PRODUCING INFORMATIVE CELL-SPECIFIC DATA USING GENERATIVE ARTIFICIAL INTELLIGENCE

A SOCIO-TECHNOLOGICAL ANALYSIS OF CHATGPT's EFFECTS ON INFORMATION SPREAD AND PERCEPTION

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Computer Science

By
Eric Xie

Fall, 2023

Technical Project Team Members: Hyun Jae Cho

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Joshua Earle, Department of Engineering and Society

Introduction

My project revolves around the detailed exploration of the sociotechnical implications that arise from the integration of ChatGPT, a notably powerful artificial intelligence tool, into various online digital platforms. The objective of my project is to comprehensively analyze the broader social, ethical, and cultural impacts. This analysis focuses on ChatGPT's influence on human behavior, particularly within a society where a significant portion of information is spread digitally and the prevalence of AI-generated content is on a noticeable rise. Overall, the question guiding my project is: “How has AI-generated content influenced human behavior, perception, and communication on social media platforms and news websites?”

Investigating the impact of AI-generated content is critical to comprehending how potent artificial intelligence, such as ChatGPT, shapes and molds our society. This understanding is vital in helping us to foreseeing and addressing potential issues that may come because of this integration, encompassing ethical dilemmas, privacy concerns, and the ever-pervasive threat of misinformation (Kreps et al., 2022). In essence, this research is rooted in a desire to understand the complex implications that are created by the integration of ChatGPT into digital information spaces.

The overall approach I take towards analyzing this issue will be to first identify the relevant social groups that are impacted by the integration of ChatGPT into digital information spaces and into society, and how this technology affects them. Afterwards, this research will employ a mixed-methods approach, combining a qualitative analysis of user interactions with quantitative data on patterns of online information dissemination. I will be utilizing the technopolitics framework as described by Kurban to contextualize the findings within the societal domain. (Kurban, 2017) I chose this framework to better understand the way that

ChatGPT has been designed, regulated, and utilized and how impacts our political ideologies and policies, especially those related to the spread of misinformation and the ethical concerns behind policies that may attempt to deal with this issue.

Technical Project

Single cell RNA sequencing, or scRNA-seq, acts as a potent analytical tool that allows for the comprehensive examination of gene expression profiles at a single-cell level. This methodology has many key applications, allowing biological researchers to better understand the specifics of the states of various cell types within different biological tissues. This, in turn, enables the precise identification of the specific cell types and the associated genetic profiles that underlie pathological conditions.

However, the main drawback to this analytical tool is that scRNA-seq tends to be cost-prohibitive and yields a relatively limited quantity of samples, especially in the context of human disease investigations. In contrast, there exists a wealth of an easily accessible alternative, bulk RNA-seq. However, bulk RNA-seq does not include any of the cell type specific information that is found within scRNA-seq data.

To address this divide and harness the potential of the abundant amount of bulk RNA-seq data, in this research endeavor, we introduce an innovative computational framework that capitalizes on the capabilities of generative AI techniques to effectively transform bulk RNA-seq data into scRNA-seq data. Our model, the “bulk to single cell” (B2SC) variational autoencoder, is trained to deconvolute the aggregated bulk RNA-seq data into their individual single-cell transcriptomes by learning the specific distributions and proportions of each cell type.

The potential implications of the B2SC approach are particularly significant when applied to extensive human disease bulk RNA-seq datasets. Providing insights at the single cell level into the underlying mechanisms behind the disease processes is essential to furthering our understanding of diseases.

STS Project

The focus of my research is: "How has AI-generated content influenced human behavior, perception, and communication on social media platforms and news websites?" The technology I examine in this research is ChatGPT, an advanced natural language processing AI that has quickly become integrated into society soon after its release.

This question holds immense significance for our society - it addresses the societal implications ingrained into the integration of ChatGPT and how it shapes our interactions within online digital spaces. Understanding ChatGPT's impact on our society is the first step towards foreseeing and addressing potential ethical and societal challenges arising from this technology, such as concerns over privacy, ethical usage, and the implications for the way that we interact with information in digital spaces moving forward (Ray, 2023). Overall, my research question will guide us towards a better understanding of the complex dynamics between ChatGPT and way we as humans interact with each other in
digital society.

Throughout my research, I have identified several crucial social groups to understand the impact of ChatGPT's integration. These include general social media users, content creators, policymakers, and the AI developers themselves (Caramancion, 2023). First, individuals who actively engage with various social media platforms, where AI-generated content is prevalent

(Abdullah et al., 2022). Social media users come from diverse backgrounds, representing different social, economic, and cultural backgrounds due to the size of this group. Second, I have also recognized content creators - journalists, bloggers, and influencers – who generate content for these digital platforms as a relevant social group. This recognition stems from their ability to use AI-generating tools like ChatGPT to help streamline their content creation process and their influence on a large audience. Policymakers directly involved in shaping regulations related to artificial intelligence, privacy, and the spread of information constitute another relevant group. Their perspectives on the impact that ChatGPT has are extremely important given that they can directly influence policies surrounding artificial intelligence. Lastly, understanding the perspectives of AI developers, especially those involved in creating generative AI like ChatGPT, is essential to this research (Abdullah et al., 2022). Comprehensively analyzing their intentions, design principles, and the considerations they have taken is vital to comprehensively analyzing this technology's impact on society.

While my research focuses mainly on those directly involved in the use of ChatGPT and similar technologies, because of the vast audience that is affected by such a powerful technology, there exist several auxiliary groups that are also impacted by ChatGPT whose perspective is also relevant to this topic. Included in these groups are educators, academic researchers, and various advocacy groups (Ray, 2023; Baidoo-Anu & Ansah, 2023; Bozkurt et al., 2023). The perspectives of these groups are important, but to maintain a feasible scope for this project, the research is limited to the identified primary groups that are directly impacted by ChatGPT.

By defining and identifying these social groups, I can tailor my research methodologies to analyze the perspectives, experiences, and any potential concerns of the identified social groups. Understanding how these social groups interact with ChatGPT and AI-generated content

will enable us to provide insights into the impact of these technologies on society (Stahl, 2023). Furthermore, engaging with the experiences from a wide variety of stakeholders will provide a more comprehensive understanding of the topic, facilitating the formulation of more informed recommendations and policies. Policies designed for ChatGPT using input from diverse groups are more likely to be both effective and inclusive for all groups (Stahl, 2023). By considering the needs and concerns of these primary social groups, we can contribute more effectively to the development of policies within this sphere.

In analyzing ChatGPT through a technopolitics lens, I will be focusing on the following key components discussed in literature: context, purpose, scale and direction, actors, and synchronization (Kurban, 2017). The rationale behind utilizing this framework resides in its ability to comprehensively evaluate the role of ChatGPT in the sociopolitical landscape that is shaped by information and communication technologies. By analyzing the context surrounding ChatGPT, we can discern how this technology fits into the broader societal backdrop. Understanding ChatGPT's purpose helps uncover the intentions behind its development and utilization, revealing whether it aligns with important values such as democratization or empowerment. Evaluating the scale and direction enables us to identify the extent to which ChatGPT may impact power dynamics and the reconfiguration of political structures. Assessing key actors allows us to see how the empowerment of individuals and dynamics between traditional and emerging institutions may shift. Finally, exploring synchronization gives insight into how ChatGPT contributes to the coordination of information and perspectives across multiple layers and spaces. Overall, the technopolitics framework is fitting to the task of analyzing ChatGPT as it allows for a nuanced examination of the societal and political implications that this technology can have.

I will accomplish these goals using the following steps. First, a literature review must be conducted, then data collection based on the identified groups, followed by an analysis of the data using both quantitative and qualitative means, the application of the technopolitics framework onto the analyzed data, and finally, the writing of the report. While I expect these steps to be fluid and overlap with each other, I estimate spending 2-3 weeks on the literature review, and a similar amount of time for the data collection. Next, the analysis of the data, as well as the application of the framework, will take around 3-4 weeks. Throughout these steps, I will be recording any information into my report, and any remaining time will be spent organizing and finishing up the final report.

Key Texts

The following relevant texts have been compiled as primary sources of information which grant a better understanding of the topic at hand.

Key Text 1: "All the News That's Fit to Fabricate: AI-Generated Text as a Tool of Media Misinformation" by Sarah Kreps, Miles McCain, and Miles Brundage (2020)

This study focuses on online misinformation, a key focus in this project, and the impact that AI capable of generating human-like text might have on the spread of misinformation. A key differentiator for this paper is that one of the authors, Miles Brundage, is the current Head of Policy Research at OpenAI, the founders of ChatGPT. This results of this study found that individuals tend to struggle with differentiating between AI and human-generated text, even with ChatGPT-2, an older version of the model. Furthermore, it concludes that partisanship influences

the perceived credibility of information generated by ChatGPT, and exposure to this text minimally alters policy perspectives. Analyzing the research conducted by an executive within OpenAI on the beliefs of various social groups is extremely relevant to this project.

Key Text 2: "ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope" by Partha Pratim Ray (2023)

This review explores several aspects of ChatGPT, discussing the development, applications, and challenges of the technology. It focuses on the impact that ChatGPT has on scientific research, such as the potential issues it may bring in terms of ethics and biases. This also notes the attention that ChatGPT has drawn in academia, showcasing the rapid adoption of this technology within the academic sphere. Overall, this research is fairly balanced in its analysis of ChatGPT, noting both the ways that artificial intelligence can revolutionize academia, while also thoroughly listing the risks that ChatGPT may bring.

Key Text 3: "The ethics of ChatGPT – Exploring the ethical issues of an emerging technology" by Bernd Carsten Stahl (2023)

This article examines the ethical issues that may emerge due to ChatGPT. It applies approaches that have been established for analyzing the emerging technologies and applies it to ChatGPT in order to complete a systematic review of the benefits and concerns of ChatGPT. Overall, this analysis finds that ChatGPT's capabilities to produce humanlike text and

interactions seamlessly could provide high-level societal and ethical benefits, however, it raises significant concerns in terms of social justice, individual autonomy, cultural identity, and even environmental issues. Overall, the implications of this paper suggest the need for diverse stockholder engagement and multi-level policy interventions in order to maximize the potential for good within ChatGPT. This research is relevant to this project as the approaches it takes throughout their research, as well as the conclusions that are drawn, are directly related to the question that is attempted to be answered in this project.

Key Text 4: "Harnessing the Power of ChatGPT to Decimate Mis/Disinformation: Using ChatGPT for Fake News Detection" by Kevin Matthe Caramancion

This paper takes a different approach towards analyzing the ethical implications of a tool such as ChatGPT. Rather than focusing on the generative AI's ability to generate false information, this research considers the possibility of using ChatGPT as a distinguisher between misinformation compared to legitimate news content. This is relevant to this project as it creates a multifaceted approach towards analyzing the impact that ChatGPT has on the spread of misinformation. The results of this paper indicated that ChatGPT could predict the legitimacy of every item with 100% accuracy. This provides a new perspective from which the impact that ChatGPT has on misinformation can be analyzed.

Bibliography

Abdullah, M., Madain, A., & Jararweh, Y. (2022, November). ChatGPT: Fundamentals, applications and social impacts. In 2022 Ninth International Conference on Social

Networks Analysis, Management and Security (SNAMS) (pp. 1-8). IEEE.

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52-62.

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., ... & Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. Asian Journal of Distance Education, 18(1).

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt. arXiv preprint arXiv:2303.04226.

Caramancion, K. M. (2023, June). Harnessing the Power of ChatGPT to Decimate Mis/Disinformation: Using ChatGPT for Fake News Detection. In 2023 IEEE World AI IoT Congress (AIIoT) (pp. 0042-0046). IEEE.

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. Journal of experimental political science, 9(1), 104-117.

Kurban, C., Peña-López, I., & Haberer, M. (2017). What is technopolitics? A conceptual schema for understanding politics in the digital age. IDP. Revista de Internet, Derecho y Política, (24), 3-20.

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems.

Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT–Exploring the ethical issues of an emerging technology. International Journal of Information Management, 74, 102700.