Case Studies : Safety 1

Case 1 : Modeling Bad Actors in Social Media Platforms

In the programming exercise you completed this week, you worked on the content moderation system for the social media platforms “Catter” and Twitter. In this case study, we expand our exploration of social media platforms to analyze how they can be affected by bad actors’ actions.

Apply the Bad Actors Modeling Strategy to identify and analyze potential harmful actions or negative consequences that could arise from bad actors on social media platforms in general (think about Instagram or TikTok for instance) using the five motivation categories (Money, Politics, Entertainment, Ideas, Self-interest). Consider the following questions :

- ➔ What harmful actions can be taken in each category?
- ➔ How might these actions impact users and the platform?

Case 2 : Analyzing Safety Implications of Autonomous Vehicle Software Using STRIDE Strategy

As a software engineer working on the development of an autonomous vehicle system, your task is to analyze potential safety implications using the STRIDE strategy.

Part 1 : For each situation, identify the associated threat from the STRIDE strategy

1. A cybercriminal intercepts the communication between the autonomous vehicle and the cloud-based control system, gaining access to sensitive information about the vehicle's operations, routes, or passengers.
2. An unauthorized individual gains access to the vehicle's software system and modifies the decision-making algorithms, causing the vehicle to behave unpredictably or dangerously.
3. A malicious actor broadcasts fake satellite signals that override legitimate signals, confusing the receiver, causing the in-car systems to incorrectly position the vehicle.
4. A malicious attacker floods the vehicle's communication channels with excessive data or requests, causing the software system to become overwhelmed and unresponsive, potentially leading to a safety-critical failure.
5. A passenger claims that the autonomous vehicle caused an accident due to a software malfunction, but there is no way to prove or disprove the claim.

Part 2 : Provide a technical countermeasure to prevent or mitigate each issue

Case 3 : Harm Modeling Strategy

The goal of this exercise is to identify and assess potential harms associated with the technologies described in provided scenarios. These scenarios are voluntarily futuristic for practice purposes. The overall goal is to create awareness about the different types of harm technology can cause and make you realize that we tend to underestimate the number or impact of those harms in real life.

Read the provided scenarios, then apply the harm modeling strategy to assess the ethical implications and potential consequences of this technology. In your analysis, don’t forget to consider four categories of use: malfunction, misuse/abuse, unintended use and intended use.

You can use the table from the strategy, reproduced below.

Category	Type of harm	Description of harms:
Humans	Physical injury	
	Emotional or psychological injury	
Resource allocation	Opportunity loss	
	Economic loss	
Human Rights	Dignity loss	
	Liberty loss	
	Privacy loss	
	Environmental impact	
Social Systems	Manipulation	
	Social detriment	

First scenario : Affect-Display Textile Garment (Sleeve)

Read the scenario here: [Sleeve - VSD Lab (vsdesign.org)]

Second scenario : Smart home technologies

In the era of technological advancements, a suite of interconnected smart home technologies has emerged, promising unparalleled convenience, safety, and comfort. The suite includes smart doorbells, connected refrigerators, adaptive robot vacuum cleaners, automatic lights and blinds and a voice-activated assistant like Google Home.
Emily and Mark are two individuals leading busy lives in a bustling city. Emily is a young tech-savvy professional who relies on smart home technology to streamline her daily routine. She installed a smart doorbell with facial recognition to enhance her home security. Her refrigerator automatically replenishes groceries through online orders, adapting to her tastes. Her robot vacuum cleaner keeps her home spotless with minimal effort, while her voice assistant controls her lighting, entertainment, and even her morning coffee.
On the other hand, Mark is skeptical of these technologies. He is a family-oriented parent juggling work and household responsibilities, who prefers manual control over his home environment and values his privacy. His refrigerator is a traditional one, and he cleans his home the old-fashioned way. He believes in limiting the data that smart devices collect about him and his children.
Over time, these smart home technologies become deeply integrated into society, revolutionizing individual lifestyles as well as the economy and urban planning as data collected from smart devices inform city infrastructure investments.
However, not everyone opts for these technologies. Individuals like Mark, who are concerned about data privacy and security vulnerabilities, choose to stick with traditional, non-connected household items. Others cannot afford to equip their home.