

Replacement Choices in Text Sanitization: Annotation Guidelines

This annotation effort is part of a larger research project that seeks to understand how to automatically remove personally identifiable information from text documents (a problem called *text sanitization*). Personally identifiable information refers to any piece of information that may directly or indirectly reveal the identity of a particular individual. Text sanitization is an important problem when dealing with sensitive documents where we need to conceal the identity of given person(s) to protect their privacy.

The result of your annotation work will be included in a new, public dataset released under an open-source license.

The Task

In this task, you are given a number of short biographies extracted from Wikipedia. To conceal the identity of the individual described in the biography, some text spans have already been marked as needing to be replaced. Each text span is shown in a drop-down menu where the values correspond to possible replacements. The original text span for which you will choose a replacement is also provided to help in the decision making process.

Your job is to select in each drop-down menu the best replacement for the text span according to the following two criteria:

1. The replacement should not disclose (directly or indirectly) the person's identity.
2. Provided that the above criteria is satisfied, the replacement should be as informative and readable as possible.

For example, in the sentence:

*PERSON 1 lives and works in **Oslo** ...*

possible choices for '**Oslo**' might include *[capital of Norway]*, *[city in Norway]*, *[city]* and *'****'*. The first choice is not general enough since it is as informative as the word Oslo. The second choice is more general, followed by the third choice, and finally the default *'****'*, which is least informative (but also least risky from a privacy perspective). Person names are by default replaced by *PERSON X* (where X is an integer).

Procedure

The annotation work consists of the following steps:

- **Step 1** Read through the text once.
- **Step 2** For each marked span, look at the list of possible replacements and pick the most appropriate one. Only one replacement can be selected for each text span.
- **Step 3** When you are done with all replacements, review the text one final time. The selected replacements should not disclose the person identity, and the text should be as informative and readable and possible.

Many suggested replacements will be incorrect or irrelevant – this is normal and expected. If none of the suggested replacements are suitable for a given text span, you should choose the default '***' option.

The '***' option

In all the dropout lists of possible replacements, there will be an '***' option. Use this if you find that no other replacement is appropriate.

Sometimes the '***' is the only suitable option, since you might encounter cases where the automatic generation of suggested replacements failed to come up with good options.

Corner cases

There might be cases where a replacement looks appropriate but does not entirely fit the form of the sentence. For example, in the following sentence:

*PERSON 1 was born on **May 18, 1943** [...]*

The possible replacements will be *[1943]*, *[date in the 1940s]* and '***'. The most suitable choices in this case are *[1943]* and *[date in the 1940s]* (although it might necessitate some rephrasing to fit the current form of the sentence), not '***'.

Example

Below you will find a step-by-step example of the annotation steps.

Start by briefly reading the text (**Step 1**)

Then for each of the spans choose one replacement (**Step 2**). Following is a possible set of replacements chosen.

For example, the two decades could be replaced with the '***' option since they provide additional information along with the rest of the personal information still left in the text (e.g. *British, gay rights activist, general secretary* etc.) that could lead to the person being identified easier, which we wish to prevent.

Mark Christian Ashton ((1960-05-19) 19 May 1960 – (1987-02-11) 11 February 1987) was a British gay rights activist and co-founder of the Lesbians and Gays Support the Miners (LGSM) support group. He was a member of the Communist Party of Great Britain and general secretary of the Young Communist League .

Submit and go to next

Figure 1: Step 1

[PERSON 1] (([DATE]) *** – ([DATE]) ***) was a British gay rights activist and co-founder of the [voluntary association] support group. He was a member of the Communist Party of Great Britain and general secretary of the [youth organization] .

Submit and go to next

Figure 2: Step 2

Note that there is no one correct solution, as long as the identity of the individual is not disclosed and the replacement choices result in an (as much as possible) informative text.

NB! You have to choose a replacement option. The original string is provided (first option in the drop-down list that cannot be chosen) in order to help choose the most appropriate one. The *Submit and go to next* button can only be clicked if replacements for all the spans have been selected.

Read the text with the selected replacements one last time (**Step 3**). Make sure that you have chosen replacements for all text spans. Click on *Submit and go to next* to continue with the rest of the texts for this task.

[PERSON 1] (([DATE]) *** – ([DATE]) ***) was a British gay rights activist and co-founder of the [voluntary association] support group. He was a member of the Communist Party of Great Britain and general secretary of the [youth organization] .

Submit and go to next



Figure 3: Step 3

A short message will appear on your screen when your assigned number of texts have been annotated.