**Activity 3.1: The Confidentiality Check (Risk-O-Meter)**

**Objective**

To assess the practical risks of sharing different types of workplace information with public, cloud-based AI tools (like free chatbots) and to develop immediate **Discernment** regarding **Data Privacy**.

**Timeline**

10 minutes

**Theme covered**

Sensitive Data, PII, Confidentiality, and the "Harm Test"

**Instructions**

1. **Form Groups:** Participants should remain in small groups (3-5 people).
2. **Define Signals:** Each group must agree on three signals to indicate the perceived risk level of the data in each scenario:
   * **Green Light (Safe/Low Risk):** **Thumbs Up**
   * **Yellow Light (Caution/Medium Risk):** **Flat Hand**
   * **Red Light (Forbidden/High Risk):** **Thumbs Down**
3. **The Challenge:** The facilitator will read out a series of real-world professional scenarios. After each one, the groups must quickly discuss the data involved and simultaneously signal the appropriate risk level **before the facilitator counts to three.**
4. **Debrief:** After each scenario, the facilitator will ask a group to justify their answer, reinforcing the difference between public, proprietary, and regulated data.

**Reflection Questions (For Final Discussion)**

* **Personal Risk:** Think about a task you often use AI for. Have you ever unintentionally shared "Yellow" or "Red" light information in the past? What one habit can you change immediately to reduce that risk?
* **Organizational Impact:** Beyond legal fines, what is the biggest consequence if our organization suffers a breach of proprietary data due to careless AI use? (E.g., loss of reputation, competitive disadvantage).
* **The Trust Problem:** The AI can't tell the difference between a client's name and the name of a fictional character. Why is this lack of judgment the single biggest reason why we, the humans, must be the gatekeepers of confidential data?

**Facilitator Script & Scenarios**

*(Read these scenarios aloud. After each, ask groups to signal their risk level.)*

| **Scenario** | **Risk Level** | **Key Rationale (for facilitator debrief)** |
| --- | --- | --- |
| **1. Drafting a Public Email:** Asking a public chatbot to write a generic, professional email announcing a new public-facing event. | **Green (Low Risk)** | No confidential data is shared; the information is already public. |
| **2. Analyzing Client PII:** Copying and pasting a client's **full legal name, personal phone number, and detailed financial history** into the AI to analyze their creditworthiness. | **Red (High Risk)** | Direct sharing of **PII** (Personally Identifiable Information) and **regulated financial data**. Violates legal and compliance rules. |
| **3. Summarizing Proprietary Strategy:** Asking an AI to summarize a **confidential, internal Q4 business strategy report** that contains unreleased market forecasts and competitor research. | **Red (High Risk)** | Sharing highly **sensitive corporate data**. Even if the AI doesn't misuse it, the data is logged by an external service, creating security and competitive exposure. |
| **4. Generating Code for a Public Project:** Using an AI to write a short, simple JavaScript function for a non-sensitive internal web application. | **Green (Low Risk)** | Simple, non-sensitive task using generic code. (Caution if the code involves specific proprietary database names or credentials). |
| **5. Testing a Policy Idea:** Describing a **hypothetical, fictional scenario** to the AI to help generate alternative phrasing for a new HR policy draft. | **Green (Low Risk)** | Use of **hypothetical examples** is a key safe practice as it avoids sharing any real data or PII. |