**Download and prep files**

1. Go to project GitHub repo
2. Download the following files to a folder where you will be able to find them later, preferably with nothing else in it:

NLP\_final\_project\_v3.dtd

mae-2.2.4-fatjar.jar

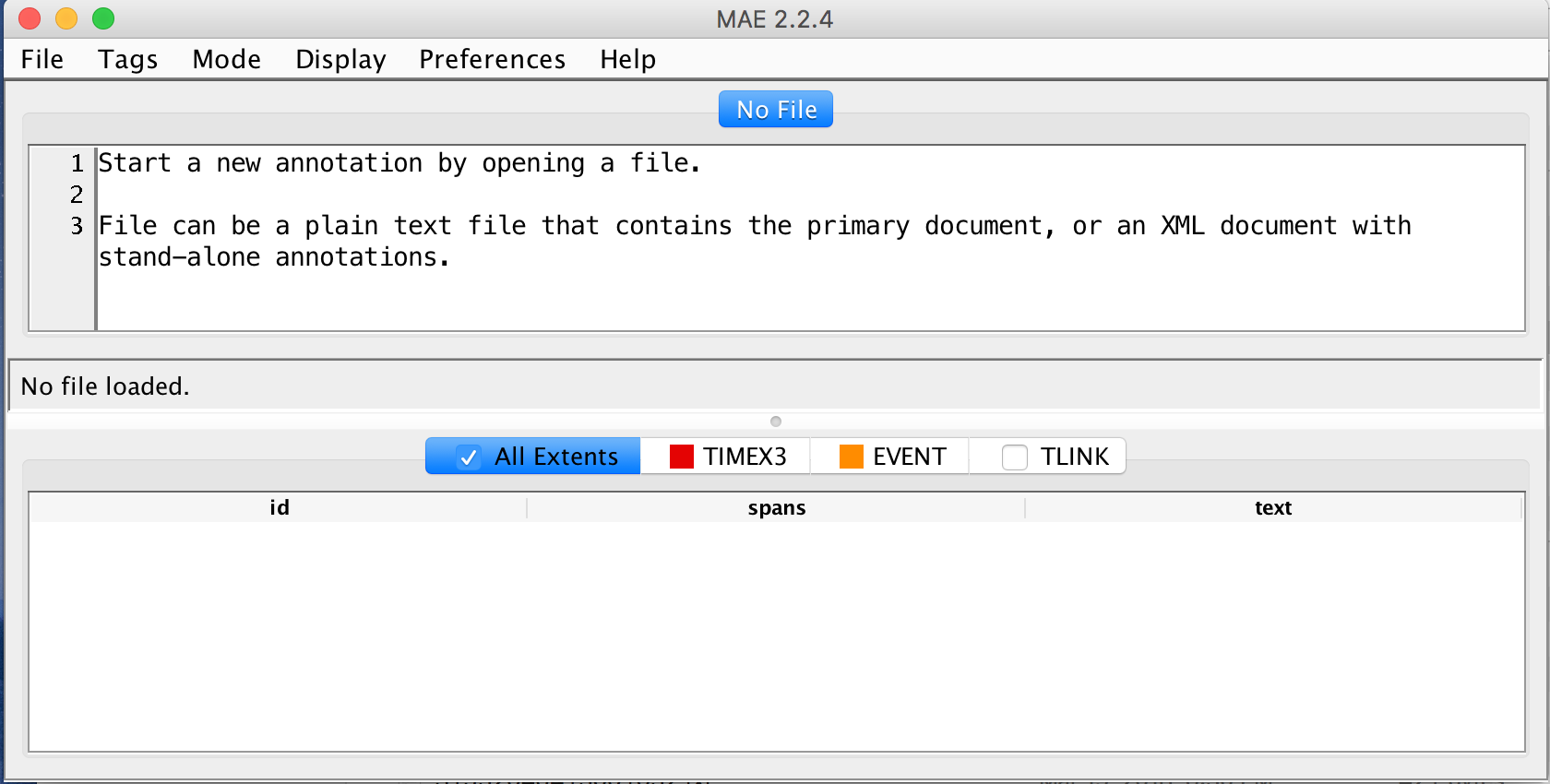
finalDCnotes.csv

1. Create a new directory on your computer called <initials>\_Annotation\_Output, i.e. “JF\_Annotation\_Output”. You will save your annotation files here.
2. Create individual text files for notes in the rows of finalDCnotes.csv you’ve been assigned. Naming convention is ds<row number>.txt, for example ds1.txt.

You can do this via Python, or by opening the CSV file and copying the note into a text editor, at your preference. I **highly suggest** the former, as it should preserve whitespace and special characters; the latter may run into character limit issues in Excel.

**Open MAE**

1. Double click on mae-2.2.4-fatjar.jar to open the Multi-purpose Annotation Editor (MAE). <https://keighrim.github.io/mae-annotation/> has more details on this open source application.



**Load NLP\_final\_project.dtd**

1. In MAE, go to File -> New Task Definition
2. Select the downloaded NLP\_final\_project.dtd file. Click Open.

**Open your assigned documents, one at a time**

1. In MAE, go to File -> Open Document. Select one of your text files to annotate and click Open.
2. Click OK in the window informing you that a new XML file will be generated (you can select “Do not show this again”.

**Annotate**

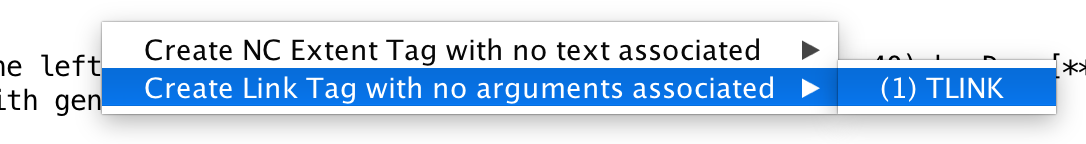
There are 3 types of annotations: TIMEX3 for dates, EVENT for events, and TLINK for the relationship between a date and an event. Focus on the relevant events (listed below) and **all explicitly stated dates** that occurred within the inpatient admission.

To annotate a span of text, highlight the minimum relevant text (for dates, this is JUST the date - not the surrounding formatting) and right click. Select Create Extent Tag with selected text -> TIMEX3 or EVENT. This should create a new row in the bottom window with an auto-generated id, the span of text highlighted, and the actual text highlighted.

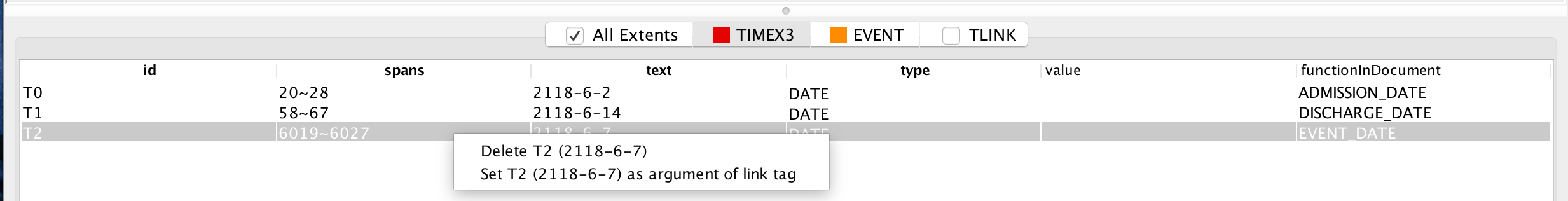
1. TIMEX3
   1. functionInDocument: leave as EVENT\_DATE **except** for the admission and discharge date statements at the top of each note and the signature and authored dates often at the bottom. Change to ADMISSION\_DATE, DISCHARGE\_DATE, or AUTHORED\_DATE as appropriate.
   2. Value: leave blank **unless** the date is incomplete and you need to add information, particularly to add the year. This is your interpretation of the full date value if it is not explicit in the text
2. EVENT
   1. eventType: we’re looking at 7 common events in the ICU
      1. Intubation/extubation
      2. DNR (do not resuscitate) start/stop
      3. DNI (do not intubate) start/stop
      4. Resuscitation / withdrawal of life support measures
      5. Feeding/OG (orogastric) / NG (nasogastric) tube placed/removed
      6. Transfer in to/out of the ICU
      7. Foley catheter placed/removed
   2. Class:
      1. Default: OCCURRENCE. Leave alone if the text is describing an occurrence of the event, that is the patient was intubated or was declared DNR.
      2. STATUS: if there is a dated declaration of a status but not when the event occurred, change this value to STATUS. For instance, if the note says “on 2118-2-1, since patient was still intubated” this is a STATUS rather than an OCCURRENCE.
      3. **Do not annotate statements that describe other things relative to these events**, e.g. “sore throat post extubation”. Do not annotate headers or field names, e.g. “Communication/code status: patient…”
   3. Polarity: all the eventTypes other than resuscitation cover the positive and negative instances, for instance intubation and extubation. Positive refers to a procedure being done (intubation, placement of feeding tube, making a patient DNR) or the more severe version (transfer to ICU). Negative refers to an intervention being stopped (extubation, removal of Foley catheter, reversing a Do Not intubate order) or the less severe version (transfer out of ICU to another unit/floor). Resuscitation, or the re-starting of someone’s heart and lungs after a code, has a positive polarity. Its negative polarity is the withdrawal of, or agreement to withhold, life support measures such as ventilation.
   4. Completion: this field is not mandatory, and is used to refers to events that were attempted but completed. For instance, if a note says “underwent a failed extubation”, this is an attempt. If it says “patient was extubated but then intubated 3 hours later”, this is a completed extubation and a completed intubation.
3. TLINK

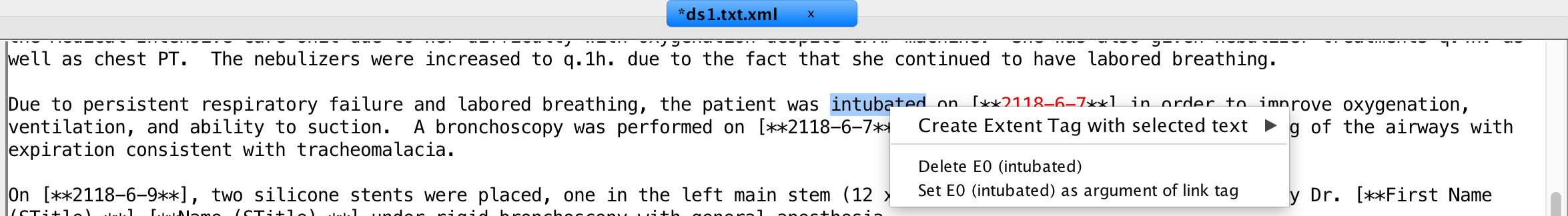
Once you have the TIMEX3 and EVENT annotations done, make a connection FROM the time TO the event. **Not all TIMEX and EVENT items will have a link** (e.g. if there was an event but no explicit date attached), and links may not occur in the same sentence.

1. Right click on the document with nothing selected and click Create Link Tag with no arguments associated -> TLINK

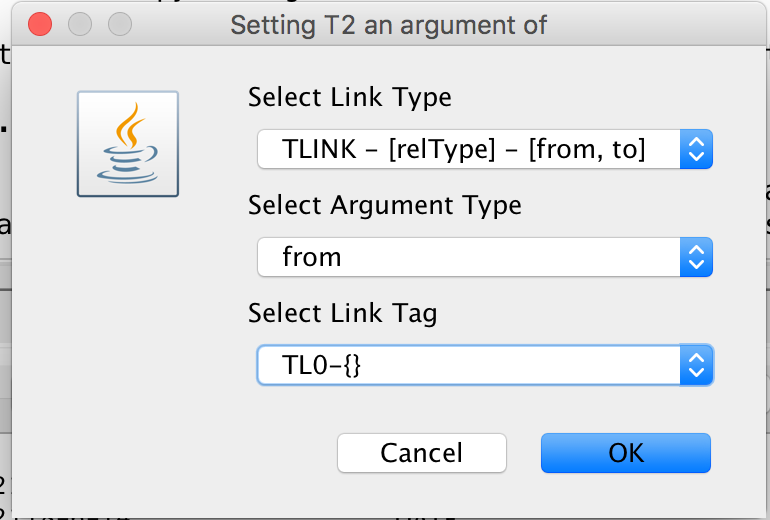


1. Right click on either the TIMEX3 event row in the bottom window, or the highlighted TIMEX3 annotation in the top window

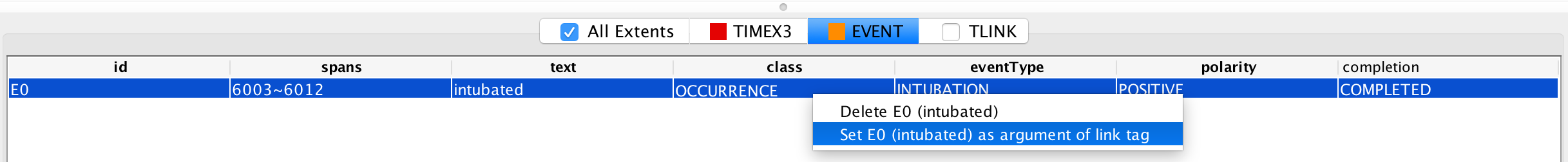


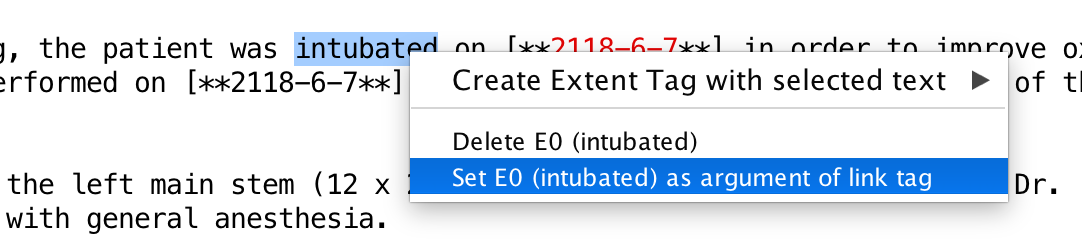


1. Click Set T<id> <value> as argument of link tag
2. In the pop up window, make sure the Select Argument Type is from and the Select Link Tag does not already have a T<id> TIMEX3 annotation. Click OK.

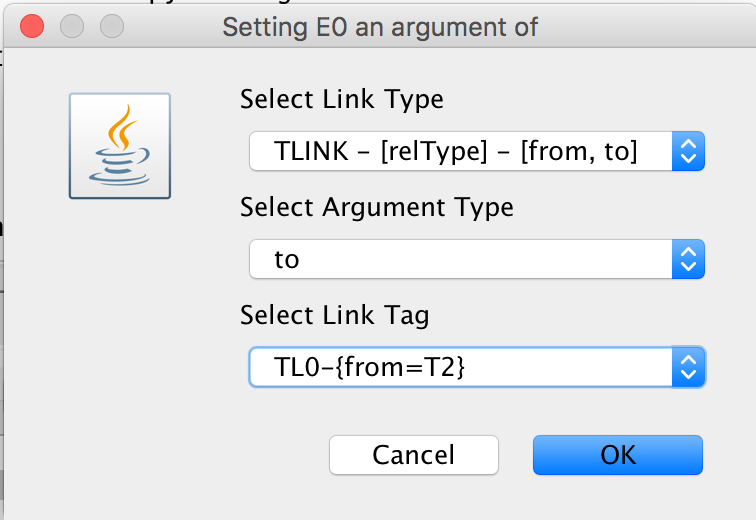


1. Right click on either the EVENT row in the bottom window, or the highlighted EVENT annotation in the top window

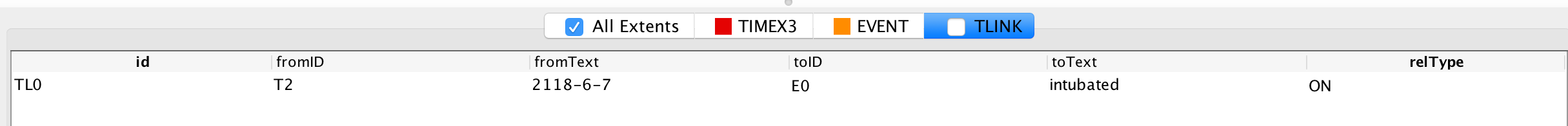




1. Click Set E<id> <value> as argument of link tag
2. In the pop up window, set Select Argument Type as to and make sure the Select Link Tag is the one the TIMEX3 was attached to: TL<id> -{from = T<id}. Click OK.



1. In the TLINK table, you should now see a TL<id> with the T<id> in the fromID field and its value in the fromText field, and the E<id> in the toID field and its value in the toText field.



1. Set the relType to ON.

4. Make sure to save your annotations as you’re working, and again once you’re finished with a note. The output should be a ds<id>.txt.xml file and should be in your <initials>\_Annotation\_Output directory.

**Upload your annotations**

Once you have a few notes done, upload your output directory to the GitHub (make sure you upload the entire directory, so your files are kept separate). Once your folder exists in the repo, you can just upload additional individual files into it as you finish annotating - please try to upload the output as soon as you finish each file, so the other team members have access to your work. There will be some notes that get annotated by multiple people, so make sure you save your output in your own folder!