Lightning Talk 8: Ethics and Professional Responsibility

SDmay25-07: Ask Captain Cyber

Client & Advisor: Doug Jacobson

Ask Captain Cyber Project Overview

Ask Captain Cyber is a cyber-security focused chatbot that will be able to answer user's questions relating to cybersecurity, with support for all levels of complexity. Questions that have not already been answered will be generate by an LLM and vetted by experts to ensure accuracy.
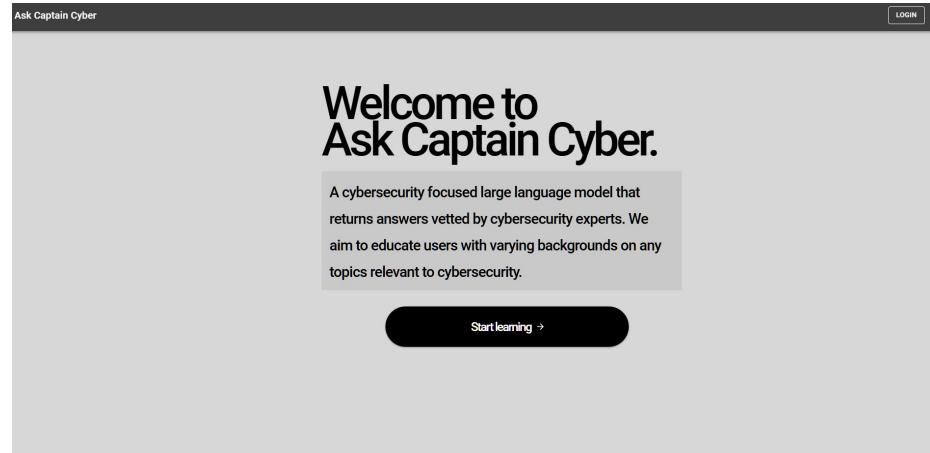

Figure 1 - "Ask Captain Cyber" Landing Page

Professional Responsibility - Financial responsibility

- Financial responsibility is one our primary ethics:
  - We need to create the best possible solution given the budget for this project
  - We have ensured this by comparing the price and features tradeoffs between different LLMs and WordPress plugins
  - This approach upholds ethical and professional responsibilities since we are considering many options and learning of their strengths to provide the most capable and reliable solution with the provided resources

Professional Responsibility - Social Responsibility

- Social responsibility is also another primary aspect of our project:
  - Since people of all cybersecurity skill levels will use this product, we have to make sure it is consistently reliable.
  - We also have to make sure it produces accurate, valid information to not mislead users.
  - We are working toward this by developing a rigours system initialization prompt.
 - We are also setting boundaries on generated responses so it consistently remains accurate and appropriate.

Broader Context Area

	Beneficence	Nonmaleficence	Respect for Autonomy	Justice
Public health, safety, and welfare	Our project will help enhance the cybersecurity knowledge of all users, creating a more informed society	The design does not allow for Ask Captain Cyber to promote incorrect responses	Give users ability to enhance their own cybersecurity knowledge in any ethical way they can imagine	Equal access to the chatbot for all internet users who navigate to the Iowa State website
Global, cultural, and social	Ask Captain Cyber can promote global security by educating users worldwide	Our design avoids harm by not promoting practices that may conflict with values and being culturally sensitive	Our design will respect cultural differences and align with local norms and practices	Ask Captain Cyber will provide only unbiased assistance to the best of its capabilities
Environmental	Design does not contribute to emissions or otherwise negatively impact the environment	Minimize power consumption through existing servers to reduce emissions	Ask Captain Cyber will promote eco-friendly cybersecurity practices users can engage in	Environmental impacts will not disproportionately affect certain communities or cultures
Economic	Improved cybersecurity knowledge would reduce financial loss from cyber attacks	Prevent Ask Captain Cyber from misuse that would lead to giving out malicious/too many details	Users are capable of using this product whenever they want to or as needed	Free access to chatbot for all internet users

Broader Context Area Cont.

Public health, safety, and welfare

- protecting sensitive information
- delivering accurate advice
- promoting cyber security awareness
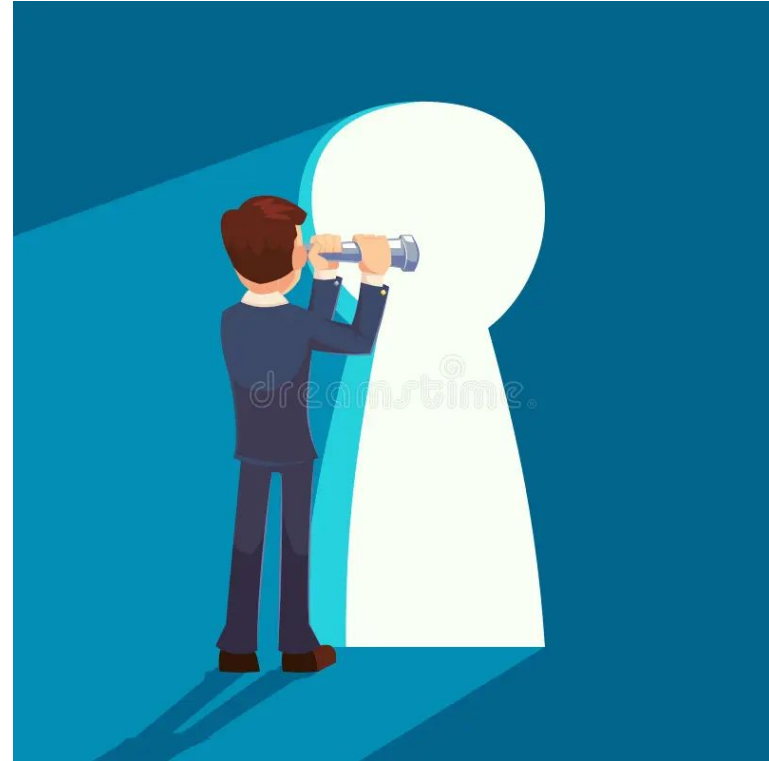
Potential Ethical Issues

- Ethical responses
  - Ethics vary from person to person
  - Ethical dilemmas are difficult for LLMs
  - Reputational damage
  - Stick to the technical details
- Giving out too much technical information
  - Too much technical details is bad
  - Bad actors could misuse this information
  - Use content moderation filters