

Thesis Project Portfolio

Enhancing AI Transparency in Cybersecurity: Tackling the Black Box Problem through Explainable AI

(Technical Report)

AI's Influence on Political Communication and Its Threat to Democratic Integrity

(STS Research Paper)

An Undergraduate Thesis

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Caroline Coughlin

Spring, 2025

Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Enhancing AI Transparency in Cybersecurity: Tackling the Black Box Problem through Explainable AI

AI's Influence on Political Communication and Its Threat to Democratic Integrity

Prospectus

Sociotechnical Synthesis

My technical project and STS research project focused on different domains but both explored pressing ethical challenges involved with artificial intelligence (AI). My technical project was completed through a cybersecurity internship and tackled the black box problem in AI, aiming to enhance the transparency and reliability of an AI system used for due diligence document analysis. Meanwhile, my STS research paper examined AI's influence on political communication, arguing that AI-driven algorithms and bots function as active agents of electoral manipulation. Although their focuses differed, cybersecurity versus political communication, both projects addressed the increasing urgency for greater transparency, accountability, and governance of AI technologies. Using Langdon Winner's framework of technological politics, my STS research offered broader insight into the political and ethical implications of AI, expanding upon the technical work I accomplished to make AI systems more explainable and trustworthy.

In my technical project, I worked to redesign and optimize the prompt engineering of an AI model used in cybersecurity compliance tasks. The AI application extracted information from client documents to answer due diligence questionnaires, a process critical for regulatory audits. By refining input prompts and implementing step-by-step logic explanations, I improved both the model's accuracy and its transparency to users. These efforts directly addressed the black box problem, helping lower the risk of incorrect outputs and improving user trust. The project demonstrated how prompt-based methods and iterative testing can make AI decision-making processes more understandable, which is critical in high-stakes fields like cybersecurity and finance.

In my STS research paper, titled “AI’s Influence on Political Communication and Its Threat to Democratic Integrity,” I argued that AI technologies are not neutral tools but political actors that reshape electoral discourse. Using technological politics as a conceptual framework, I analyzed how AI algorithms and bots amplified disinformation, reinforced ideological echo chambers, and undermined democratic values during the 2016 U.S. presidential election. My research highlighted three principles: that technologies embed political values, often develop outside democratic oversight, and structure political participation in ways that may not be immediately visible. Through this analysis, I demonstrated that AI’s design and deployment often serve the interests of political and economic elites, posing significant threats to electoral integrity both domestically and globally.

My work on these two projects deepened my understanding of the socio-technical aspects of AI development. The technical work showed me firsthand how even well-intentioned AI systems can diminish transparency and trust without careful design interventions. Meanwhile, the STS research allowed me to recognize that technical solutions are insufficient if not in tandem with broader regulatory and ethical considerations. The insights from my STS project have made me more critically aware of the veiled political effects embedded in technological developments. Moving forward, I intend to apply these insights by ensuring that any AI systems I help develop not only meet functional goals but are also designed with ethical transparency, user empowerment, and democratic accountability in mind.