

Annotation guideline of normalizing coping strategies

Version: 3, date: 15-6-2021

What is the task you will be doing?

You will be working on the normalization of coping strategies. First of all, normalization is the mapping of phrases in text to concepts in a standardized ontology [1]. This will allow us to count how many times a certain concept appears even if people talk about the same concept in multiple ways. Let me give an example: people could talk about their headache by saying 'their head is bursting', 'pounding headache' or 'feels like being hit with a hammer on my forehead'. They are all referring to the same concept but in different ways.

With coping strategies here we mean strategies that patients are recommending to each other to deal with side effects. In this annotation task, you will not have to worry about deciding if something is a coping strategy. They have already been labelled for you.

General recommendations

Generally, when annotating it is good to plan a block of time (e.g. minimal 1 hour) to be more efficient. From experience, it takes some time to get into the annotation again and therefore we do not advise you to work in smaller blocks of time. However, also do not continue annotating when you are tired or after a long day of hard work. You need to be fresh for annotation and on the long run, this saves time.

How do I get started?

First let's get you logged into the annotation tool.

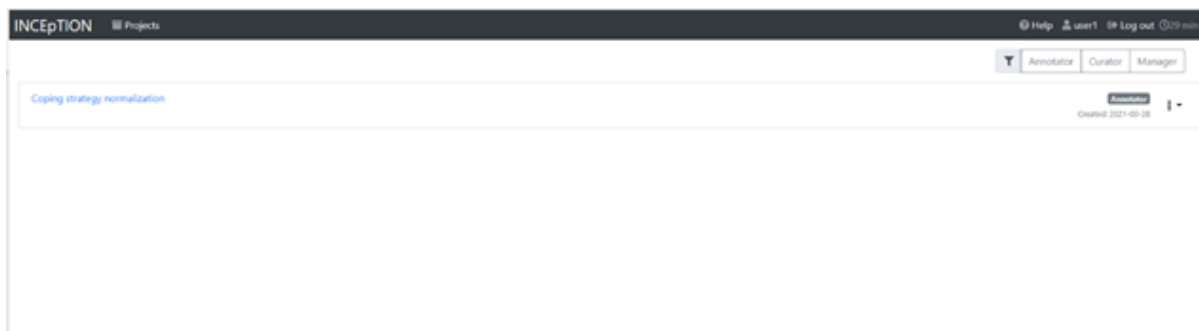
This link should get you to the log in screen

<http://textdata.nl:8080/login.html#!d=153&f=1>

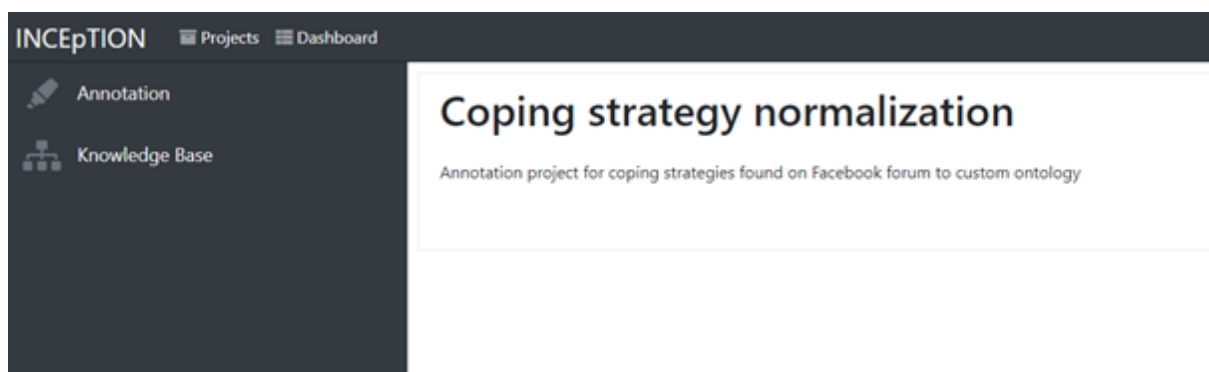
where you can log in with the user name and password you have received.

Note: If you are not active for 30 minutes you will be logged out automatically and will need to log back in but the program automatically stored what you have already done, so no worries.

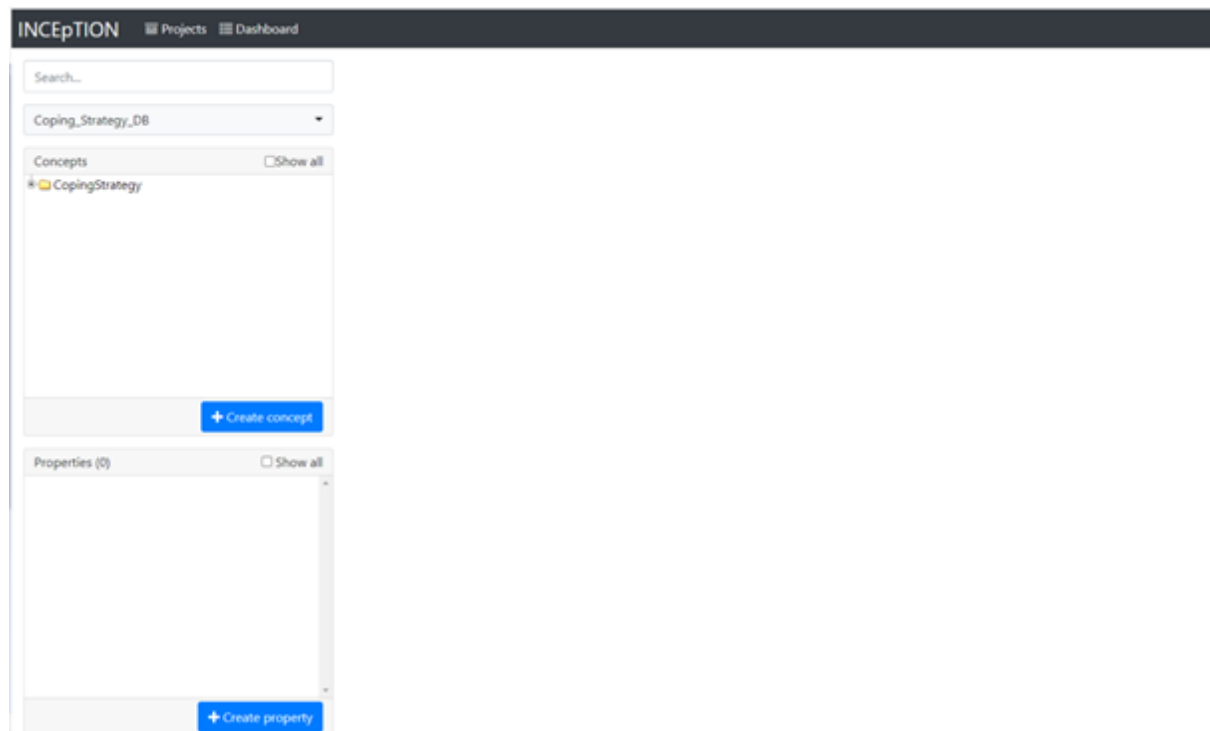
You will get a screen like this



Please click on Coping Strategy normalization (the project you have been assigned to)



You should see the above. Before you get started with annotating the data, you should take some time to explore the ontology or Knowledge Base. Click on **Knowledge Base**. This should bring you to this screen:



You see the top level concept Coping Strategy. Clicking on the plus will open the 'folder' and you can see which concepts fall under Coping Strategy. There are six categories:

1. **Alternative Therapies** : This category includes many non-medical therapies like mindfulness or hypnotherapy
2. **Eating and drinking** : This category includes any food and drink items you might consume and frequency and size of meals you consume
3. **Lifestyle**: This category includes sports and rest
4. **Medication and Supplements**: This category is very large and includes all possible ingredients that medication can contain but also includes supplements like Vitamin B12. You will see here that the top level concepts all start with **ATC**.
5. **Method of consumption drug**: This category is for any advice on how to best consume the drug so you get fewer side effects. This is divided into Dosage related advice and Temporal (or time) related advice or in other words, how to consume the drug and when to consume the drug.
6. **Physical Object**: This category contains aid devices like hearing ls, or glasses (visual aids), Personal care products like shampoo or toothpaste and clothing items.

There are also some loose concepts that do not fall under any of the above categories:

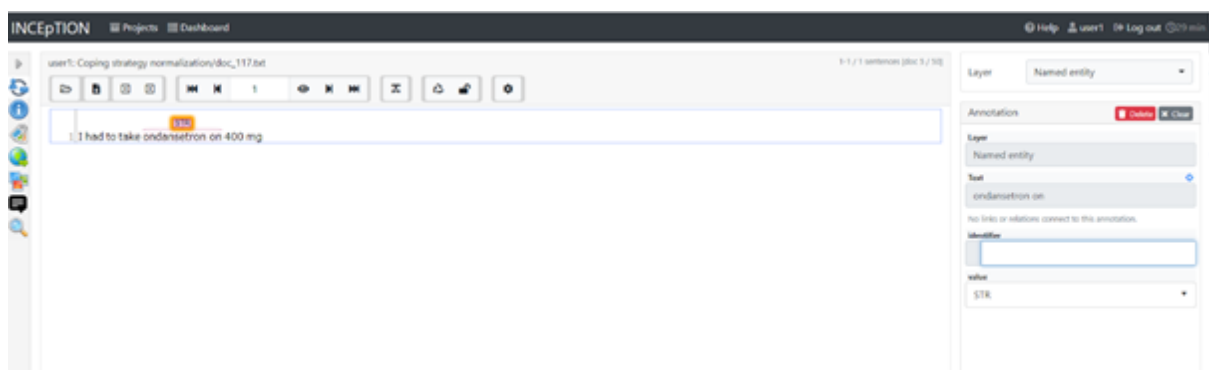
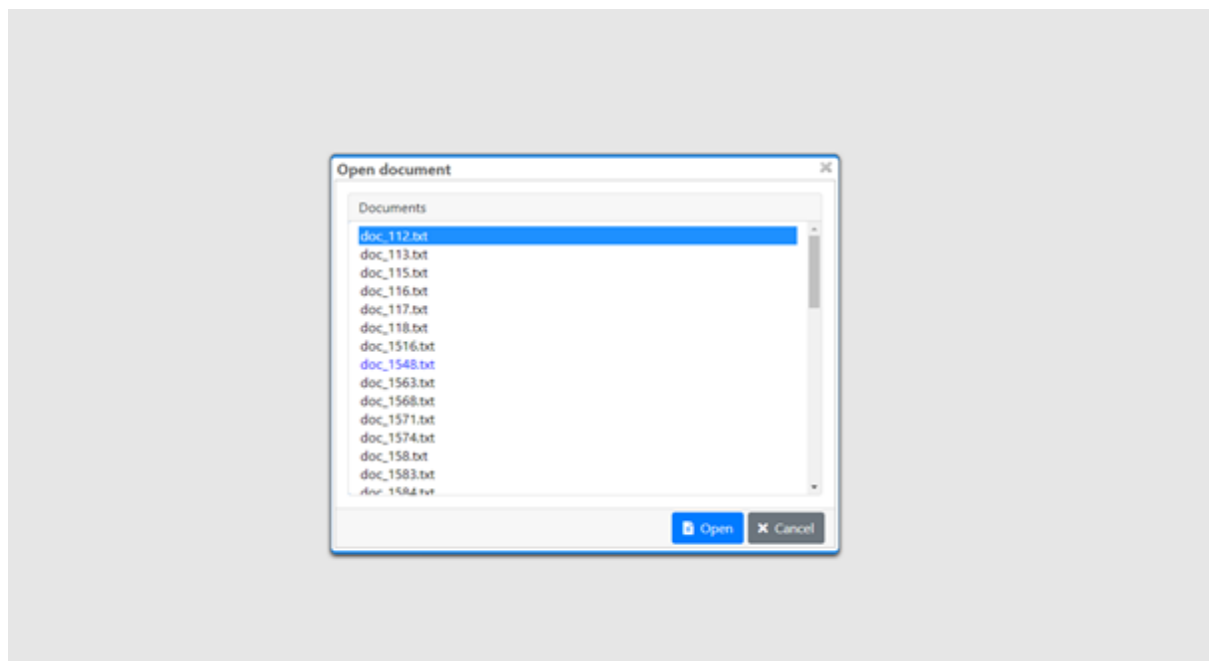
1. Acclimation (or letting your body get used to something)
2. Patience

3. Positive attitude

Have a scroll through the ontology and familiarize yourself with a rough idea of what is in there. You do not need to know all concepts. There will be a search bar available to you just like you see in the top left corner.

Starting the annotation

Click on the dashboard button on the top left of your screen to go back to the previous screen. Now click **Annotation**. This brings you to a screen with all the documents you need to annotate. Simply press **open** when a document is selected to start. The documents you have already done are blue and the ones you have not done yet are black.



This (above) is what the screen looks like when you have entered the annotation for a document.

You can see here that some words have already been highlighted in pink and labelled with STR. This stands for Coping STRategy. If you click on the STR the box on the right opens. The identifier box is where you will fill in the appropriate concept for this phrase or word. It functions like a search bar: if you start typing, the program will start searching in the ontology and you click on the search result that you think is correct.

Some cases may be very hard. Give it your best try.

When it concerns a general category of medication types such as nausea pill or vitamins, you will need the top-level ATC categories in the Medication category. Aside from the search bar, you can use a specific search bar for finding the right ATC category for a concept here:

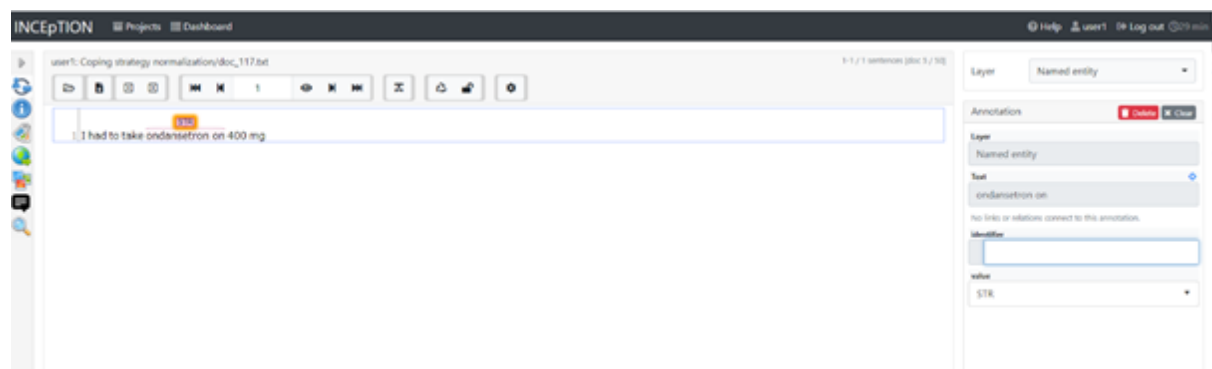
https://www.whocc.no/atc_ddd_index/

This search engine is better at working with synonyms than the search bar in the annotation tool.

What if there are multiple strategies?

If there are multiple strategies in a message, please annotate all of them.

Please make a note of any difficulties you run into to help us make the annotation guide better.



The green button is for when you have finished your annotation. **Please press this before going to the next document (!)** so that we can monitor who is finished with which documents. The orange button will bring you to the next document.

You will also encounter cases like the one below where there seem to be two coping strategies although it is highlighted as one phrase. Here 'take my imatinib a couple of hours

after I eat' and 'stay active' are two strategies which refer to different concepts in the ontology.

What if a strategy is not to DO something, but to NOT do it?

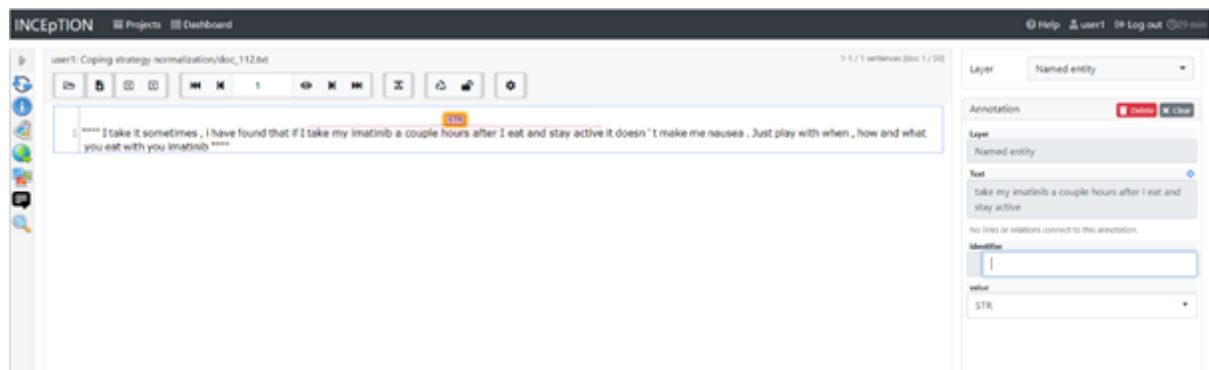
These should be labelled with STR-NEG instead of STR like the examples below. If this is not the case, please change the STR (in value) to STR-NEG

Also avoid hot water and direct sunlight .

1 Mine happened after over a year on my drug . I stopped wearing all of my jewelry and it is helped so much .

What if there are two separate pieces of advice in one entity?

This is problematic for labelling. The pieces of advice should have already been separated but if they are not you will need to fix this. They are often connected with an 'and' or an 'or' like in this example:

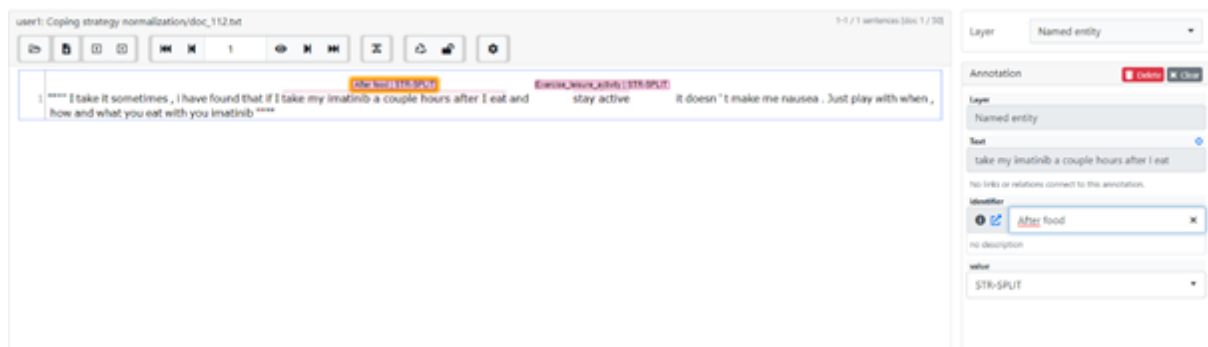


If they are not separated do the following:

1. Remember where the annotation starts and ends.
2. Press the STR and click delete (red button on the right hand side)
3. Now highlight the first piece of advice, here "take my imatinib a couple of hours after I eat"
4. Under value select STR

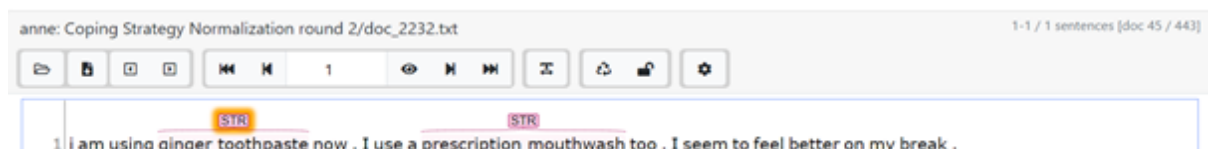
5. Now highlight the second piece of advice, here “stay active”
6. Under value select STR
7. Select the right identifiers for each strategy

It should now look like this:



What if a strategy needs multiple labels because the whole strategy is not in the ontology but separately the parts are?

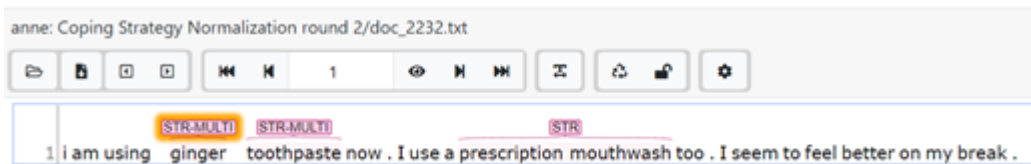
An example is 'ginger toothpaste' on the left in the example below.



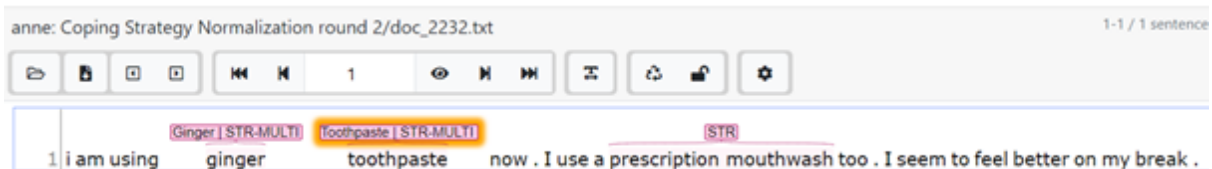
If you come across a case like you first need to split the annotation:

- Remember where the annotation starts and ends.
- Press the STR and click delete (red button on the right hand side)
- This deletes the annotation.
- Now highlight the first part of the strategy for example 'ginger'.
- Under value select STR-MULTI (for multilabel)
- Now highlight the second part 'toothpaste' and do the same thing.

It should now look like



For each part, add an identifier. It should now look like:



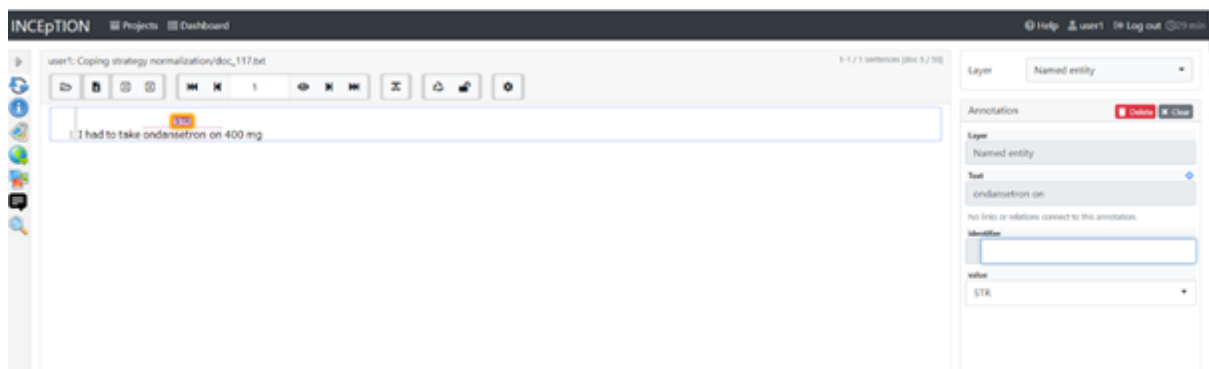
There is one exception: Indications of HOW OFTEN the strategy needs to happen like ‘daily’ or ‘monthly’. These do not need to be labelled separately and you do not need to split into a STR-MULTI for this. We will ignore them **unless it concerns when to take medication.**

What if the annotation is wrong like words are added that clearly do not belong to the strategy or words are missing?

An example is the screen below where ‘ondansetron on’ should be ‘ondansetron’. Please correct this.

You need to first delete the entity and make a new one. Now you can select STR under Value.

Please make a note of all the cases where you have changed which words belong to the strategy in a separate document.



What if I cannot find any suitable entry in the ontology?

If you have tried different keywords and cannot find any suitable entry, you can label the concept with a higher level category for example:

Maybe you come across a sport that is not in the ontology like rowing or boxing, then you can annotate it with the higher level category Exercise_Leisure_Activity.

When you label an entity with a higher level or less specific concept than it is in reality, please add the UPPER tag to the value. So for instance STR would change to STR-UPPER, STR-NEG to STR-NEG-UPPER etc.

If the concept you are labelling IS a less specific concept like 'do sports' then do not use the UPPER tag. The UPPER tag is to indicate that the exact concept is not in the ontology (not available) and so you use a more general concept.

If you also cannot find a suitable higher level category, then you can label the concept with the category UNKNOWN which has been added for this specific purpose. We will then discuss this concept in our team meetings to see if the ontology needs to be changed and to resolve these issues.

What if it is not a coping strategy for a side effect at all?

Then you can label it with a special label for these cases: NOT_A_STRATEGY.

What if I see coping strategies that have not been annotated?

Do not annotate these please. This is not the task we are working on.

Short summary of the rules:

- Please annotate all strategies in the message
- Do not annotate any additional strategies that are not already annotated (highlighted)
- You can allowed to correct which words belong to a strategy that is already annotated
- If there are two separate pieces of advice in one highlight, you can separate them and add STR (or STR-NEG) to both
- There are three tags that can be added to STR:
 - o NEG for strategies that are about NOT doing something

- UPPER for when the exact label for the strategy is not available (like basketball) and so you use a more general label (like Sports)
- MULTI for when you need to attach multiple labels to ONE strategy (like ginger toothpaste). Please ignore references to time periods (e.g. daily) unless it is about when to take medication
- If it is not a strategy, use NOT_A_STRATEGY
- IF there is no correct label and no more general label that applies to a concept, use UNKNOWN
- Press finish when done with a document

Logging your time

We have made an estimate of how long you will be spending on annotation and other tasks. However, we would like you to log the time you *actually* spend.

For any questions please e-mail me at a.r.dirkson@liacs.leidenuniv.nl

[1] a set of concepts and categories in a subject area or domain that shows their properties and the relations between them.

