

Annotation guidelines for symptom onset identification and temporal linking

Written by Natalia Viani and Sumithra Velupillai.

1. Onset extraction task

In this annotation task, we will consider only those patients who have been admitted to early intervention services.

For each patient, multiple documents related to these “early intervention” admissions are provided. Your task will be to identify symptom onset information in these documents. Specifically, we are interested in identifying the first presentation of symptoms related to psychosis.

Symptom onset information could be found in sentences (*“he first experienced hallucinations in 1999”*) as well as in longer text spans (e.g., a whole paragraph). It is important that you mark the entire textual portion that is needed to recover the onset information, i.e. the symptom and the associated date/time.

For each patient, it might happen that none of the provided documents include onset information. On the other hand, onset information could be found in more than one document. This annotation task is intended as a “document screening”, where we identify all the document types that are likely to contain the target information.

2. Temporal link task

In this annotation task, we will consider documents that have been pre-annotated with time expressions and events (symptoms).

In particular, the following symptoms have been considered:

1. hallucinations
2. delusions
3. delusional
4. thought disorder
5. paranoia
6. formal thought disorder
7. tangential
8. thought interference
9. flight of ideas
10. persecutory ideas
11. thought block
12. circumstantial
13. tactile hallucinations
14. loosening of associations
15. running commentary

16. floridly psychotic
17. somatic passivity
18. thought alienation
19. somatic hallucinations
20. derailment
21. tangential speech
22. circumstantial speech
23. auditory command hallucinations
24. systematised delusions
25. knights move thinking
26. clang associations

Note that:

- Time expressions are identified by an automatic tool, so there will be errors: you can correct them (adding missing expressions, or removing “wrong” ones)
- For symptoms, we are restricting our search to the list above. However, if you find variants of the same symptoms, you can mark them, too (e.g. “thought blocks” instead of “thought block”)
- Also, if some annotations are not really symptoms (e.g. “circumstantial”), you can delete them

In this annotation task, you should try to link each pre-annotated event to one time expression that’s written in the text, thus creating a temporal link (TLINK).

- We will assume an “overlap” relation between the event and the selected time expression.
- In some cases, it will not be possible to identify any link (e.g., “*he told us that he had experienced hallucinations before*”).
- When a symptom clearly refers to the document creation date, you should link it to the time expression representing this date

FIRST UPDATE (05/06/18)

- For each event, **only one temporal link** should be marked.
- We will only consider links to **Date** or **Age_related** time expressions.

SECOND UPDATE (08/06/18)

- You should create a temporal link if it can be easily linked to another reference date written in the document (even at the beginning of the document). For example, if the paragraph containing the event begins with “examination on admission”, and the beginning of the text includes a “date of admission: dd/mm/yyyy”, you should explicitly create this temporal link. When there are many possible reference dates, try to choose the more “likely” one.

THIRD UPDATE (03/07/18, after batch 5)

For the “MENTAL STATE EXAMINATION” paragraph

- If the title contains “ON ADMISSION” then link to the admission date
- If the title does not say explicitly “ON ADMISSION”, then:
 - If it is very obvious that the paragraph still refers to admission (e.g., the following paragraph is “formulation on admission”, the whole letter is about admission), link to admission date
 - If not (e.g., you have a discharge letter, with admission and discharge date), do not annotate

For the “FORMULATION” paragraph

- Do not link to admission date, unless it is obvious that it refers to admission

Note: there will be “ambiguous” examples where you will have to choose whether to create a link or not, based on your interpretation

E.g. *“The patient was referred in **July 2009**. He was initially assessed to be suffering from ... but further assessment suggested possible **hallucinations**”.*