



# COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

## PhD Qualifying Exam

**Wednesday April 19, 2023 at 11am via Zoom [Remote]**

**Carter Cousineau**

*Demystifying Trustworthy and Interpretable  
Artificial Intelligence (AI) for System Developers*

**Chair:** Dr. Fangju Wang

**Advisor:** Dr. Rozita Dara

**Co-Advisor:** Dr. Ataharul Chowdhury [SEDRD]

**Non-Advisory:** Dr. Daniel Gillis

**Non-Advisory:** Dr. Fattane Zarrinkalam [SoE]

### **Abstract:**

With the need and rise of Artificial Intelligence (AI) systems continuing to increase, the need for trusted and ethical AI systems also increases. This begs the question; how do we ensure AI systems have implemented trustable and ethical AI?

There is a range of diverse complexities in methodologies used to build these AI systems. There is also a broad and diverse range of existing literature on definitions and best practices for trustable and ethical AI. With no unified view on trustable and ethical AI, unclear standardization, pressures for regulatory compliance, and limited oversight on how the trustable and ethical AI are implemented in the AI system lifecycle leaves the need for further research in trustable and ethical AI.

The objective of this research is to investigate and highlight current challenges in trustable and ethical AI. It will also identify the challenges specific to system developers. We plan to focus on system developers as they are the individuals who build AI systems. To the best of our knowledge, no previous research has extensively examined AI developers' perspective. Trustable and ethical AI implementation cannot take place without ensuring the system developers are a part of the adoption process. We also aim to focus on the challenges in interpretability and explainability (XAI) of AI systems, as these two trustable and ethical concepts were commonly examined in the existing literature and many solutions have been proposed. This research will examine the importance of trust and ethics in AI systems in order to improve interpretability and XAI between system developers and the AI system lifecycle. We aim to build decision trees and taxonomies for interpretable and XAI as a way to enable and establish semantic interoperability for system developers.