

This folder contains a template to add a tokenization test.

Usage

Using the `cookiecutter` utility requires to have all the `dev` dependencies installed.

Let's first fork the `transformers` repo on github. Once it's done you can clone your fork and install `transformers` in our environment:

```
shell script git clone https://github.com/YOUR-USERNAME/transformers
cd transformers pip install -e "[dev]"
```

Once the installation is done, you can generate the template by running the following command. Be careful, the template will be generated inside a new folder in your current working directory.

```
shell script cookiecutter path-to-the folder/adding_a_missing_tokenization_test/
```

You will then have to answer some questions about the tokenizer for which you want to add tests. The `modelname` should be cased according to the plain text casing, i.e., BERT, RoBERTa, DeBERTa.

Once the command has finished, you should have a one new file inside the newly created folder named `test_tokenization_Xxx.py`. At this point the template is finished and you can move it to the sub-folder of the corresponding model in the test folder.