**Incident Classification Matrix & Guidance**

**Module 8C**

**Context:** Spanish secondary school (1º ESO → 2º Bachillerato, ages 12-18)  
**Applies to:** All AI-enabled systems used for teaching, assessment, admin, wellbeing, and safeguarding.

**1) Purpose & Scope**

This document standardises how we **identify, grade, and escalate** AI-related incidents. It covers data/privacy events, AI malfunctions, bias/discrimination, platform misuse/abuse, and safety/security issues across our school stages (ESO, Bachillerato).

**2) Definitions**

* **AI incident:** Any behaviour of an AI system (or its misuse) that risks harm, privacy breach, discrimination, service disruption, or legal non-compliance.
* **Personal data breach (GDPR):** Accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data.
* **High-risk AI (EU AI Act):** AI used to determine access/assess performance in education (e.g., grading, placement, interventions).
* **Serious incident (AI Act):** An event causing death/serious harm, serious fundamental-rights infringement (e.g., discrimination), or other serious outcomes.

**3) Classification Criteria (what we look at)**

1. **Scope/scale:** How many pupils/staff are affected?
2. **Sensitivity:** Type of data/system (e.g., contact details vs. health/biometric).
3. **Impact on individuals:** Educational harm, dignity, mental/physical safety.
4. **Impact on operations:** Lesson disruption, exam integrity, outage.
5. **Legal/regulatory triggers:** GDPR notification? AI Act serious incident?
6. **Reputation/press risk:** Is wider community trust at risk?

**4) Severity Levels (with triggers & mandatory actions)**

| **Level** | **Description** | **Typical Triggers** | **First Actions & Escalation** |
| --- | --- | --- | --- |
| **L1 - Minor** | Low impact, contained, no sensitive data, no rights impact. | Small glitch (e.g., timetable suggestion error caught early), single-user UI bug. | Log in incident register; assign fix to IT/vendor; no external notices. |
| **L2 - Moderate** | Limited cohort affected or temporary disruption; low-sensitivity data exposed to authorised users only; no material rights impact. | 1–2 classes can’t access an AI tool for a period; minor misrouting of non-sensitive data. | Notify AI Lead + IT; contain/restore; inform affected staff; log details; consider parent note if learning was notably impacted. |
| **L3 - Serious** | Material risk to rights/fairness/education outcomes OR personal data breach likely to harm individuals OR bias affecting a group. | Wrong grades for a year group; exposure of contact details/student work to unauthorised peers; algorithmic downgrading of female/ELL students. | **Immediate containment** (pause feature/system); notify Principal + DPO within **2 hours**; start GDPR assessment (72-hour clock); prepare comms to affected users/parents; notify vendor; plan remediation; consider AI Act consultation with provider. |
| **L4 - Critical (Crisis)** | Child safety risk, sensitive data at scale, severe discrimination, or incident likely to attract regulator/media attention. | Stranger access in online environment; biometric/health data leak; systemic discriminatory grading across cohorts/exams. | **All-hands:** disable system; safeguard pupils; inform Principal immediately; contact law enforcement (if safety), DPO starts GDPR notification; provider prepares AI Act serious-incident report; hold statement prepared; board informed; daily exec briefings. |

**Fast decision rules (use these if in doubt)**

* **Any child-safeguarding risk** → treat as **L4**.
* **Special category/biometric data exposed** to unauthorised parties → at least **L3** (often **L4** if broad).
* **Group-level unfairness/bias** affecting opportunities or grades → **L3/L4** depending on scope.
* **Exam integrity compromised** (e.g., proctoring failure) → **L3** minimum.

**5) Example Classifications (secondary school)**

* **11th-grade English grading bug (software error):** AI assigns lower marks to anxious/fidgeting students → **L3 Serious** (fairness + educational harm). Pause grading feature; teacher review/override; notify parents; remediate model; consider vendor’s AI Act duty.
* **Learning-app shows peers’ feedback/comments:** Confidentiality breach among one class → **L3 Serious** (GDPR risk). Take app offline; DPO risk assess; notify DPA within 72h if risk; inform affected families.
* **Minecraft/Roblox intruder contacts a student:** **L4 Critical** (child safety). Remove intruder; secure server; inform leadership and guardians; contact police; record safeguarding actions.
* **Short outage during homework window:** No sensitive data, temporary impact → **L2 Moderate**.
* **Timetable assistant double-books a room (caught early):** **L1 Minor**.

**6) Escalation & Timelines (service levels)**

* **Triage & classify:** within **1 hour** of report.
* **Containment:** L3/L4 within **2 hours**; L1/L2 same day.
* **Internal notifications:** L3/L4 to Principal & DPO **immediately**.
* **External notifications:**
  + **GDPR:** DPA within **72 hours** of awareness if risk is likely; affected individuals **without undue delay** for high risk.
  + **EU AI Act (provider duty):** serious incidents **asap** and **≤15 days**. School alerts provider the same day.
  + **Police/child protection:** **immediately** for safeguarding.

**7) Roles (who does what)**

* **Incident Coordinator (AI Governance Lead):** triage/classify; chair response.
* **DPO/Data Lead:** GDPR assessment/notifications; evidence log.
* **IT/EdTech:** containment, forensics, restoration, vendor liaison.
* **Principal/SLT:** decisions, parent/community messaging, regulator/media oversight.
* **Safeguarding Lead:** pupil protection steps, police liaison (if needed).
* **Comms Officer:** clear, accessible updates to staff/families.

**8) What to Log (minimum fields)**

* Unique ID, date/time discovered, reporter, system(s) involved, severity, description, data types, # affected, containment taken/time, notifications sent (who/when), remediation actions/owners/deadlines, closure date, lessons learned.

**9) Appendices**

**A. Contact tree & after-hours rota (what it is)**

A simple one-pager showing **who to call first**, then who they call next, **with phone numbers** and **who is on duty outside school hours**.  
**Why:** Incidents love weekends. You need a fast chain.

**Mini-template**

Primary: Incident Coordinator (AI Lead) – +34 XXX XXX XXX

Backup: IT Manager – +34 XXX XXX XXX

DPO/Data Lead – +34 XXX XXX XXX

Safeguarding Lead – +34 XXX XXX XXX

Principal – +34 XXX XXX XXX

After-hours week of <dates>: On-call = <Name>, Backup = <Name>

**B. Incident log template (fields should match Section 8)**

A standard form (Word/Sheet) everyone uses to record incidents consistently.  
**Why:** Evidence for GDPR/AI Act accountability and internal learning.

**Mini-template (table headings)**

ID | Date/Time | Reporter | System | Summary | Severity (L1–L4)

Data types | #Affected | Containment (time) | Notifications (who/when)

Actions/Owners | Deadline | Closure date | Lessons learned | Links (evidence)

**C. Decision tree (one-pager)**

A visual flow that helps staff make the **first classification** and route correctly.

**Text version that can be pasted into a box:**

Did personal data leave authorised access?

└─ Yes → Call DPO → Start GDPR 72h assessment → Continue below

└─ No → Next check

Is there any safety risk or contact from an unknown adult?

└─ Yes → Safeguarding Lead + Police NOW → L4

└─ No → Next check

Is this a high-risk AI function (grading/placement/interventions)?

└─ Yes → Pause feature/system → Notify AI Lead & Provider → Consider AI Act serious incident

└─ No → Next check

Scale/sensitivity high? Many students or special-category data (health/biometric)?

└─ Yes → L3 (or L4 if severe) → Principal + Parents informed

└─ No → L1/L2 → Log and fix

**D. Vendor expectations (contractual)**

A short list of **must-have clauses** you expect in every AI/EdTech contract.  
**Why:** During an incident, your rights depend on what’s in the contract.

Include language like:

* **Breach notice ≤ 24h** to the school (processor → controller duty).
* **No training on our data** (school data cannot be used to train vendor models without explicit, separate consent).
* **Access to system logs** for audits/incident investigation.
* **Human-in-the-loop & override** features enabled for school.
* **Decommissioning & data deletion SLAs** (clear timelines and confirmation of deletion/return upon contract end).