CYB-4203/6203

Secure and Trustworthy AI

Presentation 2.5: Assignment 1 Insights - Q1 & Q2

Monday, February 2, 2026

Assignment 1 - Q1 & Q2

Student Motivation & Interest Themes

1. Intellectual Curiosity
2. Safety, Ethics, and Law
3. AI is the Future
4. AI is so hot right now (Career Opps)
5. BUILDING Secure & Trustworthy AI
6. AI as Research Subject / Object

THEME 1: Intellectual Curiosity

Intellectual curiosity & interest in theoretical and philosophical questions related to intelligence

"...philosophically I'm intrigued by the concepts (of 'Secure' and 'Trustworthy' AI), and not confident that the goal is even possible. Blaise Pascal suggested that belief and trust are based outside of reason, foundational to knowledge, but not something we can empirically measure. Similarly, most people say they want privacy, but the evidence suggests that they really just want the illusion of privacy. Is objectively Secure & Trustworthy AI practical and obtainable, or is it more realistic to set lower goals?"

"...I would love to learn more about the processes of how AI works and learns... I think it would be interesting to see the inner working of an AI and how it deciphers binary to output human readable response."

THEME 2: Safety, Ethics, and Law

Interest in safety, legal, ethical aspects of AI systems

"...tech tends to develop faster than our laws, so we end up with many legal gaps... I'm interested in learning how AI can be used ethically and without violating privacy."

"...From my project last year, seeing how vulnerable AI models can be has made me realize how important it is to ensure the security and safety of these models early on. I want more insight as to how we could improve these models while keeping people safe."

"...while we don't yet quite know how to use or manage it, what its true capabilities are, people are already using it for nefarious purposes... low-skill people developing complex tools for hacking and malware... to hurt other people, take advantage of financial systems, or destroy things to further their ideology..."

THEME 3: AI Is The Future

Recognition of AI's massive impact potential

"...AI is here to stay... it's more important than ever to ensure that AI systems are reliable and trustworthy for people using them to make crucial decisions."

"...AI is going to impact humanity in ways not seen since the internet was introduced."

"...AI is in every aspect of our lives and uses and relies on our sensitive information..."

"...I personally believe that the current sparks of an AI revolution we are witnessing today will lead to considerable change in the way society functions within my lifetime. Plus, due to the general nature of these systems, there is seemingly a conceivable use case for AI in almost every scenario I could imagine."

THEME 4: AI Is So Hot Right Now (Career Opportunities)

Recognition that AIxCyber skills are in high demand

"...employers want it, and adding an AI specialization to a cybersecurity degree will allow pursuit of career options in both AI and cyber."

"...(because there have been) a lot of AI mishaps in the news, (it seems like secure & trustworthy AI skills will be great to add to my) toolbelt."

”…I really want to grow my understanding of AI and how to create models that are real-world applicable... I only have one year left before I start the rest of my career life and I know I want to do something with AI, but I don't know exactly what that is or how to even start."

THEME 5: BUILDING Secure & Trustworthy AI

Desire for knowledge and understanding to build secure & trustworthy AI/ML systems

"...I have spent some time AI modeling and trying to create my own agents using genAI, but my knowledge and understanding is lacking, so I am excited to take this class and do a lot of hands on projects and assignments to grow my understanding of AI."

"...I do not want to be only an AI user - I also want to be part of the group that designs and builds AI systems. My goal is to contribute to creating AI that people can trust: systems that are secure, privacy-aware, and resilient, with bias minimized as much as possible. Understanding how to build AI responsibly is essential to me, especially from a security and ethical perspective."

"...I'd like to learn more about how to apply AI in a manner that doesn't end up with deletion of my codebase."

THEME 6: AI as Research Subject / Object

Desire to conduct AI research, use AI for research and growth

"...Over time, my interests have shifted toward cybersecurity and trustworthy systems, especially as my research has moved beyond model performance research to questions of reliability and risk..."

"...many people... focus only on the immediate answers AI provides, and accept them as correct without questioning privacy, security, or underlying risks... I'm interested in how these systems work internally, what risks they carry, how they influence decision-making... the role AI can play in my academic journey, and how it can meaningfully support my education and research."

"...I use AI for my research and constantly try to improve its accuracy for the tasks that I test it with. So, I am very interested in how to move more towards secure and trustworthy AI, without having to completely start from zero and fumbling through all of it myself with trial and error."