**Activity 2.1 – When AI Gets It Wrong: Case Study Discussion**

**Objective:**

Explore real-world examples of AI systems that failed, understand *why* they failed, and reflect on what we can learn about responsible and ethical AI use.

**Group Setup**

* Form small groups of **2–4 participants**.
* Each group will be assigned (or can choose) **one case study** from the list below.
* The participants will have 10 minutes for their case study.

**Case Studies (Choose or Assign One)**

1. **COMPAS System (U.S. Criminal Justice)**  
   *Used to predict likelihood of reoffending but was found to be biased against certain groups.*
2. **Amazon’s Hiring Algorithm (2018)**  
   *Designed to screen job applicants but ended up discriminating against women.*

**Discussion Questions**

As a group, discuss and prepare brief answers to the following:

1. **What was the AI expected to do?**  
   *(What was its intended goal or task?)*
2. **What went wrong?**  
   *(Describe the issue, failure, or unintended outcome.)*
3. **Why did it go wrong?**  
   *(Think about causes — data bias, poor design, lack of human oversight, etc.)*
4. **What can we learn from this?**  
   *(How can human judgement and ethical design prevent similar issues?)*

**Sharing Back**

* Each group will share their findings in a **3-4 minute summary**.
* Try to highlight **key lessons** that apply to your own professional context — e.g., decision-making, data use, or fairness.

**Tips**

Focus on **understanding**, not technical details.  
Think about the **human consequences** of AI errors.  
Consider what *you* would have done differently as a designer, policymaker, or user.