**WEEK 3 ASSINGMENT ON AI JUDGEMENT IN REAL WORLD SCENARIOS.**

## 1: The Alphabet Advantage in answer sheet making with ai:

**The Scene**  
Imagine a school where a fancy new **AI grading system** is rolled out. The promise? Faster grading, less teacher workload, and “fair” results. But… there’s a twist.

This AI has a quirky habit: it gives **higher scores to students whose names start with letters later in the alphabet**. So a student named *Zipporah* gets an A+, while poor *Abel* barely scrapes a pass — even though Abel’s answers were just as good.

**What’s Really Happening**  
The AI is not grading based on the **content of the answers**. Instead, it’s using an irrelevant feature (the **alphabetical order of names**) to decide scores. Basically, the algorithm is playing favorites with letters, not learning.

**What Could Go Wrong?**

* **Fairness**: Students are judged by their names, not their knowledge. Totally unfair.
* **Bias**: Zipporah, Yolanda, and Zachary shine; Abel, Brian, and Clara suffer.
* **Accountability**: Who do students complain to? The teacher? The school? Or the mysterious AI that no one understands?
* **Trust**: Once students discover this bias, the entire system loses credibility.

**How to Fix It**  
The AI should be **trained and evaluated only on actual answers** — the essay content, multiple-choice correctness, or logical reasoning. Names should never even enter the model as input.

Extra step: add **transparency reports** so schools (and students!) can see *how* the AI makes grading decisions.

**Fun Bloggy Wrap-Up**  
So, next time you meet *Zipporah* with her suspiciously high marks and *Abel* sulking in the corner, remember: sometimes it’s not the student — it’s the **AI alphabet mafia** at work.

As Responsible AI Inspectors, our job is to shine a light on these shady algorithms and remind schools that fairness isn’t about A-to-Z… it’s about giving **every student an equal shot**.

**2: The Workspace Whiz… or Workplace Whack:**

**The Scene**  
A shiny new **AI productivity tracker** hits the office floor. Its mission? Measure how “hard” employees are working by counting keystrokes, mouse clicks, and time spent in apps.

At first, management cheers: “Finally, data-driven performance reviews!”

But then, things get messy.

**What’s Really Happening**  
The AI assumes **more clicks = more productivity**.

* Sara, a software engineer, spends time thinking, planning, and writing quality code — but doesn’t click much. She gets flagged as “lazy.”
* Meanwhile, Bob, who clicks around aimlessly and types nonsense to stay active, scores “highly productive.”

**What Could Go Wrong?**

* **Fairness**: Creative thinkers and deep workers get punished; button mashers get rewarded.
* **Privacy**: Every move is tracked — creepy, right?
* **Stress & Burnout**: Employees feel spied on and pressured to game the system.

**How to Fix It**  
Instead of tracking **clicks**, the AI should focus on **outcomes** (like completed projects, bug fixes, or creative deliverables). Plus, give employees transparency and control over their data — nobody likes being a digital lab rat.

The AI “workspace whiz” turned out to be more of a **micromanaging boss-bot**. Remember: productivity isn’t about how many times you smash your keyboard — it’s about the value you bring. Let’s design AI that respects humans, not reduces them to click-counters.

**3: The Hospital Helper… or Hazard?**

**The Scene**  
Hospitals adopt an **AI triage system** to decide which patients need urgent care first. Sounds life-saving, right?

Except… the AI has been trained mostly on **data from younger patients**.

**What’s Really Happening**  
The AI starts underestimating the severity of symptoms in **older patients**. For example:

* A 70-year-old with chest pain is told to “wait.”
* A 25-year-old with mild symptoms jumps the line.

**What Could Go Wrong?**

* **Fairness & Bias**: Age-based discrimination sneaks into life-or-death decisions.
* **Accountability**: Who takes responsibility if someone dies waiting because the AI said so?
* **Trust**: Patients lose confidence in hospitals and technology.

**How to Fix It**  
Hospitals must **train AI on diverse, representative datasets** — covering patients of all ages, genders, and backgrounds. Plus, keep a **human doctor in the loop** for final decisions.

In this case, the AI triage tool became less of a **guardian angel** and more of a **biased bouncer at the ER door**. The lesson? Health AI should always be tested for fairness, and humans must stay in charge when lives are on the line.