

# How best collect CASTEMO data?

- [Describe your data collection choices](#)
- ["Same as above": Referencing information content in CASTEMO knowledge graphs](#)

# Describe your data collection choices

Every data collection campaign, even the most comprehensive CASTEMO annotation, necessarily **makes choices**, and is **selective**.

In collaboration between several users, and also as time goes by, it becomes increasingly tricky to remember what data collection guidelines you used, what you included, what you skipped... while this is, obviously crucial for the **interpretation of results** based on the data collected. Therefore:

You should always describe your data collection choices.

Furthermore:

It is generally better to keep the description of data collection choices very close to the data themselves.

In CASTEMO knowledge graphs, we recommend to append the description of the collected data as part of the **metadata of the Territory** which holds this CASTEMO data. There is a **pre-defined section Protocol** under each Territory which allows you to capture the basic ones. Beyond that, you might want to describe your choices in a more narrative and comprehensive way in a **document** (webpage, Google Doc), which can be linked from the Protocol.

# "Same as above": Referencing information content in CASTEMO knowledge graphs

## Referring to the content of another

Documents and statements often make references to other documents and statements to express that the content is the same, different, or related in other ways.

... same, different. Treat this here.

## Referring to temporal and spatial information of another document

...

## Epistemic level in the referencing property

General recommendations for epistemic level apply. Therefore:

- If the reference is **contained in the text**, the whole property should be at the textual epistemic level.
- If in this property, you are using an **analytical property type** (e.g. C "same content") to which something corresponds in the text, but the reference is not done with the same words, the involvement of this property type in the property statement will be at **interpretive level**.
- If the **property value** is a document (i.e. "the same thing as text X"), while the actual text is referring to what a previous *person* (not *document*) said, then its involvement in the property statement will be **interpretive**; if the document is being referenced and in virtually the same words as the label of the document has it, then the property value's involvement in the property statement will be **textual**.