

## **Thesis Project Portfolio**

### **TranSec: Multi-Language Vulnerability Detection Using Transfer Learning**

(Technical Report)

### **Cultural and Economic Impacts of Music Synthesis Systems**

(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

**Liam Brennan**

Spring, 2023

Department of Computer Science

## **Table of Contents**

Sociotechnical Synthesis

TranSec: Multi-Language Vulnerability Detection Using Transfer Learning

Cultural and Economic Impacts of Music Synthesis Systems

Prospectus

## **Sociotechnical Synthesis**

My technical and STS projects focus on the growing landscape of AI in the modern world. Following my interest in AI, I wanted to work with and analyze many current State of the Art AI systems. In my technical project, I get hands-on experience with AI projects by building an AI system that uses novel architectures to detect vulnerabilities in source code. In my STS project, I combined my love for AI systems and music to analyze musical AI systems' potential effect on the music industry and across different social groups. While these two projects were not focused on the same subset of AI, many technologies and techniques, such as similar model architectures and data collection, are used for both. The same technical knowledge gained through my vulnerability detection AI project can be applied to many other AI applications, such as music. Additionally, similar technologies mean that many AI issues and implications, such as the fight over data privacy, may be shared between domains. Working on both projects gives me a broad overview of the current AI landscape and allows me to apply knowledge to other fields within AI.

For my technical project, I introduce a vulnerability detection platform called TranSec. Vulnerable code detection is one of the most critical yet challenging tasks in the modern age. Leveraging knowledge of similar vulnerabilities across different programming languages provides an opportunity for richer vulnerability detection. For this project, I built a machine learning framework, TranSec, that extracts high-level information about vulnerabilities to perform vulnerability detection. Building from recent works on Natural Language Processing (NLP), I designed a new model that can detect similar vulnerabilities across many different programming languages by learning language-independent features of various vulnerabilities.

Transec introduces a new pre-training scheme and architecture that uses the novel Semantic Sovereignty Objective (SSO) on top of other standard loss functions. As a result, TranSec outperforms various vulnerability detection approaches on various vulnerability detection datasets while requiring significantly less training data. Additionally, TranSec is exceptional at extracting high-level information, implying that multi-language tasks other than vulnerability detection, such as language translation, can leverage TranSec. In the future, I plan to release a complete vulnerability detection system based on the TranSec model.

In my STS project, I look into the potential effects of musical AI, especially across different social groups and the music industry. Music Synthesis AI systems (systems that can create full-length songs without human input) have rapidly improved. With the rise in generative AI models such as ChatGPT that are beginning to perform generative language at a human level, music will inevitably follow. In my paper, I show that musical AI will not replace artists but will supplement the creative process, similar to AI in other domains and musical technologies across history. This analysis contrasts much of the apocalyptic sentiment towards musical AI and other generative systems. However, I also show that currently and historically, developers of musical AI systems have almost exclusively catered to Western music. The lack of representation of non-western music within current Musical Synthesis AI systems is an issue that needs addressing as technology progresses. I perform my analysis by combining my research into different sources, such as academic papers on the latest musical AI systems and books about musical social psychology, with the Social Construction of Technology (SCOT) framework, which posits that technology is a product of society.

Working on these projects simultaneously allowed me to perform more nuanced STS analysis and build a better vulnerability detection AI system. AI is a vast and complex domain,

and working on a Musical Synthesis AI STS project is extremely challenging without all of the appropriate prior technical knowledge. Working on large-scale technical AI projects simultaneously allowed me to not just learn about the technicalities behind many AI systems, but helped drive my analysis as I was faced with many of the same problems, such as how to collect data, that developers of Musical Synthesis AI systems face. In particular, I used a similar model architecture in my technical project that is often used for musical systems, which helped inform my analysis. Additionally, working on my STS project alongside my technical project helped me drive many technical decisions as I learned about many different design decisions that Musical AI systems developers make during their development process. For example, reflecting on the different unintended consequences of AI models that I learned about in my STS project helped me watch for unintended consequences of my design decisions.