Thesis Project Portfolio

Pixel to Platform: Reforming Online Toxic Communities

(Technical Report)

Ethical Concerns of Artificial Intelligence in Online Communities

(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

Dominic DaCosta
Spring, 2024
Department of Computer Science

Table of Contents

Sociotechnical Synthesis

Deep learning models, sophisticated subsets of machine learning algorithms, have proven transformative across a myriad of industries, particularly within the domain of online gaming. These algorithms, which grow more effective with the increasing size and complexity of data sets and models, serve as the backbone of my technical project. The technical report is based upon an implementation of a moderation system incorporating these models, where they will aid in mitigating the amount of negative interactions.

Parallel to the technical advancements, my STS research explored the ethical and broader sociotechnical impacts of deploying artificial intelligence within online gaming communities. I employed several STS frameworks to examine the inherent biases of AI technologies and their implications for fairness and effectiveness in moderation practices. This investigation highlighted the complex ethical considerations involved in AI deployment, particularly in environments that attract diverse and often young audiences.

Moreover, the political and ethical dimensions of AI implementation in gaming were scrutinized to assess how decisions regarding AI use are made and the consequences of these decisions. The STS thesis explored the ethicality of building these systems, and reducing its bias, particularly within online communities. This analysis underscored the need for more rigorous regulatory frameworks to ensure that AI tools are used in ways that genuinely benefit the public and do not merely advance corporate interests. The broader goal of these papers was to share the significance of developing the tools necessary to mitigate toxic behavior between players. The technical portion encompasses developing an initial system to mitigate these behaviors through a

warning system.