

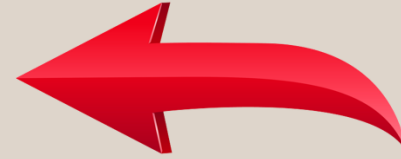
Tema 1. Introducción a entornos interactivos de aprendizaje

Autores: José Antonio Ruipérez Valiente (jruiperez@um.es), Manuel Jesús Gómez Moratilla (manueljesus.gomezm@um.es), Mariano Albaladejo González (mariano.albaladejog@um.es)



Contenidos

1.1 Entornos interactivos de aprendizaje



1.2 Datos educacionales

1.3 Explotación de datos educacionales

1.4 Ética y privacidad



UNIVERSIDAD
DE MURCIA

Tema 1. Introducción
a entornos interactivos
de aprendizaje

Analítica de aprendizaje y minería
de datos educacionales

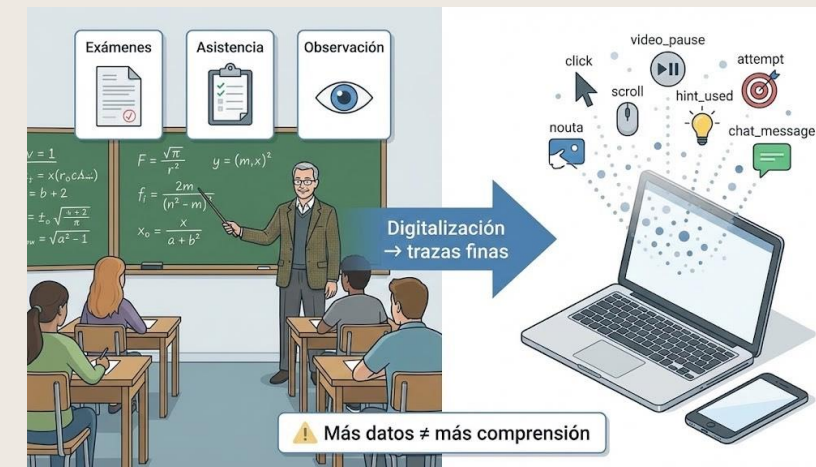
1.1 ¿Por qué esto importa?

- Learning Analytics (LA) y Educational Data Mining (EDM): usar datos educativos para comprender, mejorar y tomar decisiones
 - **NO** va a ser aplicar ML a notas de estudiantes
- Lo que veremos hoy:
 - Entornos que generan datos
 - Qué tipos de datos existen y cómo “se vuelven analizables”
 - Explotación de datos educacionales
 - Ética, privacidad y gobernanza (no como extra, como requisito)
- 3 ideas guía para todo el tema:
 - Sin contexto, los datos engañan
 - Medir $\neq$ aprender (actividad no siempre es aprendizaje)
 - Analítica sin acción/evaluación no produce impacto
- Objetivo: Habilitar puente entre tecnología/datos $\leftrightarrow$ aprendizaje $\leftrightarrow$ decisiones



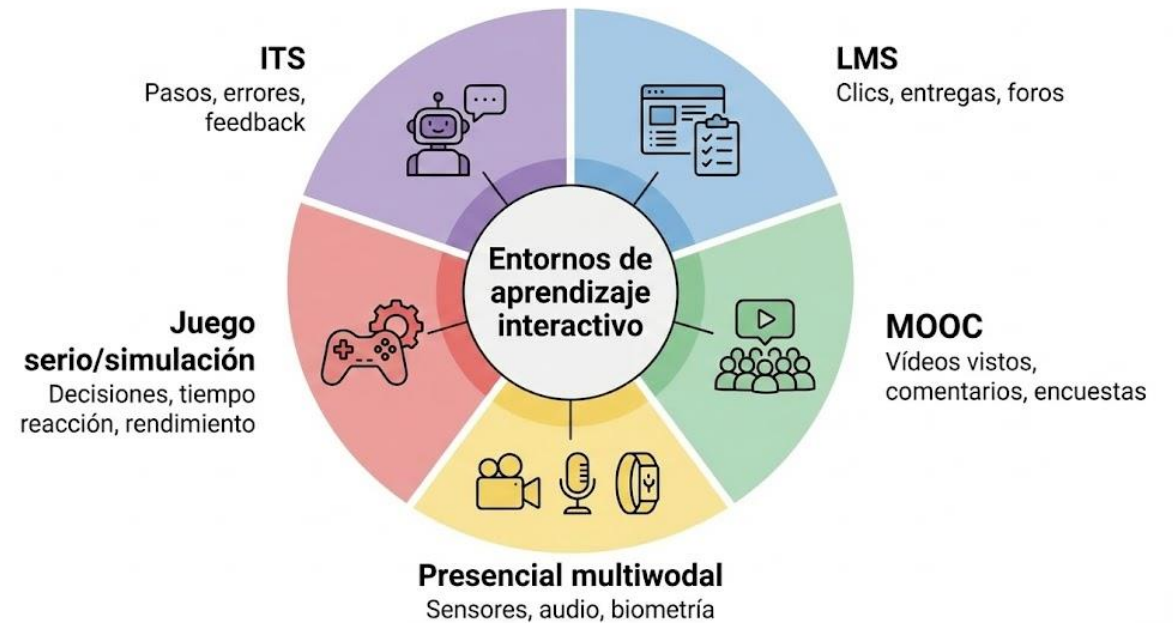
1.1 De aula tradicional a entornos mediados por tecnología

- Antes (aula tradicional):
 - Evidencias macro: examen final, asistencia, trabajos, observación...
 - Trazas “pocas” y costosas de obtener (observación humana)
- Ahora (entornos digitales):
 - Evidencias micro: eventos, secuencias, tiempos, elecciones, rutas
 - Escala: miles de estudiantes y millones de eventos
- Qué cambia realmente:
 - Granularidad: del resultado al proceso
 - Temporalidad: análisis casi en tiempo real
 - Personalización: rutas/feedback adaptativo
 - Automatización: alertas, recomendaciones...



1.1 Tipos de entornos interactivos (taxonomía práctica)

- LMS (Moodle/Canvas): participación, progreso, puntos de atasco
- MOOC (edX/Coursera...): heterogeneidad, abandono, escala
- ITS (tutores inteligentes): “evidencia fina” de conocimiento procedimental
- Juegos serios/simulaciones: stealth assessment y competencias complejas
- Multimodal/presencial instrumentado: sensibilidad y complejidad (privacidad + interpretación)



UNIVERSIDAD
DE MURCIA

Tema 1. Introducción
a entornos interactivos
de aprendizaje

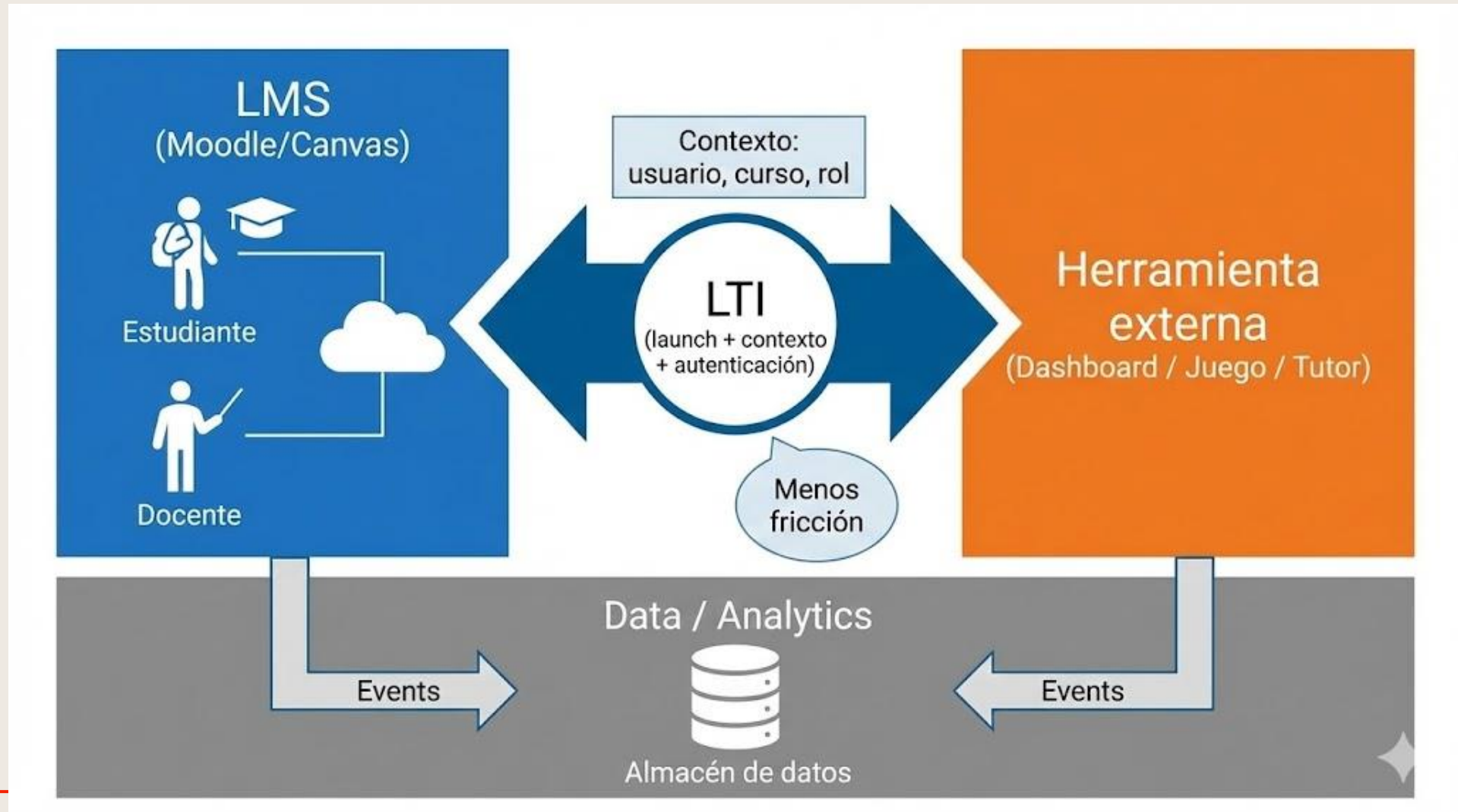
Analítica de aprendizaje y minería
de datos educativos

1.1 Interoperabilidad: por qué LTI existe

- Problemas reales sin interoperabilidad:
 - Múltiples herramientas → múltiples identidades → datos fragmentados
 - Pérdida de contexto: curso/rol/grupo/actividad
 - Baja adopción: “demasiado esfuerzo” para docentes/estudiantes
- LTI (idea funcional):
 - Lanzar herramienta externa desde el LMS con SSO
 - Transferir contexto (usuario, rol, curso, actividad)
 - Facilitar integración en el flujo docente (no “otra web más”)
- Beneficios para LA/EDM:
 - Trazabilidad consistente entre herramientas y semántica de eventos
 - Posibilita ecosistemas (LMS + simulador + dashboard + LRS)



1.1 Interoperabilidad: por qué LTI existe



UNIVERSIDAD
DE MURCIA

Tema 1. Introducción
a entornos interactivos
de aprendizaje

Analítica de aprendizaje y minería
de datos educativos

1.1 Ejercicio: Inicia el cambio de tu perspectiva

- Entra en Sakai como estudiante y recorre estas secciones del curso (si existen):
Resources/Recursos, Assignments/Tareas,
Gradebook/Calificaciones, Announcements/Anuncios
- Define una aplicación útil para el docente (elige solo una)
- Para esa aplicación, propón 3 eventos clave a recolectar (máx. 3):
 - Evento (nombre corto)
 - Dónde ocurre (herramienta Sakai)
 - Campos mínimos (user_id, course_id, timestamp, y el id del objeto: resource_id/assignment_id...)



Contenidos

1.1 Entornos interactivos de aprendizaje

1.2 Datos educativos



1.3 Explotación de datos educativos

1.4 Ética y privacidad



UNIVERSIDAD
DE MURCIA

Tema 1. Introducción
a entornos interactivos
de aprendizaje

Analítica de aprendizaje y minería
de datos educativos

1.2 Por qué ahora hay LA/EDM: volumen + proceso

- Motores del crecimiento:
 - Digitalización (LMS/online), movilidad (apps), masificación (MOOCs)
 - Instrumentación fina (ITS, juegos, simulaciones)
 - Sensores y multimodalidad (aula, remoto)
- Cambio conceptual:
 - De datos administrativos (nota final) a datos del proceso (trayectoria)
- “Nuevas preguntas” posibles:
 - ¿Dónde se atascan? ¿Qué secuencias llevan al éxito?
 - ¿Qué perfiles de aprendizaje emergen?
 - ¿Qué intervención funciona y para quién?



1.2 Tipos de datos educativos (fuentes, ejemplos y qué capturan)

Fuente de datos	Ejemplos típicos	Qué “captura” bien	Qué suele faltar / ambigüedades
Interacciones / logs (lo que hace)	abrir recurso, descargar, ver tarea, entregar, click en enlace, timestamps	Actividad observable, secuencias, regularidad, exposición a contenidos	“Abrir ≠ leer/entender”; tiempo en página ≠ atención; acciones fuera de la plataforma no se ven
Evaluación (resultado formal)	nota, rúbrica, feedback, estado de entrega (a tiempo/tarde)	Evidencia de rendimiento en hitos del curso	No explica el proceso (cómo llegó); puede mezclar aprendizaje con estrategia de examen
Encuestas / autoinforme (lo que dice)	satisfacción, motivación, carga, estrategias, percepción de dificultad	Significado subjetivo (lo que el log no muestra)	Sesgo de respuesta; baja frecuencia; memoria imperfecta
Datos del entorno / contexto	dispositivo, navegador, localización aproximada, horario, co-presencia (si aplica)	Contexto operativo (condiciones de acceso/uso)	Contexto pedagógico real (por qué) no siempre se infiere
Biométricos / fisiológicos (estado interno)	HR/EDA, seguimiento ocular, etc.	Señales posibles de carga/estrés/activación (con mucha cautela)	Inferencia débil sin calibración; alta variabilidad individual
Social / comunicación	posts, replies, menciones, redes de interacción, colaboración	Participación social, dudas, apoyo entre pares, dinámica de grupo	“Ver” un hilo ≠ participar; calidad del mensaje no se mide sola
Datos administrativos (quién es / estructura)	matrícula, grupo, rol, itinerario, calendario	Denominadores, segmentación, cohortes, planificación	No refleja aprendizaje; puede inducir sesgos si se usa mal

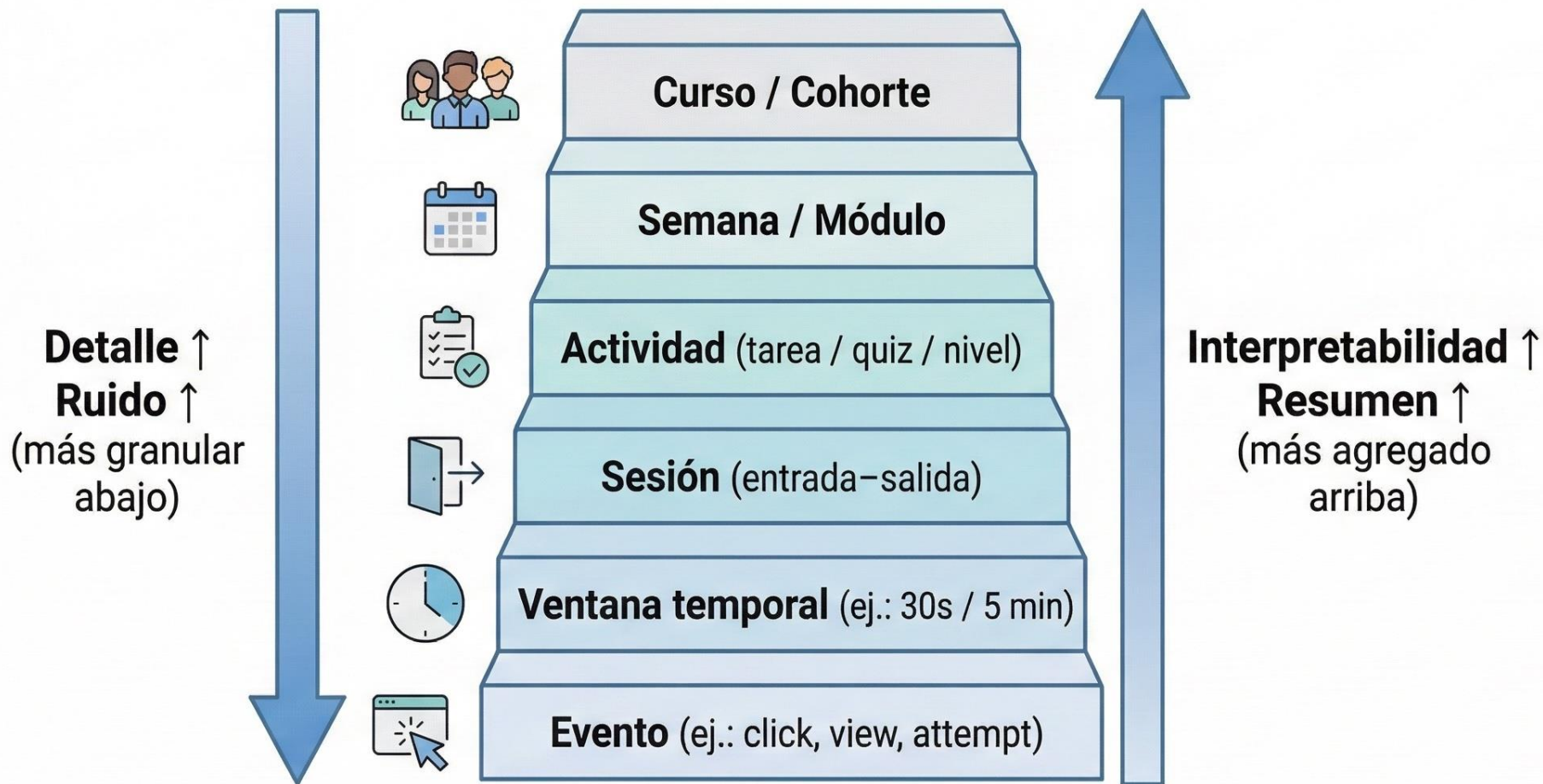


1.2 Granularidad: la unidad de análisis se diseña

- Niveles comunes: Evento (click, attempt), ventana temporal (30s, 5 min), sesión (entrada/salida), actividad (tarea, quiz, nivel), curso/cohorta (semanas)
- Pregunta decisiva: ¿qué decisión quiero soportar?
 - Intervención inmediata → granularidad fina
 - Rediseño del curso → agregación por actividad/semana
- Errores típicos:
 - “tiempo en recurso” = pestaña abierta (no estudio real)
 - “muchos intentos” = persistencia o frustración (necesita contexto)
 - agregaciones que borran subgrupos (promedios engañosos)



“Granularidad (grain size): la unidad de análisis se diseña”



Ejemplo: 'Tiempo en recurso'

Evento

"open_resource" + "close_resource"
→ tiempo = diferencia

⚠ pestaña abierta ≠ atención

Sesión

suma de tiempos en varios recursos
→ "estudio en bloque"
(más robusto, aún imperfecto)

Semana

promedio por semana → tendencia
(bueno para diseño del curso,
malo para intervención fina)

⚠ promedios ocultan subgrupos

El grano depende de la decisión



Intervención inmediata

necesita grano fino (evento/ventana)



Seguimiento del progreso

grano medio (sesión/actividad)



Rediseño del curso

grano alto (módulo/semana/curso)

1.2 Formatos: logs “a medida” vs xAPI (y por qué importa)

- Logs “ad hoc” (CSV/JSON propios, por ej, [Open edX](#)):
 - Pros: rápidos de implementar
 - Contras: semántica ambigua, difícil integrar, cambia con versiones
- Necesidades en ecosistemas reales:
 - combinar herramientas (LMS + juego + simulador + app)
 - mantener trazabilidad y contexto
- xAPI como patrón de evento:
 - Actor + Verbo + Objeto (+ Resultado, Contexto, Timestamp)
- Ejemplos (en texto, simple):
 - “Ana completó Quiz 2 con 80% en Curso LA101”
 - “Ana usó pista conceptual en Puzle 17 (nivel 2)”



1.2 Ejercicio: Midiendo la colaboración

- Contexto: tienes un trabajo en grupo (3-5 estudiantes) durante 3 semanas. Quieres estimar el nivel de colaboración real (no solo “dividirse tareas”) usando datos de distintas fuentes
- Define el tipo de colaboración a medir del proyecto
- Propón al menos 3 indicadores de 3 fuentes de datos diferentes:
 - Fuente (¿de dónde sale?)
 - Qué medirías (evento/variable)
 - Cómo lo convertirías en un indicador (regla simple o fórmula)
 - Qué limitación tiene (sesgo/ambigüedad)



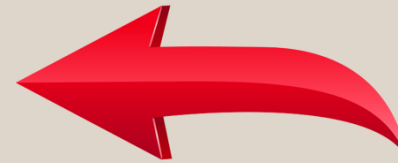
Contenidos

1.1 Entornos interactivos de aprendizaje

1.2 Datos educacionales

1.3 Explotación de datos educacionales

1.4 Ética y privacidad



UNIVERSIDAD
DE MURCIA

Tema 1. Introducción
a entornos interactivos
de aprendizaje

Analítica de aprendizaje y minería
de datos educacionales

1.3 Madurez: de “reporting” a intervención evaluada

- Nivel 0: datos para operación (logs para soporte/errores)
 - Nivel 1: descriptivo (qué pasa): participación, progreso, funnel
 - Nivel 2: diagnóstico (por qué): segmentación, patrones, cuellos
 - Nivel 3: predictivo (qué pasará): riesgo, probabilidad, alertas
 - Nivel 4: prescriptivo (qué hacer): recomendaciones y nudges
-
- Condición de calidad LA:
 - accionable (alguien puede actuar)
 - usable (encaja en flujo docente)
 - evaluable (medimos impacto)



1.3 Learning Analytics: definición y stakeholders

Learning Analytics

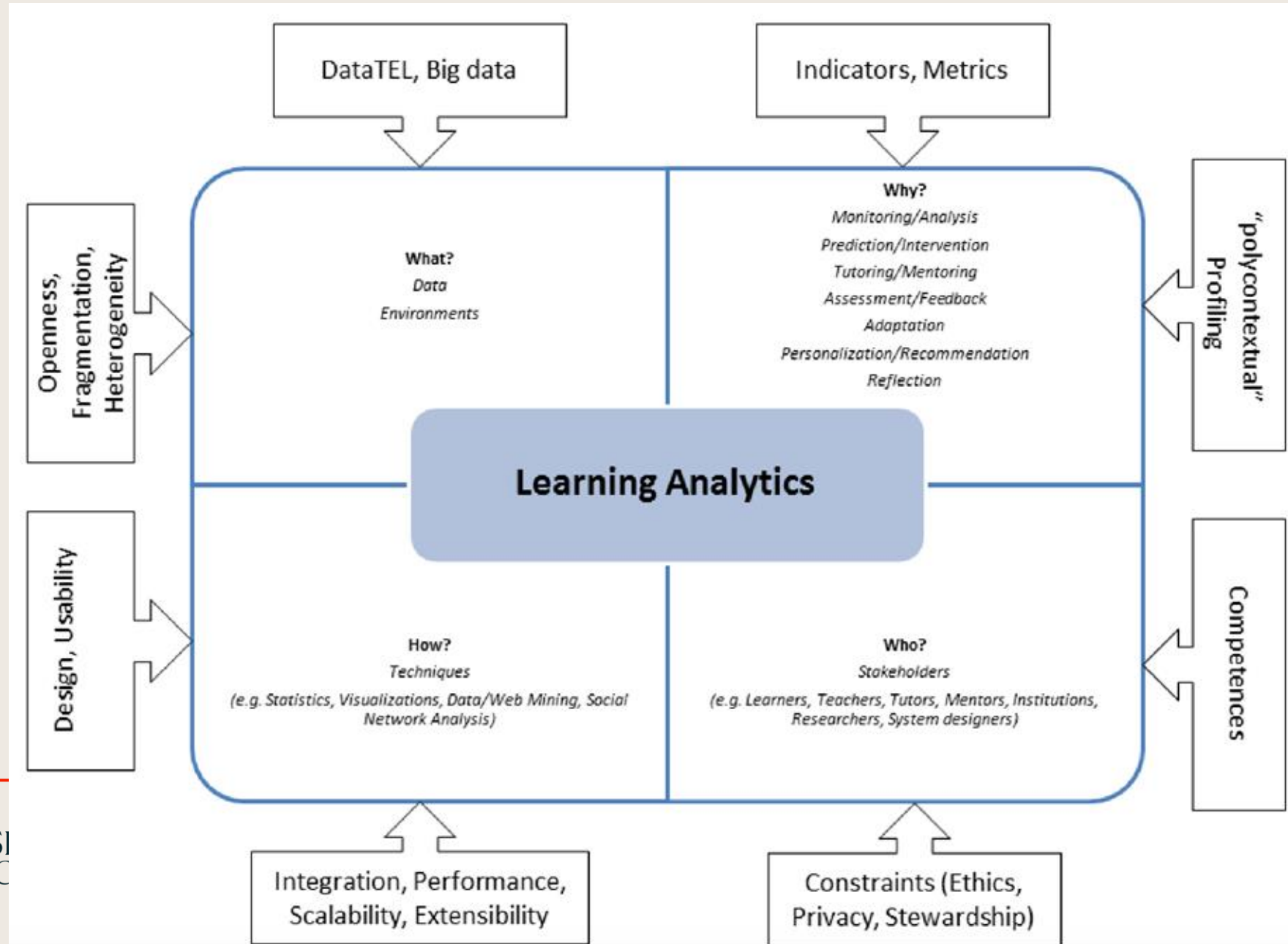
...collection, analysis and reporting of **data** about learners and their contexts, for purposes of understanding and optimising **learning** and the **environments** in which it occurs...

Fuente: [First Learning Analytics and Knowledge Conference](#)

- Stakeholders clave:
 - Estudiantes, docentes, instituciones e investigadores
- Tensiones frecuentes:
 - apoyo vs vigilancia
 - personalización vs equidad
 - predicción vs estigmatización



1.3 Modelo de referencia de Chatti et al. (2012)



1.3 Marco What/Who/Why/How (plantilla de proyecto LA)

- WHAT (datos/entorno)
 - LMS logs, quizzes, foros; juego/ITS eventos; multimodal...
- WHO (actores)
 - estudiantes, docentes, tutores, coordinadores
- WHY (objetivos)
 - monitorizar, detectar riesgo, apoyar autorregulación, mejorar diseño
- HOW (métodos)
 - descriptivo/visualización, reglas, ML, redes, secuencias



1.3 Aplicaciones LA: del “qué” al “qué hacer” (sólo algunos ejemplos)

- Monitorización (descriptivo):
 - funnel: inscritos → activos → entregan → completan
 - mapas de calor por recurso/semana
- Detección temprana (predictivo):
 - estudiantes en riesgo de abandono / bajo rendimiento
 - alertas calibradas (probabilidad + umbral)
- Feedback y autorregulación (estudiante):
 - metas, ritmo, recordatorios, recomendación de recursos
- Mejora de curso (diseño instruccional):
 - identificar recursos “problemáticos” (alta tasa de abandono/errores)
- Reglas de oro:
 - predicción → intervención → evaluación (si no, no vale)
 - mostrar “por qué” de forma comprensible



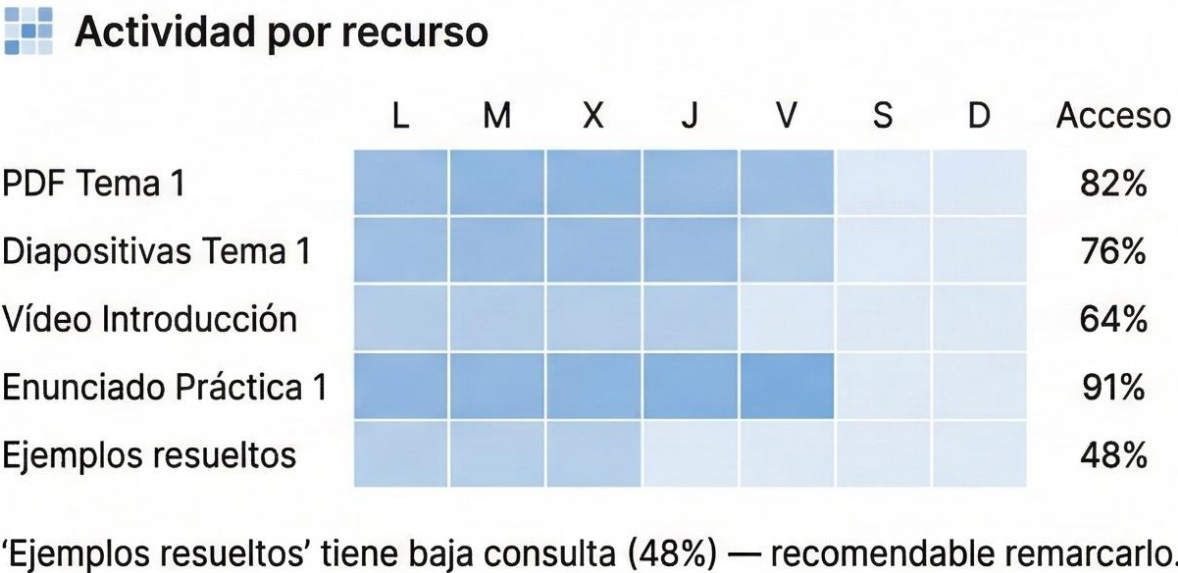
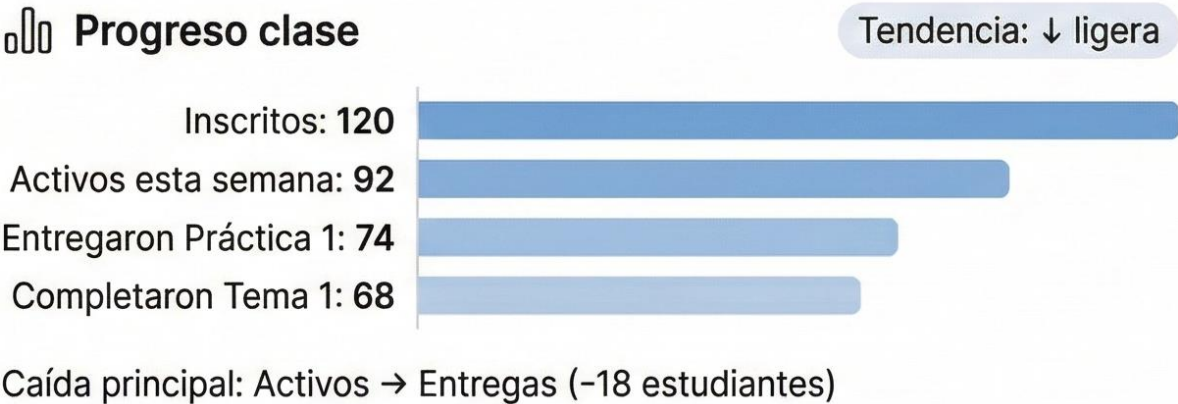
Dashboard docente (ejemplo)

Curso: LA101

Semana: 3

Grupo: Todos

💡 Accionable: sugerencias de intervención



Estudiantes en riesgo

Riesgo alto: 3 | Medio: 4 | Bajo: 0

	A. García	Sin acceso a recursos clave (7 días)	Ver detalle
	M. López	No entregó Práctica 1	Ver detalle
	S. Pérez	Baja regularidad (1 sesión/semana)	Ver detalle
	D. Ruiz	Acceso bajo a “Vídeo Introducción”	Ver detalle
	L. Martín	Resultado bajo en Quiz inicial	Ver detalle
	I. Sánchez	No inició “Ejemplos resueltos”	Ver detalle
	C. Torres	Participación foro baja	Ver detalle

- Sugerencias de intervención
- ✉ Enviar recordatorio a quienes no abrieron “Ejemplos resueltos”.
 - 📢 Publicar anuncio: aclaración de Práctica 1 + enlace directo.
 - 📅 Ofrecer tutoría breve a riesgo alto (3 estudiantes).

1.3 Áreas cercanas: LA vs EDM vs AIED (definiciones)

- Learning Analytics (LA): Analiza y comunica datos de estudiantes y contextos para **comprender y optimizar** el aprendizaje y la enseñanza (orientación a **decisión/acción**).
- Educational Data Mining (EDM): Aplica métodos de descubrimiento (KDD/ML/estadística) para **encontrar patrones y construir modelos** sobre datos educativos (orientación a **modelado/descubrimiento**).
- AI in Education (AIED): Diseña sistemas educativos que usan IA para **adaptar, tutorizar, recomendar o interactuar** con el estudiante (orientación a **intervención automatizada y diseño de sistemas**).
- Relación rápida:
 - EDM suele “producir” modelos/patrones
 - LA suele “operacionalizar” (dashboards/decisiones) esos hallazgos
 - AIED suele “actuar” en tiempo de aprendizaje (adaptación/feedback)



1.3 Áreas cercanas: LA vs EDM vs AIED (tabla comparativa)

Área	Pregunta típica	Producto/salida	Métodos típicos	Usuario principal
LA	¿Qué está pasando y qué hacemos ahora?	dashboards, alertas, indicadores accionables	visualización, métricas, modelos predictivos interpretables	docente, estudiante, institución
EDM	¿Qué patrones/modelos existen en los datos?	modelos predictivos, clustering, patrones de secuencia	ML, estadística, KDD, minería de secuencias	investigador/analista (y luego LA/AIED)
AIED	¿Cómo adaptamos/tutorizar automáticamente?	tutor inteligente, recomendador, feedback adaptativo	planificación, RL, modelos del estudiante, NLP, sistemas inteligentes	estudiante o docente (en interacción)



1.3 Ejercicio: Enmarca un caso con WWWH

- Caso: “En un MOOC hay alto abandono en la semana 2”
- WWWH completo (mín. 2 ítems por cuadrante)
- Define 2 KPIs accionables (no solo “visitas”)
- Propón 1 intervención concreta
- Diseña cómo evaluar impacto (A/B, cuasi, antes-después)



Contenidos

1.1 Entornos interactivos de aprendizaje

1.2 Datos educacionales

1.3 Explotación de datos educacionales

1.4 Ética y privacidad



1.4 Ética: por qué es estructural en educación

- Educación: alto impacto (oportunidades, permanencia, autoestima, becas)
- Riesgos éticos comunes:
 - estigmatización (“en riesgo” como etiqueta)
 - vigilancia y chilling effect (cambio de conducta)
 - sesgos (acceso digital, idioma, horarios, discapacidad)
 - automatización opaca (decisiones sin explicación)
- Principios guía:
 - proporcionalidad (solo lo necesario)
 - beneficencia (beneficio educativo claro)
 - no maleficencia (evitar daños plausibles)
 - justicia/equidad (no amplificar desigualdad)



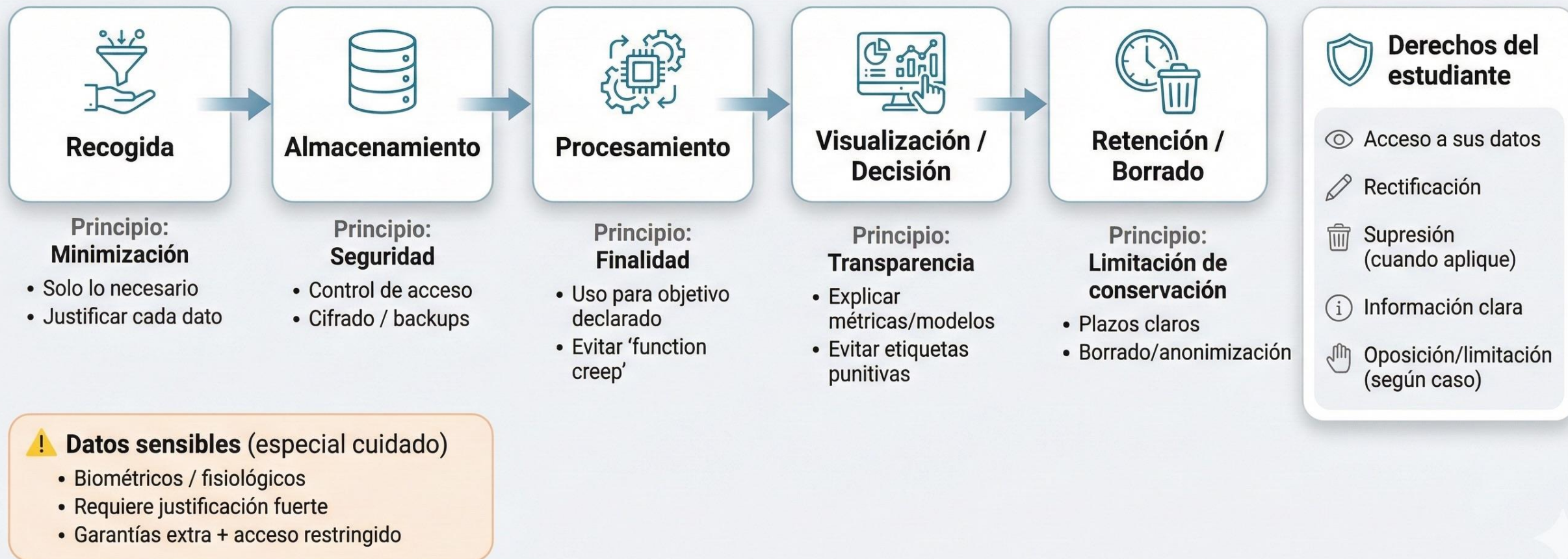
1.4 Privacidad (UE/GDPR): principios prácticos aplicados

- Principios clave (versión práctica):
 - minimización (menos dato, mejor)
 - finalidad (no reutilizar sin base)
 - transparencia (explicar qué/para qué)
 - seguridad (acceso, cifrado, auditoría)
 - retención limitada (borrar cuando no sea necesario)
- Consentimiento:
 - opt-in/opt-out y el problema de “consentimiento no libre” en contextos obligatorios
- Datos sensibles:
 - biométricos/fisiológicos → justificación fuerte + garantías extra



“Privacidad (UE/GDPR): principios prácticos aplicados”

Del diseño de datos a la acción docente (**privacy-by-design**)



1.4 IA responsable en LA/EDM: equidad + explicabilidad + humano

- Equidad:
 - evaluar rendimiento por subgrupos cuando sea apropiado
 - monitorizar drift (cambios de comportamiento/curso)
- Explicabilidad útil:
 - razones accionables (“baja regularidad + no intenta quizzes”)
 - evitar explicaciones técnicas incomprensibles para el usuario final
- Human-in-the-loop:
 - IA como apoyo; decisiones de alto impacto no totalmente automáticas
- Evaluación de impacto:
 - no basta “AUC alto”; importa “¿mejoró algo en educación?”



IA responsable en educación: caso de uso

Equidad + Explicabilidad + Humano-en-el-bucle + Evaluación de impacto



1) Datos del entorno (simulados)



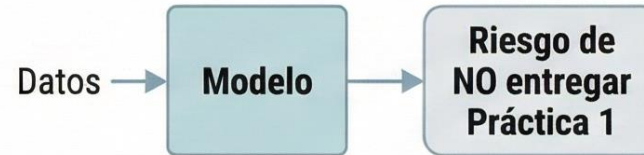
- **Eventos:** abrir recurso, ver tarea, entregar
- **Señales:** regularidad semanal, preparación, retrasos previos
- **Contexto:** semana del curso, deadline

Minimización: solo lo necesario

Datos simulados



2) Modelo: predice riesgo (no sentencia)



- **Salida** = probabilidad (0–100%)
- **Umbral** configurable (ej.: 70%)

Evitar fuga temporal (no usar datos del futuro)



3) Lista priorizada (con razones)

- A. García – 82%
 - No abrió recursos clave (7 días)
 - Vio la tarea pero no avanzó
 - Baja regularidad (1 sesión/semana)
- S. Pérez – 74%
- D. Ruiz – 62%
- L. Martín – 55%

Explicación práctica: señales observables

mejora continua



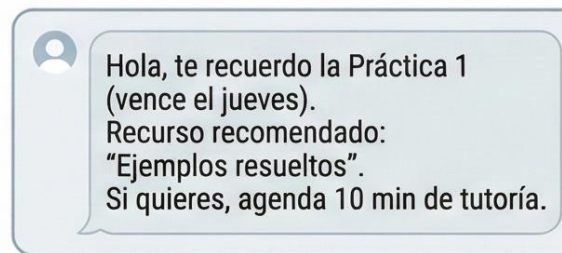
4) Docente revisa y decide



- La IA sugiere; el docente decide
- Se registra la decisión para mejorar el sistema



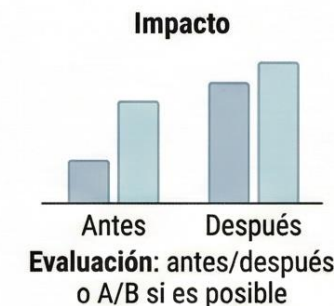
5) Intervención (apoyo, no castigo)



Evitar estigmatizar: no usar "estás en riesgo" en el mensaje



6) Evaluación de impacto y equidad



Equidad: Rendimiento del modelo por subgrupos (simulado)

- Grupo 1:** F1 = 0,78
- Grupo 2:** F1 = 0,74
- Grupo 3:** F1 = 0,61 (⚠️ revisar)

Si un grupo sale peor, revisar datos/umbral/intervención

Equidad: comprobar diferencias



Explicable



Supervisado (humano)



Evaluado (impacto real)



Justo (monitorizado)

1.4 Ejercicio 4: Auditoría ética y privacidad de un sistema

- Caso: sistema que detecta “desmotivación” datos:
 - logs + (opcional) audio
 - acción: mensajes automáticos + panel para docente
- Posibles riesgos
 - éticos
 - de privacidad
- Posibles mitigaciones
 - técnicas
 - de proceso
- Define quién ve qué (estudiante / docente / admin) y justifica



