The Intersection of Ethics and Artificial Intelligence in U.S. Federal Cybersecurity: An In-Depth Analysis

STS 4500 Prospectus

School of Engineering and Applied Science

Computer Science

The University of Virginia, Charlottesville

Name: Kevin Carlson

Technical Advisor:

STS Advisor: Alice Fox

Projected Graduation Date: May 19, 2024

Submission Date: May 9, 2023

The Intersection of Ethics and Artificial Intelligence in U.S. Federal Cybersecurity: An In-Depth Analysis

Overview:

The United States government faces increasing cyber threats, occupying a threshold larger than human technical resources can deal with (Norris et al., 2015). Recent advances in the field of artificial intelligence create new methods for defense organizations to mitigate/prevent threats and execute offensive attacks in accordance with ethical laws and customs held by the U.S. government. My technical project focuses on an innovation in artificial intelligence that has been exploited to promote disinformation by manipulating audio and visual characteristics to fabricate fake images and videos. My technical project explores a solution involving an artificial intelligence algorithm for recognizing auditory and visual nuances for distinguishing between real and fake videos. The manipulation of this technology for ill-intended purposes is just one example of how AI can be recklessly abused without proper technological boundaries. One note of this thesis will highlight necessary AI training methods in creating a technology capable of existing within the dynamic legal and ethical boundaries it is functioning in a system.

Positionality:

A quick Google search for "Richest Counties in the U.S." yields published lists with Loudoun County at the top. The county I have grown up in since birth. While I don't see myself as having lived a spoiled or lavish lifestyle, I do recognize the position my parents were fortunate enough to put our family in has provided excellent opportunities for all of us. As a child and teenager, I was able to define my identity through my experiences and explore my academic, social, and athletic interests with very little bounds. Early on in academia, I found my natural

ability in the sciences and math. It set me on a path in school that would push me into accelerated classes and eventually a magnet high school. I was introduced to STEM and engineering concepts early and was able to explore anything and everything within them due to the human and technological resources around me.

Northern Virginia was much of what I knew and held to be the truth. It wasn't until later on in my life that I was able to gain the perspective necessary to see the area as the cultural and socioeconomic "bubble" that it is. It surrounded me on a daily basis and clouded my perception of the world as a whole. With this understanding, I knew I needed to make use of the opportunities and skills I had been given to make a difference in whatever way I could. Utilizing my passion for cybersecurity and the connections around me, I knew it was my journey to develop applications and methodologies for the typical consumer to help combat the growing landscape of cybercrime.

Problematization:

Artificial intelligence is a revolutionary breakthrough in modern technology and the applications the technology possesses are limitless. When such a powerful tool is utilized in such a sensitive field as cybersecurity, action must be taken to uphold ethics. Federal cybersecurity is a growing area of concern as superpowers around the world become increasingly fluent in their cyber capabilities, both offensively and defensively. Whether AI is being used to assist in calculated offensive attacks or defend the sensitive networks it belongs to, the system must adhere to ethical bylaws put in place to prevent unjust maleficence or unprecedented actions. The paper highlights countries around the world and their use cases for AI in cybersecurity, as well as the ethical literature surrounding such use. The analysis will conclude with the most reasonable

approach the U.S. government should take to involve future AI technologies in their cybersecurity divisions.

Guiding Question or Main Argument:

What is the preeminent ethical framework for the U.S. government to employ for ensuring responsible AI in federal cybersecurity?

Projected Outcomes:

The research completed in this project will provide insight into the current ethical frameworks for AI employed by governments around the world. It will elaborate on such policies and focus their implementation on the U.S. government. The project will conclude with a suggestion for an ethical framework to be instilled by the U.S. federal government for future advances in AI technology within their cybersecurity ventures.

Technical Project Description:

The digital world is constantly evolving, and news stories and videos are created and shared almost continuously. Recent developments in AI have made it possible to harness video and voice characteristics, giving ill-intended people the ability to manipulate what a person does or says. For example, a person could use AI to create a rendering of a presidential address where a convincing digital avatar of the president is recorded saying something false. Should such a video be shared enough without proper means of identifying the falsity and lackluster education on the dangers of this technology, we could see the wide belief of this video manifest in changes in political trends and individual actions. My technical project focuses on an application utilizing AI capable of recognizing visual and verbal nuances of legitimate characters to differentiate them

from fabricated alternatives to protect the public from frighteningly convincing disinformation. Such a technology would be comprehensively useful in the social media space for distinguishing between real and fake videos and flagging the fake ones to slow the spread of disinformation everywhere. Further, this has the power to protect our national and foreign interests should false, fabricated videos be shared with principalities globally.

Preliminary Literature Review & Findings:

Current literature provides a wide range of insight on the ethical implications of artificial intelligence in the scope of cybersecurity with a common theme: As AI evolves, it is imperative the ethical guidelines surrounding it evolve in an equal proportion. AI and cybersecurity are complex fields in their own right, and the literature surrounding them individually indicates the inevitable amalgamation of the two. Among AI research in government, current literature points to the competitive global marketplace for advances in AI technology and the different utilization techniques each country employs (Kwon, Lim, 2021). As AI becomes increasingly relevant in various disciplines, STS researchers have observed the need for foundational guidelines to address the concerns of privacy, accountability, transparency, and bias (Shneiderman, 2021). In the field of cybersecurity, researchers around the globe recognize the importance of secure domestic cyberspace and the possibilities that arise in cyber attacks due to the computerization of society. Current literature in the field points to improving community protocols and ethical responses to critical events (Baskerville, 2022). Researchers have also concluded that, in order to establish greater security in domestic cyberspace, state and federal governments must work together to combat the three main weak points: Insufficient funding, governance, and policy (Norris et al., 2015). In current literature that has tied the two fields together for future use, researchers have analyzed the current methodologies for ensuring responsible AI use and how

they address the ethical issues that arise in each field individually. Some researchers have presented potential ethical frameworks to be employed by organizations looking to use AI for their cyber operations. The current research in this area of study provides a deep insight into how artificial intelligence must be approached to uphold ethical principles and the current state of national ethics around the globe in such areas (Timmers, 2019). It will be greatly useful in developing the necessary components of a practical, efficient ethical framework for AI in cybersecurity for the United States federal government.

STS Project Proposal:

“Science and Technology in Society” is how physical and intellectual design creations have been molded by the society it’s created in as well as how such artifacts continue to restructure society. In many ways, such systems are founded in societal framework biases and under obligations to cultural, legal, and ethical constraints and catalysts. This project focuses on the legal and ethical conundrums that come into play when the U.S. government uses AI in its cybersecurity operations. Technological systems have, for a long time, been created under the ethical constraints of the society they have been developed for and have had little mobility away from such. Artificial intelligence is a revolutionary technology capable of synthesizing its own decisions and must adhere to the ethical constraints of society as they evolve on a continuous plane. The development of AI in cybersecurity raises important ethical questions about the impact of these technologies on society; such questions involve privacy, security, and surveillance (Cole, 2022). These technologies have the potential to improve security around the U.S. as a whole greatly but also raise concerns about the ethical implications of using AI to make decisions that could impact people's lives.

The problem will be approached through the lens of ethics and values and their relation to technological policy. Focusing on such factors in a piece that involves three main areas of study relies on the diverse connections between them all. Primarily, the main authors of this piece will be professionals in the field of artificial intelligence and cybersecurity. These individuals are vital to the work for their technical expertise and their ability to provide insight into the challenges and opportunities of using AI in the context of federal cybersecurity, as well as the ethical implications of such technology. It will also be useful to derive this piece from ethicists and legal experts. Such characters provide specialized knowledge on the ethical nuances of cutting-edge technologies in the appropriate landscapes. While these authors inherently come from different disciplines, the interdisciplinary nature of this piece requires such diversity in order to integrate all necessary components most effectively.

The research question of this piece will be investigated using ethical analysis. The specific courses of action necessary for utilizing this framework will incorporate an evaluation of ethics in artificial intelligence in nations worldwide and how such choices have influenced their AI applications in cybersecurity. Analyzing ethical AI guidelines around the world (such as the EU's Ethics Guidelines for Trustworthy AI and similar documents around the world) will provide a useful foundation of values and efficacy rates among individual guidelines. Contributions from the ethical analysis will provide a useful framework for incorporating ethical considerations into the design and deployment of AI systems used in U.S. federal cybersecurity. I expect the results of analyzing such a framework to align with the research question analyzed under the lenses previously described as both heavily consider the ethical implications of new technologies. An ethical perspective will further provide insight into the state of ethics in U.S. cybersecurity today, and how it can be further integrated in the future to promote consistent, responsible deployment

of AI technologies.

Barriers & Boons

As a student, my knowledge of such complex systematical intricacies within artificial intelligent implementations is limited. I can only learn so much in the classroom about such technology and have used my knowledge to fuel my research. It is imperative I conduct significant research in order to find the most appropriate studies and literature for crafting a professional thesis. One method of finding more effective research is to consult professors and industry professionals in the field in order to discover well-reviewed works for the most accurate and convincing foundations for my arguments. Additionally, I am ill-prepared to draw conclusions on policy and legal reform as a student in the field of computer science. Such a problem can be mitigated by introducing works of literature that draw on the complexities of foreign relations and cyber activity related to specific countries of interest.

References

Baskerville, Richard, et al. ‘Organizing Cybersecurity in Action: A Pragmatic Ethical Reasoning Approach’. *Organizing in a Digitized World*, edited by Stefano Za et al., Springer International Publishing, 2022, pp. 190–203.

Cole, Matthew et al. “Politics by Automatic Means? A Critique of Artificial Intelligence Ethics at Work.” *Frontiers in artificial intelligence* vol. 5 869114. 15 Jul. 2022, doi:10.3389/frai.2022.869114

Lim, Ji Hun, and Hun Yeong Kwon. ‘A Study on the Modeling of Major Factors for the Principles of AI Ethics’. *DG.O2021: The 22nd Annual International Conference on Digital Government Research*, Association for Computing Machinery, 2021, pp. 208–218, https://doi.org10.1145/3463677.3463733. DG.O’21.

Norris, D., et al. ‘Cybersecurity Challenges to American State and Local Governments’. *Proceedings of the European Conference on E-Government, ECEG*, vol. 2015, 01 2015, pp. 196–202.

Shneiderman, Ben. “Viewpoint Responsible AI: Bridging From Ethics to Practice: Recommendations for Increasing the Benefits of Artificial Intelligence Technologies.” *Communications of the ACM*, vol. 64, no. 8, Aug. 2021, pp. 32–35. *EBSCOhost*, https://doi.org/10.1145/3445973.

Timmers, Paul. “Ethics of AI and Cybersecurity When Sovereignty Is at Stake - Minds and Machines.” *SpringerLink*, Springer Netherlands, 11 Oct. 2019, https://link.springer.com/article/10.1007/s11023-019-09508-4.