**Case Studies for Reflection**

**Module 10E**

**Case 1: Automated Firing at Uber (GDPR Article 22)**

**Synopsis (1-2 lines).**  
An algorithmic fraud-detection system automatically deactivated driver (“robot-firing”). A Dutch court found the lack of transparency and human oversight unacceptable and ordered reinstatement and compensation. The case was brought under **GDPR Article 22**, which protects individuals from significant, solely automated decisions.

**What went wrong**

* Solely automated, high-stakes decisions affecting livelihoods
* Insufficient transparency and explanation for those affected
* No reliable, documented **human-in-the-loop** review prior to action

**Governance fixes (apply to any organization)**

* Declare when a decision is automated; provide plain-language reasons and recourse
* Build **human oversight checkpoints** before any significant effect (suspension, dismissal, denial of access/benefits)
* Keep decision logs: inputs, model score, human reviewer, final rationale
* Train staff on when and how to **override** an AI recommendation

**Reflection prompts**

* Where in our institution could a system drift toward “auto-decisions” (e.g. auto-flagging misconduct, auto-withholding services)?
* What would a **minimum acceptable review** look like before action is taken?
* Which artifacts prove compliance (logs, notices, SOPs) if challenged?

**Case 2: Algorithmic Bias in Dutch Childcare Benefits**

**Synopsis (1-2 lines).**  
A government risk-scoring system unfairly targeted families (e.g., by nationality proxies), causing widespread harm. Regulators found privacy and fairness violations; the scandal catalyzed stronger AI oversight in Europe.

**What went wrong**

* Use of sensitive or proxy variables leading to **discriminatory impact**
* Weak data governance (provenance, quality, retention)
* No robust, ongoing **bias testing** or impact assessments

**Governance fixes (apply to any organization)**

* Prohibit the use of sensitive attributes (and proxies) unless legally justified and tightly controlled
* Run **bias & impact assessments** pre-deployment and after updates; document metrics and mitigations
* Ensure **transparency**: why people are flagged, how to challenge it, how a human reviews outcome
* Embed **data minimization and retention limits**; audit data pipelines

**Reflection prompts.**

* Which school workflows risk **unequal impact** (placements, access to enrichment, discipline, attendance interventions)?
* What **bias metrics** would we track for those workflows, and how often?
* How do affected students/families **appeal** or request human review?

**“Translate to School” Mini-Exercises**

**A) Automated Action Drill (30 minutes)**

* **Scenario:** An AI tool auto-flags students for academic dishonesty and instantly blocks portal access.
* **Task:** Redesign the flow to insert human oversight, create the appeal path, and specify evidence to retain.
* **Outputs:** 1-page **Oversight SOP** + template **Decision Log** (fields: reason code, reviewer, final action, notice sent).

**B) Bias Audit Sprint (30 minutes)**

* **Scenario:** A support-allocation model is under-selecting students from one background for advanced programs.
* **Task:** Define two fairness metrics, a sampling plan, mitigation options (policy threshold, data balancing), and a re-check cadence.
* **Outputs:** **Bias Test Plan** + **Mitigation Notes** + quarterly re-evaluation schedule.

**Quick Compliance Map (use in your Capstone)**

* **GDPR Art. 22 (automated decisions):** human intervention, contesting, meaningful information about logic
* **EU AI Act (high-risk in education):** risk management, **human oversight**, documentation, post-market monitoring
* **Governance evidence:** transparency notices, oversight SOPs, training logs, bias reports, decision logs, incident reviews