

- [Doccano Installation and Usage Instructions](#)
 - [Installation](#)
 - [Usage](#)
 - [Annotation](#)

Doccano Installation and Usage Instructions

Installation

To Install Doccano, follow the instructions listed here:

<https://github.com/doccano/doccano>

The easiest one is to just install through pip: `pip install doccano`

After installation, run the following commands to initialize the doccano system:

```
# Initialize database.
doccano init
# Create a super user with username admin and password pass
doccano createuser --username admin --password pass
```

Then, start the doccano webserver with:

```
# Start a web server.
doccano webserver --port 8000
```

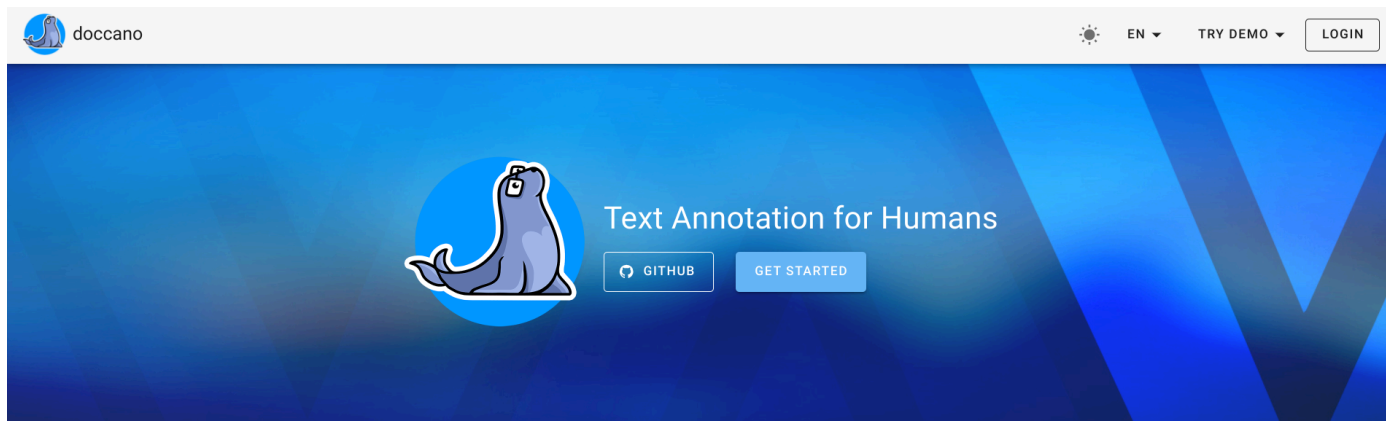
And run this in a separate terminal:

```
# Start the task queue to handle file upload/download.
doccano task
```

Now Doccano is running on <http://127.0.0.1:8000/>.

Usage

On `http://127.0.0.1:8000/`, you will see this screen:

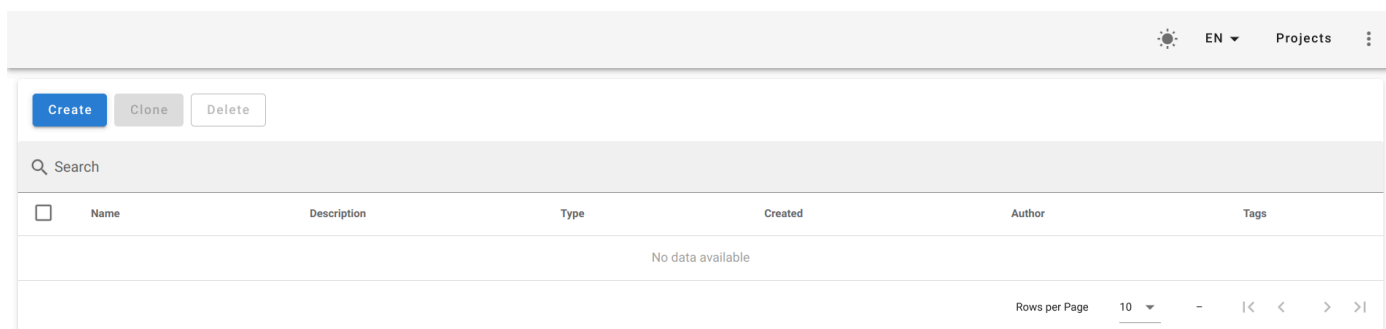


The best features

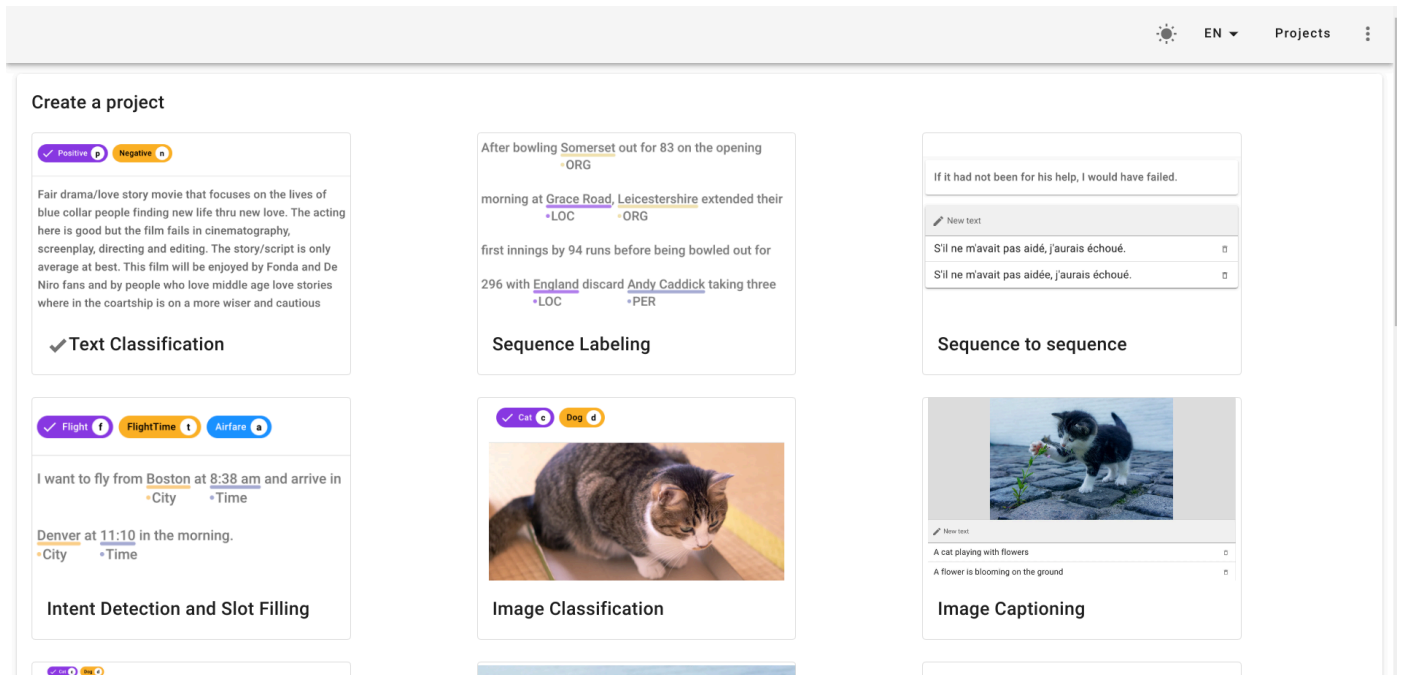


To annotate the course documents, first login using the account you created when installing doccano (the command above created an account with username `admin` and password `pass`).

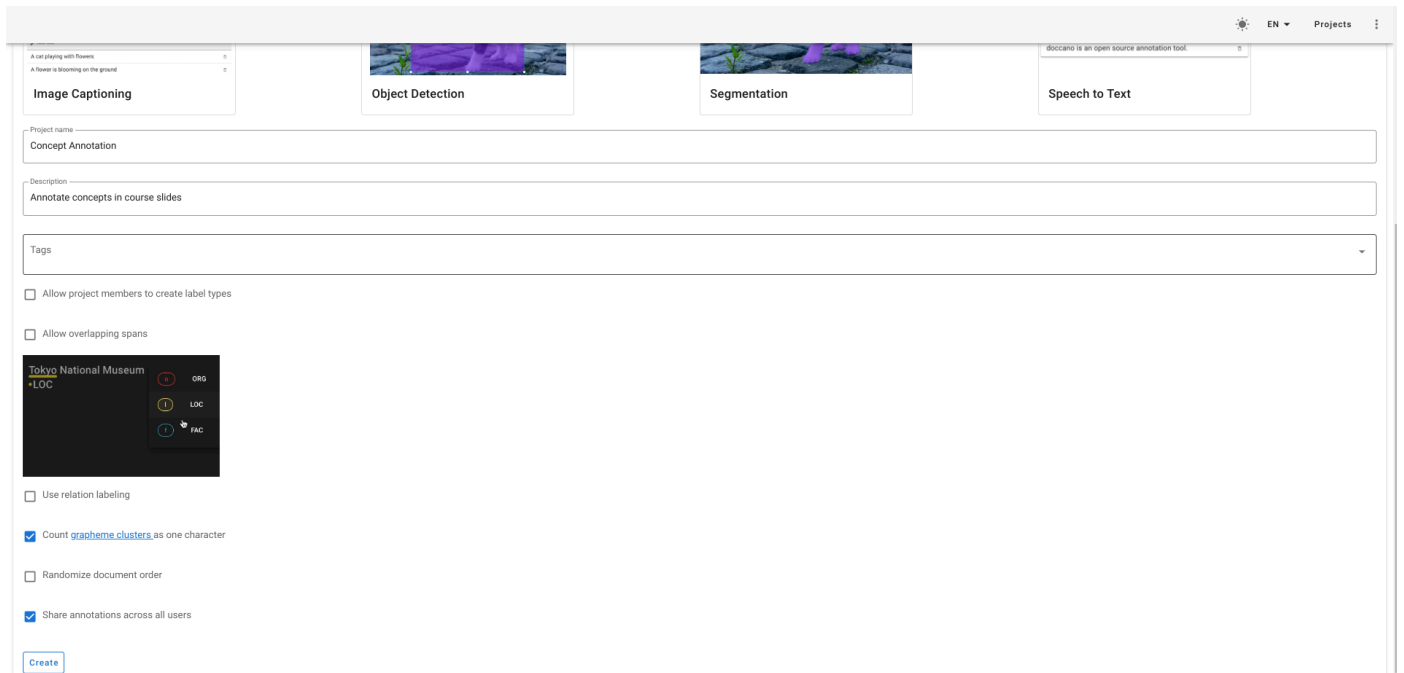
Then create a new project by hitting the create button:



Concept Labeling is a Sequence Labeling task, so select Sequence Labeling.



Scroll down and give the project a name and ensure that the checkboxes are exactly as in this picture.



And then hit create. You should see the screen below. First, go over to the labels tab on the left.

☰

Concept Annotation

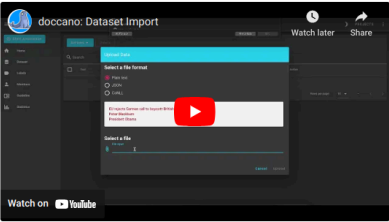
🌞 EN ▾ Projects ⋮

Start Annotation

HomeDatasetLabelsMembersCommentsGuidelineMetricsSettings

Welcome to Doccano!

1 Import a dataset



CONTINUE

CANCEL

2 Create labels for this project

3 Add members for collaborative work

4 Define a guideline for the work

5 Annotate the dataset

6 View statistics

7 Export the dataset

☰

Concept Annotation

🌞 EN ▾ Projects ⋮

Start Annotation

HomeDatasetLabelsMembersCommentsGuidelineMetricsSettings

Actions ▾Delete

Create Label

Import Labels

Export Labels

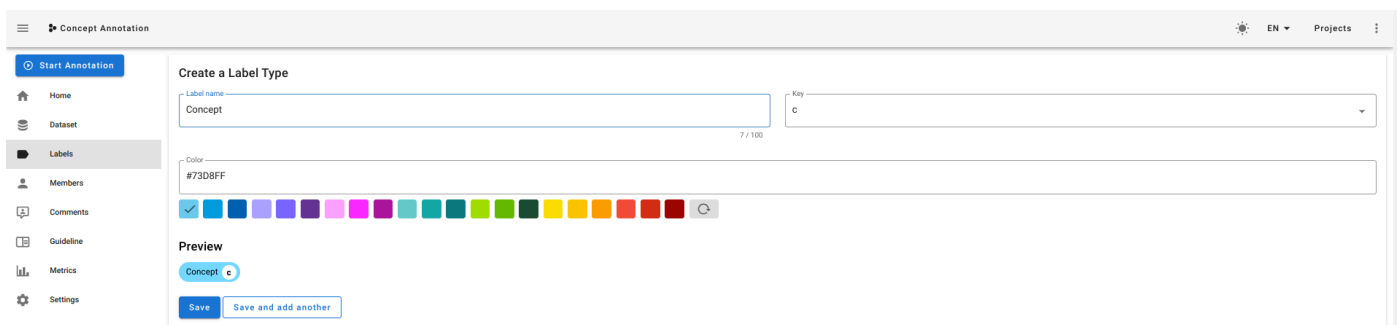
Shortcut	Color	Actions
No data available		

Rows per Page 10 ▾ - |< < > >|

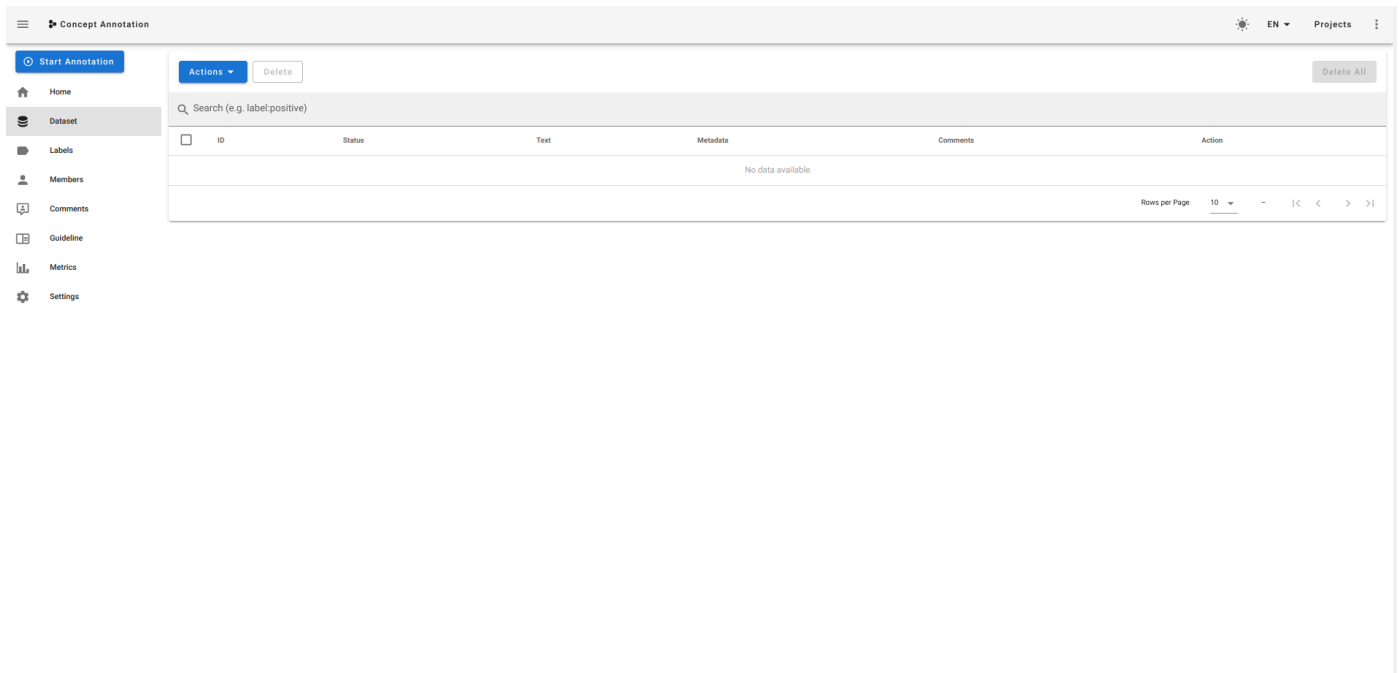
Hit actions, then import labels.



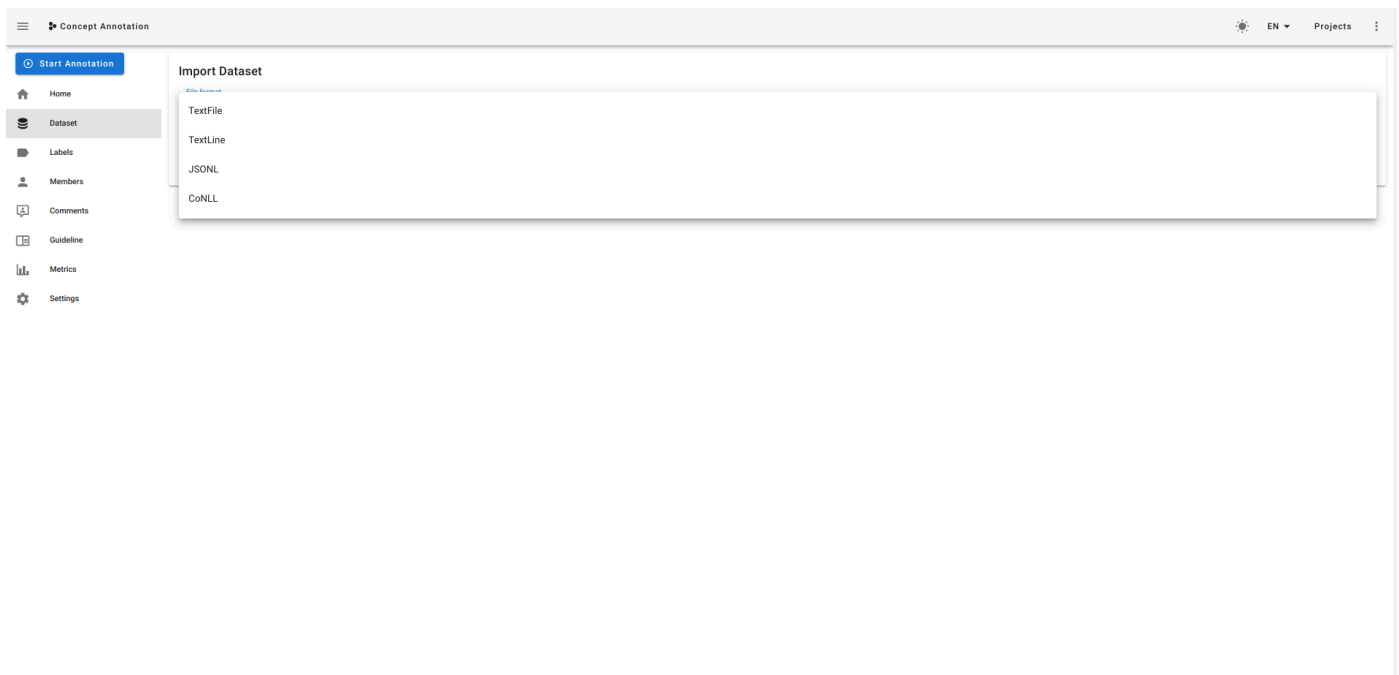
Click File input and select the `label_config.json` file provided. Then hit import. Alternatively, copy the label below exactly and hit save.



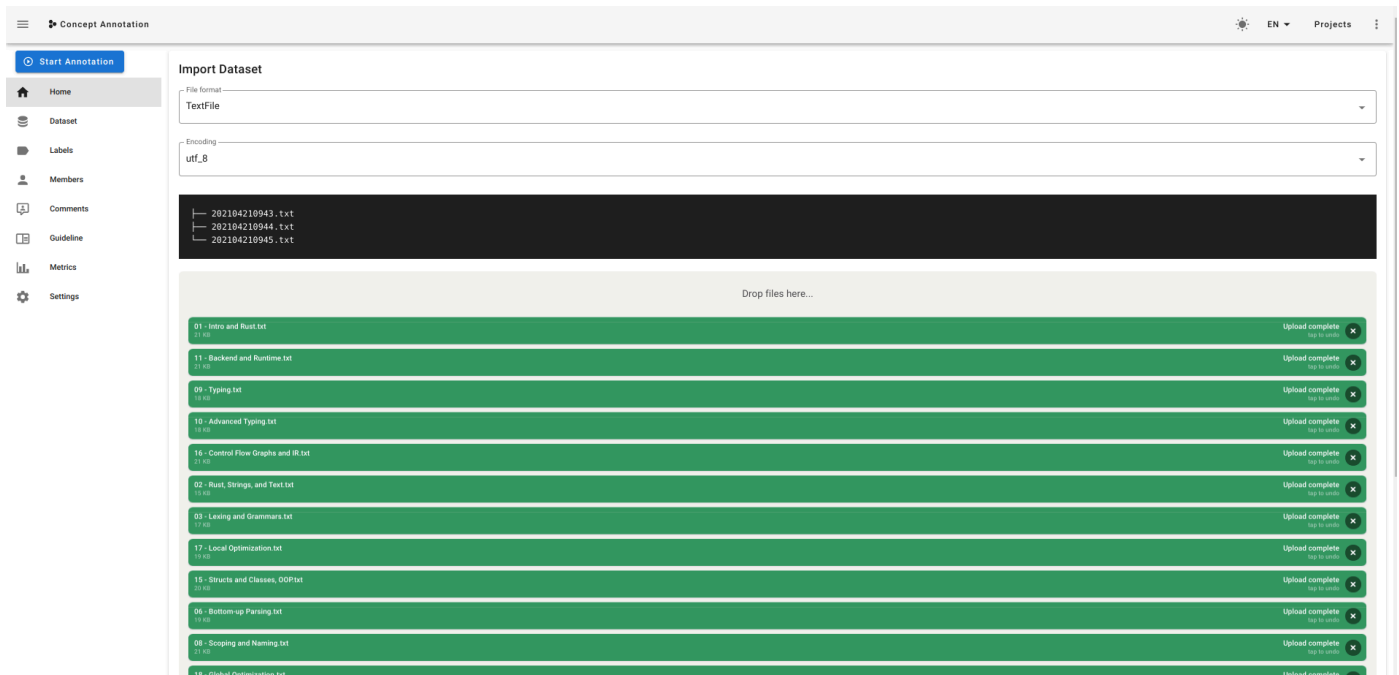
Now, go to dataset. Use the Actions drop down menu and hit Import Dataset.



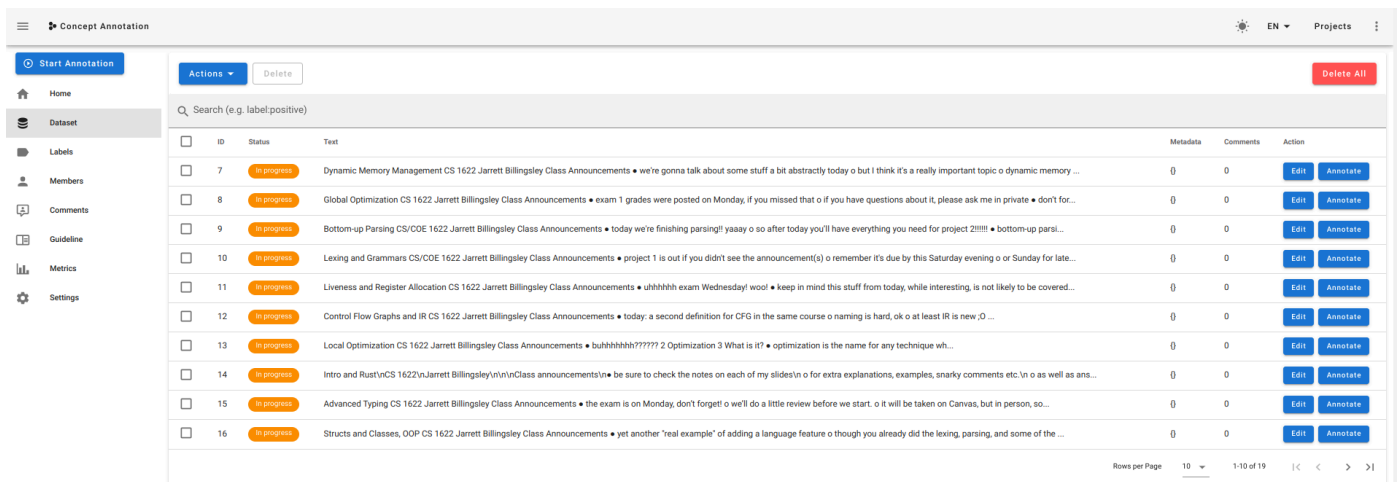
Click File format, and then TextFile.



And then click Drop Files here to initiate a popup window to select the presentation text files you want to be annotated.



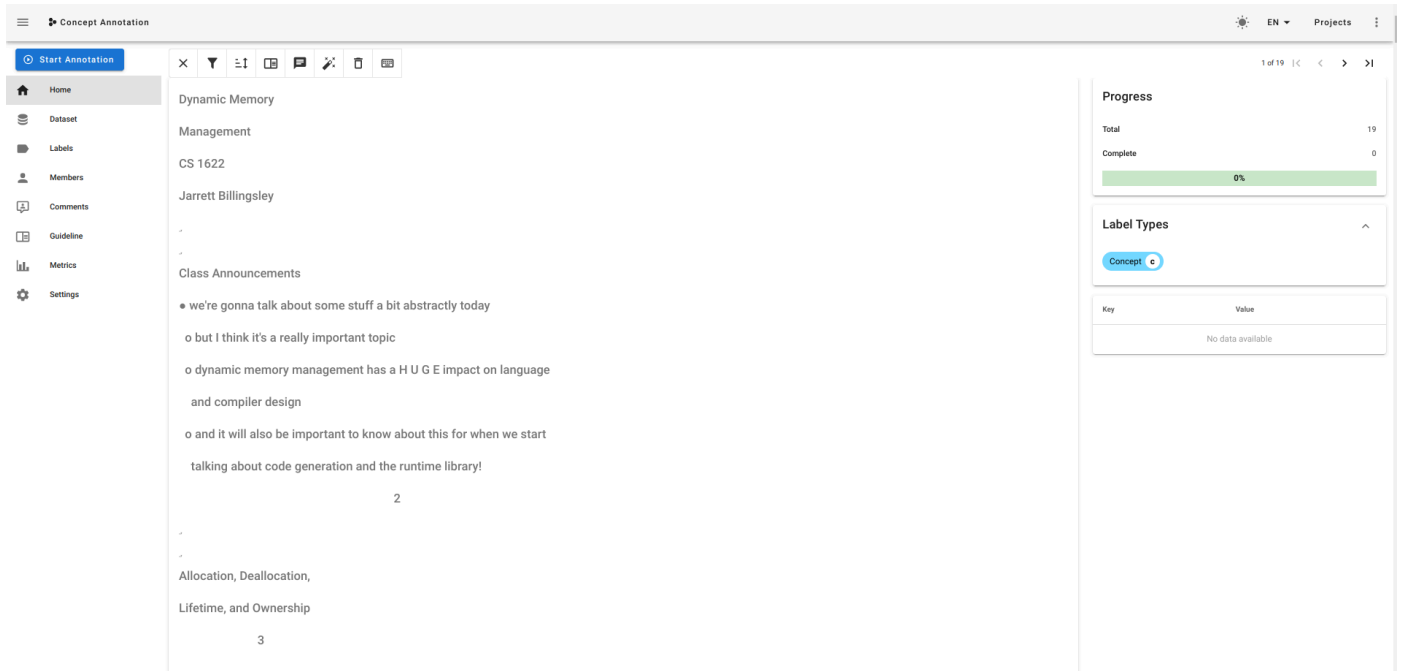
Wait for all of the uploads to complete, and then scroll down and hit the import button. You will be transported back to the Dataset tab.



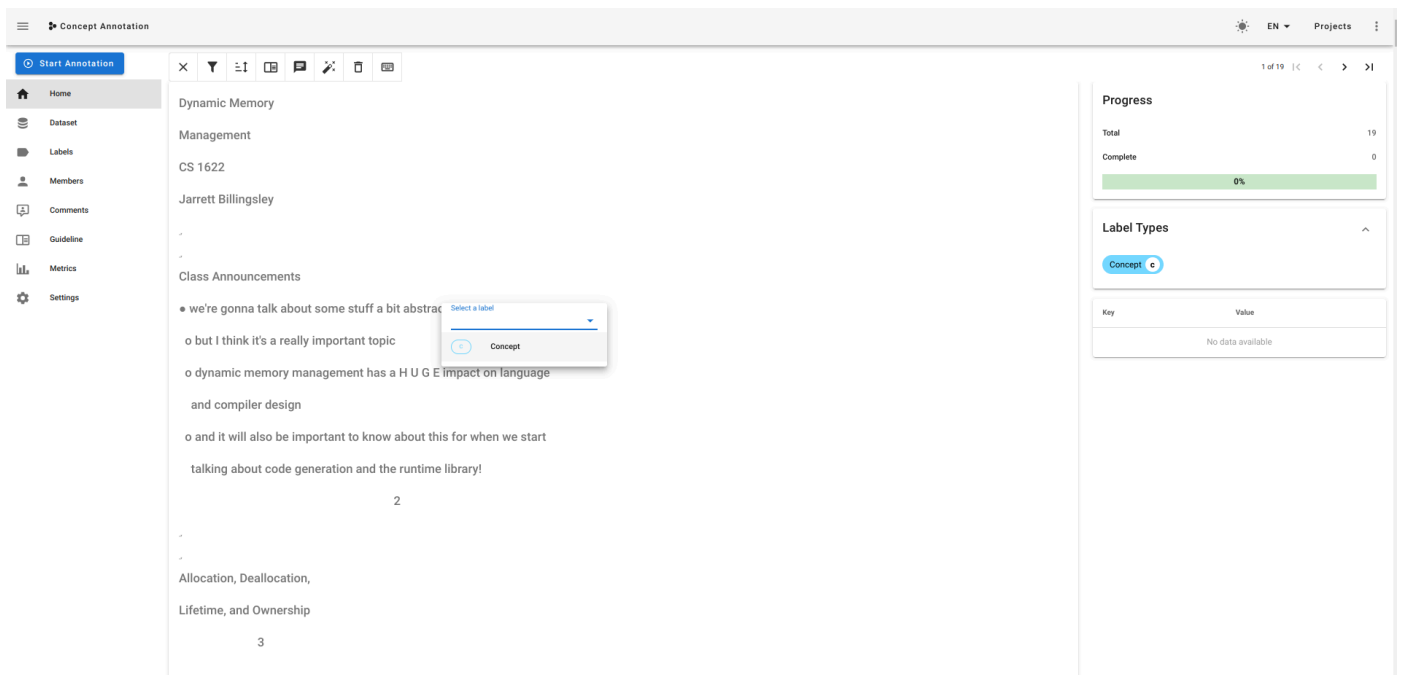
The files are ready for annotation.

Annotation

On the file you wish to annotate, click annotate.



To mark a portion of text as a concept, you can either highlight it and click the concept label on the dropdown



Or press c and then highlight:

Start Annotation

HomeDatasetLabelsMembersCommentsGuidelineMetricsSettings

Dynamic Memory Management CS 1622 Jarrett Billingsley

Class Announcements

we're gonna talk about some stuff a bit abstractly today

Concept

o but I think it's a really important topic

Concept

o dynamic memory management has a H U G E impact on language

Concept

and compiler design

o and it will also be important to know about this for when we start talking about code generation and the runtime library!

2

Allocation, Deallocation, Lifetime, and Ownership

Progress

Total19Complete00%

Label Types

Concept

Key	Value
No data available	

When you are done annotation, hit datasets, and then go to **Actions > Export Dataset**

Start Annotation

HomeDatasetLabelsMembersCommentsGuidelineMetricsSettings

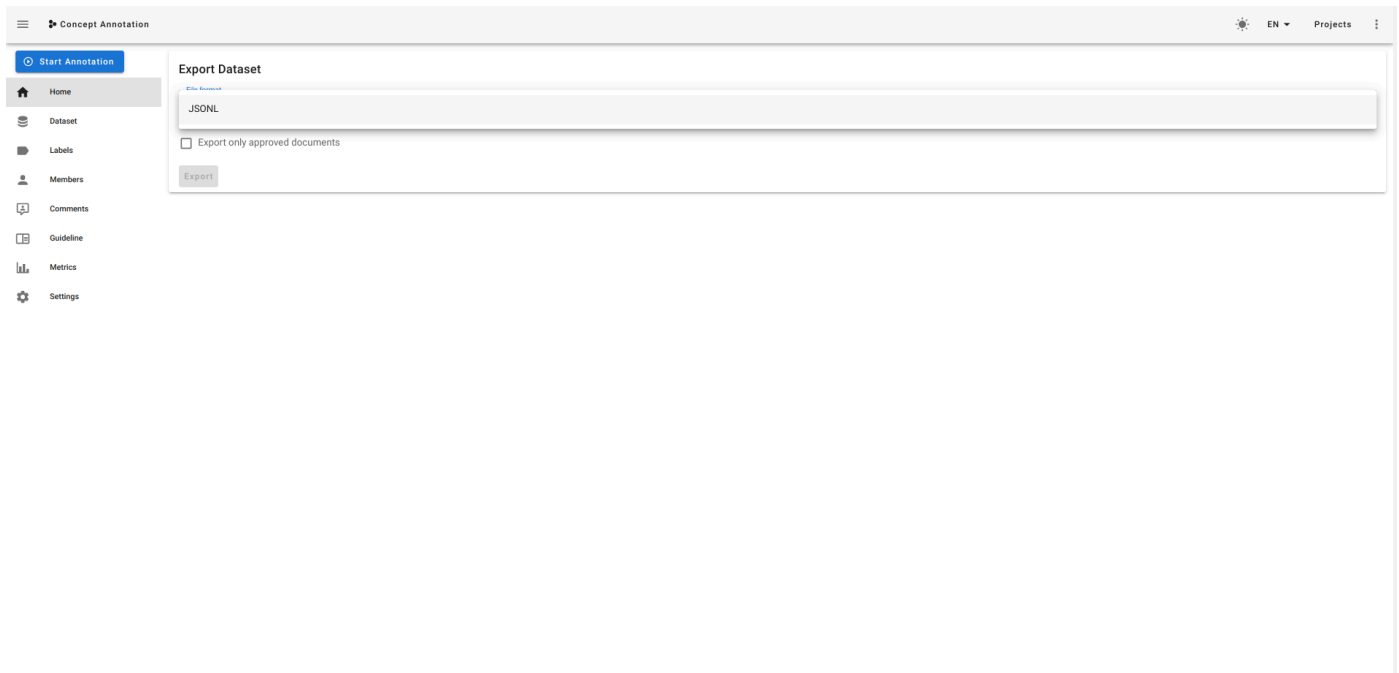
ActionsDelete

Import DatasetExport Dataset

	Text	Metadata	Comments	Action
<input type="checkbox"/>	7 Finished Dynamic Memory Management CS 1622 Jarrett Billingsley Class Announcements • we're gonna talk about some stuff a bit abstractly today o but I think it's a really important topic o dynamic memory ...	0	0	Edit Annotate
<input type="checkbox"/>	8 In progress Global Optimization CS 1622 Jarrett Billingsley Class Announcements • exam 1 grades were posted on Monday, if you missed that o if you have questions about it, please ask me in private • don't for...	0	0	Edit Annotate
<input type="checkbox"/>	9 In progress Bottom-up Parsing CS/COE 1622 Jarrett Billingsley Class Announcements • today we're finishing parsing! yaay o so after today you'll have everything you need for project 2!!!! • bottom-up parsi...	0	0	Edit Annotate
<input type="checkbox"/>	10 In progress Lexing and Grammars CS/COE 1622 Jarrett Billingsley Class Announcements • project 1 is out if you didn't see the announcement(s) o remember it's due by this Saturday evening o or Sunday for late...	0	0	Edit Annotate
<input type="checkbox"/>	11 In progress Liveness and Register Allocation CS 1622 Jarrett Billingsley Class Announcements • uhhhhhh exam Wednesday! wool • keep in mind this stuff from today, while interesting, is not likely to be covered...	0	0	Edit Annotate
<input type="checkbox"/>	12 In progress Control Flow Graphs and IR CS 1622 Jarrett Billingsley Class Announcements • today: a second definition for CFG in the same course o naming is hard, ok o at least IR is new :D ...	0	0	Edit Annotate
<input type="checkbox"/>	13 In progress Local Optimization CS 1622 Jarrett Billingsley Class Announcements • buhhhhhhhh????? 2 Optimization 3 What is it? • optimization is the name for any technique wh...	0	0	Edit Annotate
<input type="checkbox"/>	14 In progress Intro and Rust/vCS 1622/vn.Jarrett Billingsley/vn/vnClass announcements/vn• be sure to check the notes on each of my slides/vn o for extra explanations, examples, snarky comments etc./vn o as well as ans...	0	0	Edit Annotate
<input type="checkbox"/>	15 In progress Advanced Typing CS 1622 Jarrett Billingsley Class Announcements • the exam is on Monday, don't forget! o we'll do a little review before we start. o it will be taken on Canvas, but in person, so...	0	0	Edit Annotate
<input type="checkbox"/>	16 In progress Structs and Classes, OOP CS 1622 Jarrett Billingsley Class Announcements • yet another 'real example' of adding a language feature o though you already did the lexing, parsing, and some of the ...	0	0	Edit Annotate

Rows per Page101-10 of 19

And then select the JSONL file format to export.



Then hit export. You will get a zip file containing a JSON file with all of the annotations called **all.json**.

This format will then be converted into the format we will use to train the model.