Ethical Report

Group 2 – LLM Insights, CY2  
FHICT

# Introduction

The purpose of this document is to describe both the process as well as the results of acquiring ethical requirements.

# Process

This chapter will present the process chosen as a means to gather information on the ethical requirements of the group project.

The chosen approach for the process of gathering ethical requirements is the one of the TICT ethics game, using the *Tarot Cards of Tech.*

## Picked cards

The cards that have been deemed most appropriate for the group project are:

* “What is the biggest negative effect your product can cause?” - Ed Tenner (author)
* “What if your product becomes wildly successful?” - Jack Dorsey (Twitter founder)
* “Does your product invade the privacy of users?” - Edward Snowden (Privacy Hero)
* “What could a bad actor do with your technology?” - Anonymous (Hacktivist collective)

# Results

This chapter will present the results produced by the *Process* step. It will summarize and, if present, also list the possible solutions of the ethical requirements.

### “What is the biggest negative effect your product can cause?” - Ed Tenner

**Question:** Who or what disappears if your product is successful?

**Answer:** The success of the product will imply that information is easily parsed and can at all times be supplied to users (e.g., students) in short. Given this situation, it is a certainty that places where people distribute general information about an institution (e.g., school) would become less relevant and eventually, with the success of the created application, disappear. That includes places such as info desks.

**Question:** What's the worst that can happen if people use your product (wrongly)?

**Answer:** Using the product wrongly can be done at two different stages of the application. To begin with, at the school’s administration level, wrong use can imply improper management of the documents. This means that information can be leaked or exposed to the wrong audience. Secondly, misuse at the level of the end user (e.g., student) can be represented by an improper means of communication with the chatbot which will prompt suboptimal information, which may be redundant, or in the worst-case scenario, provide false information for the individual.

**Question:** What negative effects can you imagine in a future?

**Answer:** The biggest flaw of this product lies in the core concept. This means that there is yet another layer of technology added to a process. While, in itself, readily available information is desirable, in the greater scheme of things, this is another step which does not require any more social interaction. In short, the biggest negative effect of the application will be the fact that it reduces the amount of required socialization to an even lower degree than before, which, when products of this nature add up in an environment, tend to have unpleasant consequences.

### “What if your product becomes wildly successful?” - Jack Dorsey

**Question:** How would a large community of your passionate users change the world?

**Answer:** The product is designed for a specific purpose of providing up-to-date information based on documents. It does not aim to have any significant impact on the users’ lives, therefore it can't be said that them using it would “change the world”. It would certainly bring about a higher level of convenience when searching for and retrieving university-related information, but this makes no global impact.

**Question:** What if 80% of your friends use this product?

**Answer:** If 80% of the friends in our immediate surroundings use the product, the university-related information they receive will be sourced from the same source, meaning that everybody will be able to keep up to date with the university's policies.

**Question:** What if your product is bought by Facebook, Google or Amazon?

**Answer:** In that case, most likely the system would be integrated into different platforms such as Microsoft Teams, Canvas portal, website of the universities, etc. It would receive a lot of funding which will skyrocket the project's development. It is highly unlikely that the product would be purchased by tech giants as they already offer similar products.

### “Does your product invade the privacy of users?” - Edward Snowden

**Question:** Does your product make people feel that their privacy is invaded?

**Answer:** The product stores a lot of user information which is typically regarded as confidential and personal, namely chat logs of messages exchanged between the users and the LLMs. The stakeholders have expressed the need for a feature to be able to view those chat logs, in order to improve the product. For this reason, it is extremely important that the users are well-informed about the privacy policy of CY2, so that they know specifically how their data is handled.

**Question:** Does your product collect data in a non-GDPR compliant way?

**Answer:** GDPR compliance has not been a concern for the project's lifespan so far. However, since our group will be implementing features that rely heavily on user-provided data, it is a good idea that the stakeholders take initiative to outline specifically how they intend to deal with GDPR. It might be the case that the use of the product is already covered by universities’ data policies, so it is a good idea to discuss it with the stakeholders.

**Question:** Does your product really need all the data it collects?

**Answer:** The product stores data for the sole reason of providing insights, facilitating a improving mechanism that can be used by the developers/administrators to improve the product. The core of the software can function without storing any user data, but there would be no way to monitor if the product is actually useful or not.

### “What could a bad actor do with your technology?” - Anonymous Hacktivist collective

**Question:** What is the worst headline about the product that you can imagine?

**Answer: “***School shooter found the university floor plans through the university’s chatbot***”.** This headline is an over exaggeration of what might happen if wrong information is leaked through the chatbot. Having a bad actor maliciously access information which is not indented for public access is a possibility and might have dire consequences.

**Question:** How can fakers, thieves or scammers abuse your technology?

**Answer:** Acts of malicious social engineering have the same degree of occurrence as they would with the barebones LLM models used. Although, this security flaw is one directly linked to the LLM providers. Given that the documents provided to the LLMs are handled properly, no extra information could be extracted, because the LLM does not have access to it.

**Question:** Can your technology be used against certain (ethnic) groups?

**Answer:** It is possible that the technology can be biased against certain ethnic groups – large language models are typically trained on a large corpus of text, which usually contains content from the internet and books. These data sources can contain text that is considered discriminatory. However, as our technology does not train its own LLMs, but rather relies on third party providers such as OpenAI, it is out of our jurisdiction to handle biases such as that.