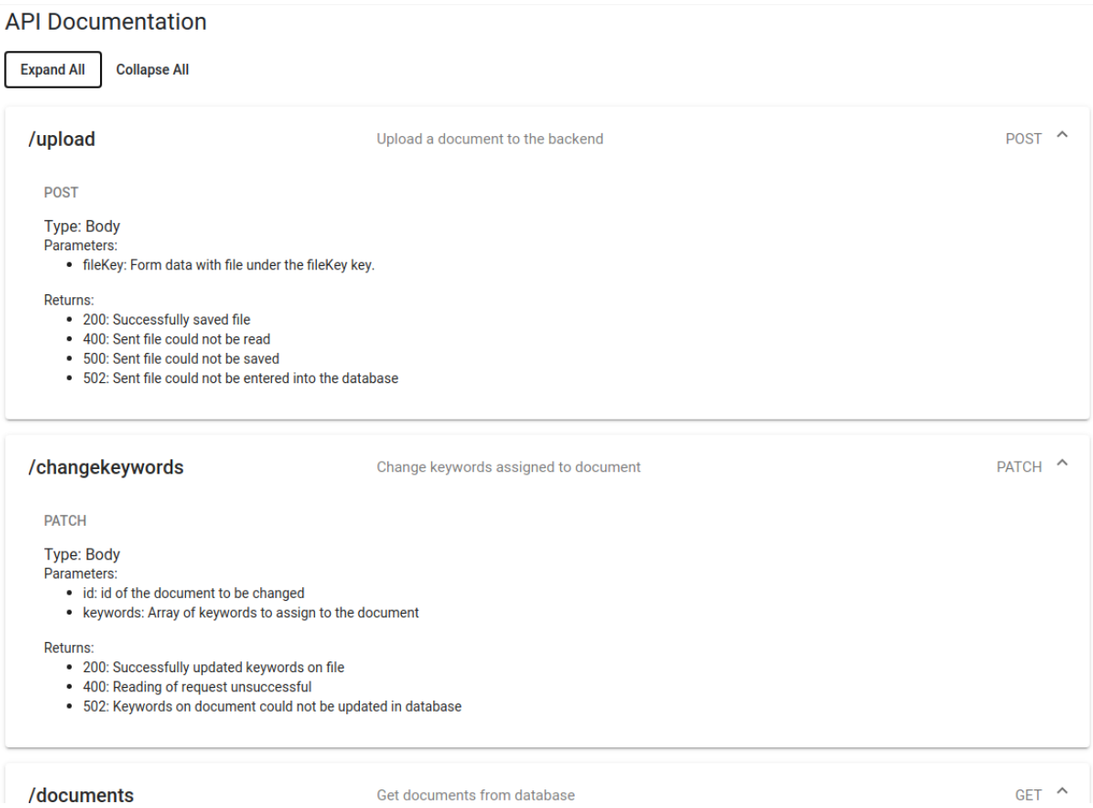

Figure 5: API Documentation

5 Further Documentation

Clicking on the `Documentation` link will provide the user with three options as seen in Figure 4:

1. API Documentation. This provides the user with an overview of all API endpoints for the back end, along with the required parameters and expected return results. A partial example can be seen in Figure 5
2. Front End Documentation. This is an external link which will show the user the actual code documentation for the front end.
3. Back End Documentation. This is another external link which will show the code documentation for the server.

6 Utilities

Important to notice is that the InfiniTag software comes with a utils folder. These are, in part, helper functions for the developers. However much more importantly are the scripts for categorizing and clustering documents. These scripts provide functionality for k-means clustering, hierarchical clustering, LDA, and TF-IDF. These are experimental features which will be integrated into the core solution as soon as possible.

7 License

InfiniTag Copyright © 2020 AMOS-5 Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions: The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software. THE SOFTWARE IS PROVIDED "AS

IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.