



Large Language Models for Public Health Communication

Bachelor Project

Research Groups: Digital Health - Machine Learning & Digital Global Public Health

Project Partner: Robert Koch Institute

Background: The COVID-19 pandemic has highlighted the critical importance of effective public health communication in safeguarding the health of the population. Amid this crisis, three aspects were central. First, the digital age has presented both opportunities and challenges for public health communication with online platforms, social media, and apps being instrumental in disseminating health advisories and guidelines. Second, the pandemic saw the rise of 'infodemics,' an overflow of both accurate and false information, stressing the need for accurate and credible information from trusted sources. Finally, trust serves as the foundation for successful public health communication. The rapid spread of misinformation and disinformation, for example, leads to confusion and skepticism among the public. To address this, public health authorities must prioritize transparent, evidence-based messaging, consistent updates, and engagement with communities. Building trust requires fostering open dialogue, acknowledging uncertainties, and addressing concerns, thus empowering individuals to make informed decisions and follow recommended guidelines. The digital landscape demands ongoing vigilance in countering the spread of misinformation online, while addressing issues of accessibility, equity, and digital literacy to ensure inclusive communication.

Project description: The Robert Koch Institute (RKI), Germany's national public health institute, is a source of trusted public health information covering a broad spectrum from infectious diseases, biological hazards, non-communicable diseases and their risks for the general population, as well as the global health situation. By legal mandate RKI communicates with and advises the expert, scientific and political community, therefore over 95% of its information is generated for this specific target group. While the RKI has an active social media presence covering multiple platforms, the bulk of the relevant public health information is found on its website. Hundreds of evidence-based documents have been and continue to be published as the need arises. While some target groups actively seek out information, many find navigating through a website challenging, and seek out more user-friendly alternatives. This also applies to the sheer quantity of the

information that is available which makes searching for information a time-consuming practice. In addition, the non-scientific and non-expert communities also look towards the national public health institute for sound information, which in view of inclusive communication should provide the general public with an opportunity to find the right information using modern digital technology.

Recent advances in large language model (LLM) research promise to enable the development of new kinds of language-based applications. For public health communication, LLMs could be a valuable asset in simplifying access to trusted information, both through their emergent natural language understanding (NLU) and generation (NLG) capabilities. We propose to develop an LLM-based tool to integrate public health information from the RKI website. Our goal is to provide concise and easily understandable explanations on various topics while appealing to different target groups. Such a tool would help overcome the challenges of sorting out through vast amounts of information manually. By harnessing the available technology of LLMs, public health organizations like the RKI can deliver crucial information in a user-friendly manner, increasing accessibility and empowering individuals to make informed decisions about their health.

What you will do:

- Conduct a thorough requirements analysis to determine the specific user needs and engineering challenges.
- Work with state-of-the-art LLMs and LLM frameworks to create and deploy a user-centered digital tool for public health communication.
- Conduct user evaluations to determine project success and areas for improvement.

What you will learn:

- Acquire a comprehensive understanding of LLM-powered applications, contemporary frameworks, and their potential in improving knowledge and risk communication.
- Develop skills in data preprocessing, model fine-tuning, and evaluation.
- Master techniques for effective requirement analysis, team collaboration and project management.

What you should bring with you:

- A strong interest in natural language processing and large language models.
- Proficiency in Python programming and/or Frontend / Web Development.
- An interest in public health communication and a motivation to make a meaningful contribution.

HPI contacts:

Christoph Lippert
Chair Digital Health - Machine Learning
Christoph.Lippert@hpi.de
G-2.1.23

Lothar H. Wieler
Chair Digital Health - Digital Global Public Health
Lothar.Wieler@hpi.de
G-2.2.23

Sina Rampe
Research Assistant Chair Digital Health - Machine Learning
Sina.Rampe@hpi.de
G-2.2.16

Esther-Maria Antao
Scientific Lead Chair Digital Health - Digital Global Public Health
Esther.Antao@hpi.de
G-2.2.34

Florian Borchert
Research Assistant & PhD Candidate Chair Digital Health - Machine Learning
Florian.Borchert@hpi.de
G-2.2.15