

Annotation Guidelines for Literary Characters  
Nov. 20. 2021

**a. Definition**

A literary character is an animated being that is important to the plot (Jahan and Finlayson, 2019).

By “**an animate being**”, it is:

- a single person indicated by a proper name (Tom Sawyer) or common entity (the boy); or a set of people, an animal, or an anthropomorphic figure, etc. that can be resolved to a proper name, such as “the Dursleys”.

By “**important to the plot**”, it:

- play an active role in the course of the narrative, adapting Jahan et al. 2018’s definition, plot importance describes animacy as “the characteristic of independently carrying out actions in the story world (e.g., movement or communication)”. This definition will be taken into account for post-processing and analytical work and less important when conducting raw annotation work.

**b. Annotation guidelines**

This annotation follows ACE version 6.6, section 3.1’s guidelines on PER annotations, and extends the tag PER to CHAR standing for characters

Annotate:

1. Given a literary character should have significance to the plot, we only focus on mentions that could resolve to **named entities** in the text.
2. Distinct or set of people, animals, or anthropomorphic beings specified by **name** (“John Smith”), **family relation** (“dad”), **definite descriptions** (i.e. “the old hobbit”, “the butcher”), **pronoun** (“he”), or some **combination of these**.
3. Dead people and anthropomorphic beings that can be resolved to named entities.
4. When there are nested entity mentions in the text, one should tag the **maximal noun phrase/determinative phrase** of the character mentioned if the spans could be nested. For example, if “Frodo Baggins” appears within the phrase “young heir Frodo Baggins”, we should tag the latter as the maximal span.
5. Collective characters that can be resolved to a named entity: for example, “The Dursleys” and “The Hobbits”.

Do Not Annotate:

1. Appositives (i.e. “Bilbo Baggins, a hobbit,”)
2. relative clauses (i.e. “the hobbit who lives in the hole”)

3. preposition phrases (i.e. “Bilbo of the Shire”)
4. First and second-person pronouns or referents used within a speech for direct addressing should not be tagged. For example, “you” in “you have better come to live here” (second-person pronouns) or “we” in “we should celebrate birthdays together” should not be tagged.
5. Collective characters are only to be tagged if they can be resolved to a proper name. “devoted admirers” or “some hobbits” , or “the hobbits of unimportant families” should not.
6. Possessive pronouns (i.e. “his” in “his money”, “his journey”)