


NAME: Critical Analysis Tool (CAT)

DATE: January 17, 2024 3:47 PM


DESCRIPTION OF TECHNOLOGY

Text analysis tool that uses a large language model to find sections that fall into a category of the user's choice in order to enhance the user's critical thinking in relation to the text.




HUMAN VALUES

The technology may reveal obvious biases and misconduct in reporting, which may make users more suspicious of what they are reading or writing. This is, in fact, the goal of the product, for the user to think critically about a text. However, if the user takes it too seriously or is extremely trusting and reliant on reporting before, it may shake their worldview and make them paranoid.




TRANSPARENCY

Yes, the technology shows explanations to the user as to why which category is marked and the tool shows a description of how it operates.




IMPACT ON SOCIETY

The tool can help people who want to analyse a text more thoroughly, gain a better understanding of it, and find points open to discussion.




STAKEHOLDERS

- Journalists
- Public speakers
- Columnists
- Conspiracy Theorists




SUSTAINABILITY

The energy use has so far not been taken into account. It would require a finished product and measuring which requires extensive research and experiments. At the moment the LLM is used through the OpenAI API. "OpenAI's infrastructure runs on Azure, and Azure will be running on 100% renewable energy by 2025." (OpenAI) Additionally, use of a Vector Storage for frequently occurring requests can minimise traffic. <https://community.openai.com/t/carbon-emissions-of-the-...>



HATEFUL AND CRIMINAL ACTORS

There is no way the tool could break the law or be used to break the law. No data is saved, the technologies used are not saving data. Even if a hacker intercepts user prompts, no sensitive data can be retrieved from the prompts.




DATA

The biases are heavily impacted by the prompts created and the large language model (LLM) used. We try to keep the prompts as logical as possible to avoid creating bias and adhere to standards (links below). However, we cannot influence the LLM which is subject to constant change since it gets updated and is learning.

Links:

- <https://prompt-engineering.co.za/interacting-with-ai/avoidin...bias/>




FUTURE

Since the LLM's AI learns from user input, the model and thereby the tool could improve over time. On the other hand, people's mistakes can reinforce flaws that the model learns, including political and societal bias.

Examples:

- <https://www.internationalhealthpolicies.org/featured-article/what-happened-to-all-human-beings-are-born-free-reflections-on-a-chatgpt-experiment/>



PRIVACY

NO



INCLUSIVITY

Depends on the biases of the LLM, if this has a built-in bias, they will appear in the CAT (see examples). Also if bias appears in the prompts, the CAT will be biased.

Examples:

- <https://www.internationalhealthpolicies.org/featured-article/what-happened-to-all-human-beings-are-born-free-reflections-on-a-chatgpt-experiment/>
- <https://www.brookings.edu/articles/the-politics-of-ai-chatgp...>



FIND US ON WWW.TICT.IO

THIS CANVAS IS PART OF THE TECHNOLOGY IMPACT CYCLE TOOL. THIS CANVAS IS THE RESULT OF A QUICKSCAN. YOU CAN FILL OUT THE FULL TICT ON WWW.TICT.IO





NAME: Critical Analysis Tool (CAT)

DATE: January 17, 2024 3:47 PM

DESCRIPTION OF TECHNOLOGY

Text analysis tool that uses a large language model to find sections that fall into a category of the user's choice in order to enhance the user's critical thinking in relation to the text.




HUMAN VALUES

How is the identity of the (intended) users affected by the technology?


To help you answer this question think about sub questions like:

- If two friends use your product, how could it enhance or detract from their relationship?
- Does your product create new ways for people to interact?...

TRANSPARENCY


Is it explained to the users/stakeholders how the technology works and how the business model works?

- Is it easy for users to find out how the technology works?
- Can a user understand or find out why your technology behaves in a certain way?
- Are the goals explained?
- Is the idea of the technology explained?
- Is the technology company transparent about the way their...

IMPACT ON SOCIETY

What is exactly the problem? Is it really a problem? Are you sure?


Can you exactly define what the challenge is? What problem (what 'pain') does this technology want to solve? Can you make a clear definition of the problem? What 'pain' does this technology want to ease? Whose pain? Is it really a problem? For who? Will solving the problem make the world better? Are you sure? The problem definition will help you to determine...

STAKEHOLDERS

Who are the main users/targetgroups/stakeholders for this technology? Think about the intended context by...


When thinking about the stakeholders, the most obvious one are of course the intended users, so start there. Next, list the stakeholders that are directly affected. Listing the users and directly affected stakeholders also gives an impression of the intended context of the technology.

...

SUSTAINABILITY


In what way is the direct and indirect energy use of this technology taken into account?

One of the most prominent impacts on sustainability is energy efficiency. Consider what service you want this technology to provide and how this could be achieved with a minimal use of energy. Are improvements possible?

HATEFUL AND CRIMINAL ACTORS

In which way can the technology be used to break the law or avoid the consequences of breaking the law?


Can you imagine ways that the technology can or will be used to break the law? Think about invading someone's privacy. Spying. Hurting people. Harassment. Steal things. Fraud/identity theft and so on. Or will people use the technology to avoid facing the consequences of breaking the law (using trackers to evade speed radars or using bitcoins to launder...

DATA

Are you familiar with the fundamental shortcomings and pitfalls of data and do you take this sufficiently into...


There are fundamental issues with data. For example:

- Data is always subjective;
- Data collections are never complete;
- Correlation and causation are tricky concepts;
- Data collections are often biased;...

FUTURE


What could possibly happen with this technology in the future?

Discuss this quickly and note your first thoughts here. Think about what happens when 100 million people use your product. How could communities, habits and norms change?

PRIVACY

Does the technology register personal data? If yes, what personal data?

If this technology registers personal data you have to be aware of privacy legislation and the concept of privacy. Think hard about this question. Remember: personal data can be interpreted in a broad way. Maybe this technology does not collect personal data, but can be used to assemble personal data. If the technology collects special personal data (like...


INCLUSIVITY

Does this technology have a built-in bias?

Do a brainstorm. Can you find a built-in bias in this technology? Maybe because of the way the data was collected, either by personal bias, historical bias, political bias or a lack of diversity in the people responsible for the design of the technology? How do you know this is not the case? Be critical. Be aware of your own biases....

FIND US ON WWW.TICT.IO

THIS CANVAS IS PART OF THE TECHNOLOGY IMPACT CYCLE TOOL. THIS CANVAS IS THE RESULT OF A QUICKSCAN. YOU CAN FILL OUT THE FULL TICT ON WWW.TICT.IO

Fontys
University of Applied Sciences

