Universiteit Utrecht

Informed Consent Page

Welcome, with this study we will investigate the quality of Large Language Models (LLMs) generated code documentation. As a participant, you will need to assess one specific Python file with different versions of code documentation on its quality of code documentation. The focus will be on the class and function documentation. Participation will take approximately 30 minutes.

This study was created by Jordi van der Lem, Rick van Grootveld and Oussama el Kandoussi for the course 'Software Production' at Utrecht University. During the study, no personal information will be stored. The answers given will be processed and stored anonymously. As a participant, you participate in this study on a voluntary basis. You have the right to stop participating in this study at any time without accountability. You have the right to view, rectify, delete your data recorded (if possible). There are no ethical or privacy risks with participating in this study. If you have any questions, please contact: o.elkandoussi@students.uu.nl.

Please use a **laptop or computer** when conducting this experiment.

Are you 18 years of age or older?

○ Yes, I am at least 18 years old.
○ No, I am *not* at least 18 years old

Do you agree to participate in this research?

○ Yes, I have read and understood the information provided above about the study and my rights, and I do consent to participate in this research.

○ No, I do *not* consent to participate in this study.

Background - Python and LLM experience

Background Information

Before we start the experiment, we want to ask you some questions about your background information regarding your experience in Python programming. This is used to investigate relationships and correlations in assessing LLM-generated documentation quality.

How much experience do you have with Python?

○ No experience
○ Less than 1 month
○ 1 to 6 month
○ 6 month to 1 year
○ 1 to 2 years
○ More than 2 years

How often do you use Python?

○ Never
○ Daily
○ Weekly
○ Monthly
○ Irregulary during the year(s)

For what purposes do you use Python?

- [ ] For school/higher education purposes
- [ ] Personal or learning purposes
- [ ] Working purposes
- [ ] [   ] Other ...

Stripped Code

Experiment Explanation

Below, you will find a snippet of Python code that will be used throughout the experiment. The Python code has currently no code documentation. To analyze the code file, click the link with the file name. You will analyze the files in new tabs within your browser, while answering the questions on this survey tab.

First, analyze the code below and try to understand what the code is about.

Second, Once you have accomplished that, five versions of the same code will be shown in the following pages, along with the corresponding code documentation. For each version of Python code you will be asked to evaluate the following factors:

Accuracy: the degree to which the code documentation precisely covers the essence of the code.
Understandability: the degree to which the code documentation helps in grasping the code provided in the file.
Completeness: the degree to which the code documentation contains every function/class level documentation.
Relevance: the degree to which the code documentation adds value to the source code by actually describing the code without hallucinations.

Finally, we ask you to rank the five versions of code documentation based on their quality with a reasoning about your choice.

Analyze the 'File S' python code and try to understand what the code is about before proceeding further.

Link 'File S'

Block 3

Evaluation Quality Code Documentation

File A (1/5)

Analyze code file 'File A' on its quality of code documentation.

Link 'File A'

How correctly describes the documentation the code?

The documentation is a complete misconception and/or inaccurate about the code	The documentation contains mostly misconceptions and/or inaccuracy about the code	Approximately half of the documentation contains misconceptions and/or inaccuracy about the code	The documentation contains some misconceptions and/or inaccuracy about the code	The documentation interpreted all the concepts that the code contains correctly
1				5

How easy can you perceive the documentation that has been provided?

	The documentation is hard to read	The documentation is mostly unreadable	The documentation is readable, but needs attention	The documentation is mostly readable	The documentation is easy to read
	1				5

Is everything of the function/class contained in the documentation?

	Completely missing documentation	Most functions missed documentation	Approximately half of the code is documented	Not all the functionalities of the code were contained in the documentation	All of the code was documented
	1				5

Is there no hallucination generated (not going off topic, stick to the subject that matters)?

	The documentation is not relevant to the code	Some of the documentation is relevant to the code	Half of the documentation is relevant to the code	Most of the documentation is relevant to the code	All of the documentation is relevant to the code
	1				5

Block 8

File B (2/5)

Analyze code file 'File B' on its quality of code documentation.

[Link 'File B'](#)

How correctly describes the documentation the code?

The documentation is a complete misconception and/or inaccurate about the code	The documentation contains mostly misconceptions and/or inaccuracy about the code	Approximately half of the documentation contains misconceptions and/or inaccuracy about the code	The documentation contains some misconceptions and/or inaccuracy about the code	The documentation interpreted all the concepts that the code contains correctly

1 ○   5

How easy can you perceive the documentation that has been provided?

The documentation is hard to read	The documentation is mostly unreadable	The documentation is readable, but needs attention	The documentation is mostly readable	The documentation is easy to read

1 ○   5

Is everything of the function/class contained in the documentation?

Completely missing documentation	Most functions missed documentation	Approximately half of the code is documented	Not all the functionalities of the code were contained in the documentation	All of the code was documented
1				5

Is there no hallucination generated (not going off topic, stick to the subject that matters)?

The documentation is not relevant to the code	Some of the documentation is relevant to the code	Half of the documentation is relevant to the code	Most of the documentation is relevant to the code	All of the documentation is relevant to the code
1				5

Block 7

File (3/5)

Analyze code file 'File C' on its quality of code documentation.

Link 'File C'

How correctly describes the documentation the code?

The documentation is a complete misconception and/or inaccurate about the code	The documentation contains mostly misconceptions and/or inaccuracy about the code	Approximately half of the documentation contains misconceptions and/or inaccuracy about the code	The documentation contains some misconceptions and/or inaccuracy about the code	The documentation interpreted all the concepts that the code contains correctly

1 ○

5

[ ]

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ▶

How easy can you perceive the documentation that has been provided?

The documentation is hard to read	The documentation is mostly unreadable	The documentation is readable, but needs attention	The documentation is mostly readable	The documentation is easy to read

1 ○

5

[ ]

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ▶

Is everything of the function/class contained in the documentation?

Completely missing documentation	Most functions missed documentation	Approximately half of the code is documented	Not all the functionalities of the code were contained in the documentation	All of the code was documented

1 ○

5

[ ]

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ▶

Is there no hallucination generated (not going off topic, stick to the subject that matters)?

	The documentation is not relevant to the code	Some of the documentation is relevant to the code	Half of the documentation is relevant to the code	Most of the documentation is relevant to the code	All of the documentation is relevant to the code
1	○				5

Block 6

File (4/5)

Analyze code file 'File D' on its quality of code documentation.

[Link 'File D'](#)

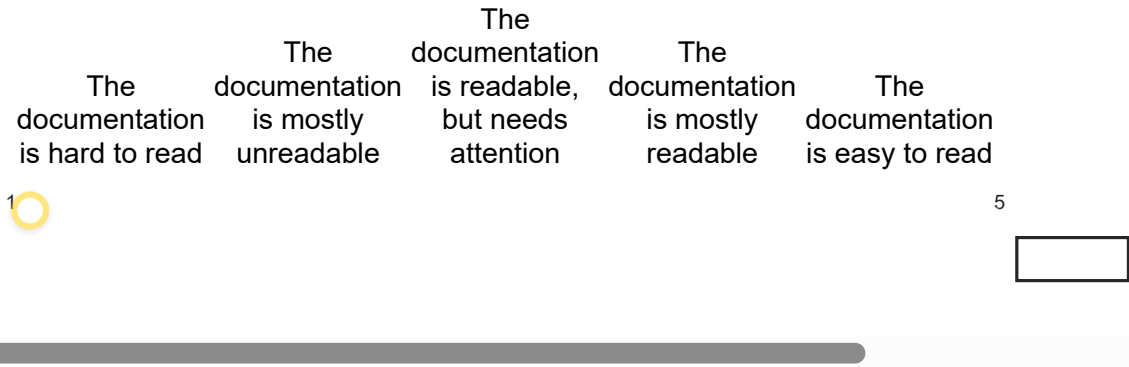How correctly describes the documentation the code?

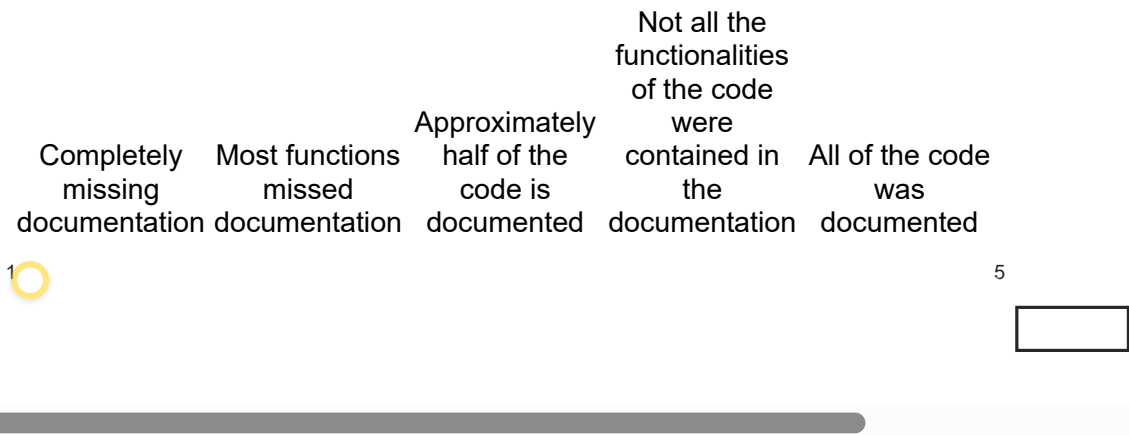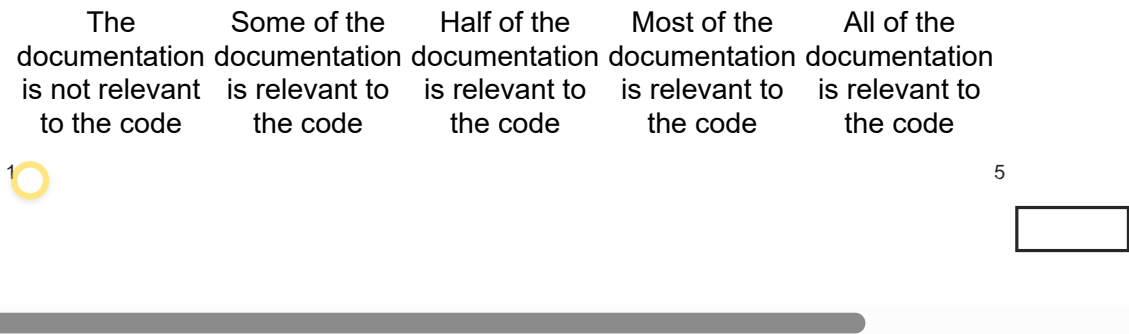|	The documentation is a complete misconception and/or inaccurate about the code	The documentation contains mostly misconceptions and/or inaccuracy about the code	Approximately half of the documentation contains misconceptions and/or inaccuracy about the code	The documentation contains some misconceptions and/or inaccuracy about the code	The documentation interpreted all the concepts that the code contains correctly
1	○				5

How easy can you perceive the documentation that has been provided?

	The documentation is mostly unreadable	The documentation is readable, but needs attention	The documentation is mostly readable	The documentation is easy to read
The documentation is hard to read				

1 ◯ 5

Is everything of the function/class contained in the documentation?

Completely missing documentation	Most functions missed documentation	Approximately half of the code is documented	Not all the functionalities of the code were contained in the documentation	All of the code was documented

1 ◯ 5

Is there no hallucination generated (not going off topic, stick to the subject that matters)?

The documentation is not relevant to the code	Some of the documentation is relevant to the code	Half of the documentation is relevant to the code	Most of the documentation is relevant to the code	All of the documentation is relevant to the code

1 ◯ 5

Block 5

File E (5/5)

Analyze code file 'File E' on its quality of code documentation.

Link 'File E'

How correctly describes the documentation the code?

The documentation is a complete misconception and/or inaccurate about the code	The documentation contains mostly misconceptions and/or inaccuracy about the code	Approximately half of the documentation contains misconceptions and/or inaccuracy about the code	The documentation contains some misconceptions and/or inaccuracy about the code	The documentation interpreted all the concepts that the code contains correctly

1 ○   5

How easy can you perceive the documentation that has been provided?

The documentation is hard to read	The documentation is mostly unreadable	The documentation is readable, but needs attention	The documentation is mostly readable	The documentation is easy to read

1 ○   5

Is everything of the function/class contained in the documentation?

Completely missing documentation	Most functions missed documentation	Approximately half of the code is documented	Not all the functionalities of the code were contained in the documentation	All of the code was documented
1 ○				5

Is there no hallucination generated (not going off topic, stick to the subject that matters)?

The documentation is not relevant to the code	Some of the documentation is relevant to the code	Half of the documentation is relevant to the code	Most of the documentation is relevant to the code	All of the documentation is relevant to the code
1 ○				5

Block 4

Rank Code Documentation

You have currently analyzed all five versions of the Python code. In this section, rank which files have, according to you, have the best code documentation quality.

Rank the code documentation files from **best** code documentation (placed at the top) to **worst** code documentation (placed at the bottom) by drag and drop.

File A

File B

File C

File D

File E

What were important factors/reasoning that influenced the choice of the ranking above?

Feel free to make references to the specific files.