Fundamentos de la Ciencia de Datos: PEC2

Víctor Suesta Arribas

Noviembre 2025

Pregunta 1

1.1. ¿Crees que en el contexto descrito en el artículo se cumplen las condiciones para asegurar que se realizaría una gobernanza del dato?, ¿y una gobernanza de la inteligencia artificial?

En el caso que describe Ercilla, creo que sí existen bases razonables para hablar de gobernanza del dato, aunque todavía no esté del todo madura. Por un lado, hay un marco legal claro que define funciones y procesos (LOPJ y regulación de los cuerpos generales, junto con el Real Decreto-ley 6/2023 sobre actuaciones automatizadas, proactivas y asistidas) y órganos de coordinación como el CTEAJE y el Directorio General de Información Tecnológica Judicial. Esto encaja bastante bien con la idea de tratar el dato como un activo que debe gestionarse de forma controlada a lo largo de todo su ciclo de vida.

Además, el texto insiste en la seguridad, la trazabilidad y la confidencialidad de la información judicial, lo que es coherente con los principios de privacidad y protección de datos que se trabajan en la asignatura. El problema es que aún quedan retos importantes, como la formación del personal o las diferencias entre comunidades autónomas. Por eso, diría que las condiciones para la gobernanza del dato están diseñadas, pero todavía en fase de despliegue.

En cuanto a la gobernanza de la IA, también vemos una base estructurada. El contexto está alineado con el AI Act europeo, que considera de alto riesgo muchos usos de la IA en justicia y obliga a aplicar salvaguardas específicas. Además, la política del CTEAJE define qué casos de uso de IA están permitidos y bajo qué condiciones, lo que se parece bastante a la idea de gobierno de la IA como conjunto de procesos y controles para garantizar un uso seguro, explicable y responsable. Aun así, persisten dudas sobre la responsabilidad ante daños y sobre cómo garantizar una explicabilidad real en contextos penales, de modo que la gobernanza de la IA también puede considerarse incipiente.

1.2. ¿Qué objetivos perseguiría la gobernanza del dato y de la IA en el departamento de justicia?

En un departamento de justicia, la gobernanza del dato debería perseguir algunos objetivos muy concretos. El primero es asegurar que jueces, letrados y funcionarios acceden a información consistente y actualizada (expedientes, plazos, notificaciones) sin contradicciones entre sistemas. Esto exige reducir silos, estandarizar estructuras y trabajar con datos de calidad a lo largo de todo su ciclo de vida.

El segundo objetivo es proteger la confidencialidad e integridad de los datos judiciales, especialmente en procedimientos sensibles, garantizando el cumplimiento de la normativa de protección de datos y de seguridad de la información. Finalmente, la gobernanza del dato debería permitir usar esa información para mejorar la gestión del servicio: detectar cuellos de botella, repartir cargas de trabajo y planificar recursos de forma basada en evidencia.

La gobernanza de la IA tiene objetivos complementarios centrados en el uso de esos datos dentro de sistemas inteligentes. En este caso, se trataría de evitar sesgos injustos, asegurar que la decisión final sigue siendo humana, y garantizar que las recomendaciones de los agentes de IA son trazables y explicables. También debe dejar claro quién es responsable de los errores y cómo se supervisa el comportamiento de los modelos a lo largo del tiempo.

1.3. ¿En cuál de las seis fases del modelo de madurez del data governance estaría el departamento de justicia después de implantar la iniciativa?

El modelo de madurez de gobierno del dato que se presenta en la asignatura distingue varias fases, desde la inexistencia de gobierno del dato hasta un estado optimizado con mejora continua y automatización amplia.

En el escenario descrito por Ercilla no partiríamos de cero, porque ya existen normas específicas, órganos de coordinación y una preocupación expresa por la seguridad y la trazabilidad. Sin embargo, tampoco parece realista hablar de una fase completamente optimizada, ya que la implantación de agentes de IA es gradual, la automatización total está limitada por diseño y aún hacen falta años de experiencia, métricas consolidadas y ajustes.

Por estos motivos, mi valoración es que el sistema se situaría en una fase 4 "avanzada": el dato y la IA se consideran elementos críticos, hay políticas y estructuras definidas, y se empieza a medir y auditar de forma más sistemática. El paso a una fase 5 requeriría consolidar la mejora continua y un uso aún más estable y extendido de estos mecanismos.

1.4. ¿Qué modelo de data governance seguiría el departamento de justicia después de implantar la iniciativa?

En los materiales se presentan distintos modelos de data governance, desde entornos indisciplinados hasta un modelo gobernado, en el que existen políticas corporativas claras, roles definidos y una gestión del dato coherente a nivel global.

Si la iniciativa de agentes de IA se implanta como propone Ercilla, el objetivo razonable sería avanzar hacia ese modelo gobernado. Hay apoyo institucional (Ministerio de Justicia, CTEAJE), un marco que obliga a compartir aplicaciones entre administraciones y una preocupación explícita por seguridad, calidad y trazabilidad de la información judicial.

En la práctica, esto implicaría que las decisiones sobre datos y sobre IA dejarían de tomarse de forma aislada en cada juzgado y pasarían a seguir políticas comunes, con plataformas compartidas y mecanismos de seguimiento. La gobernanza del dato y de la IA se integraría así en el funcionamiento habitual del sistema de justicia, y no quedaría limitada a proyectos puntuales.

Referencias utilizadas: [1, 2, 3, 4, 5, 6].

Pregunta 2

Una startup quiere lanzar un nuevo producto en el mercado internacional que calcule el riesgo de reincidencia:

2.1. ¿Cómo debería competir la startup con los datos bajo el paraguas de la privacidad y la seguridad?

En un producto que predice riesgo de reincidencia, la forma de competir no puede ser "cuantos más datos mejor", sino convertir la confianza en la principal ventaja competitiva. Los materiales de seguridad insisten en que una brecha de datos afecta directamente a la reputación, a la pérdida de clientes y a posibles sanciones, de modo que la seguridad deja de ser sólo un tema técnico y pasa a ser un tema de negocio. Algo parecido ocurre con el gobierno del dato: si tratamos la información como un activo, con impacto económico y estratégico, la empresa tiene incentivos para gestionarla bien a lo largo de todo su ciclo de vida.

Para la startup, esto implica aplicar privacidad por diseño y por defecto: usar sólo los datos estrictamente necesarios (minimización), configurar por defecto accesos restringidos y tiempos de conservación limitados, e incorporar técnicas como pseudonimización o enmascaramiento para reducir riesgos cuando se trabaja con datos penales sensibles. Desde el punto de vista comercial, puede presentarse como un producto "seguro y respetuoso" con las personas, siguiendo la idea de ethics by design que proponen organismos como el World Economic Forum.

Además, la empresa debería contar con un programa de gobierno del dato y de la IA: clasificación de información, control de accesos, auditoría de modelos y revisión periódica de sesgos, tal y como sugieren los marcos de gobierno de la IA de IBM. En el ámbito penal, también es clave que el modelo sea explicable y justo, porque una evaluación de riesgo puede influir en decisiones que afectan a la libertad de una persona. Por tanto, competir "con datos" aquí significa ofrecer un sistema seguro, transparente y alineado con principios éticos, más que explotar la información sin límites.

2.2. ¿Qué normativas de datos deberían tener presentes para Estados Unidos y Europa?. Cumplimenta la respuesta con fuentes de información adicionales a los recursos del aula.

En Estados Unidos, el marco de protección de datos es fragmentado y sectorial. El material de seguridad recuerda que existen distintas leyes según el tipo de dato: por ejemplo, HIPAA para datos de salud, otras normas para datos financieros o de menores, y leyes estatales de notificación de brechas y uso de cifrado. A esto se añaden regulaciones como la CCPA en California, que da a los consumidores derechos de acceso y supresión y obliga a informar sobre cómo se usan sus datos. Para la startup, esto significa que debe identificar qué tipo de datos emplea (penales, clínicos, financieros...) y ver qué leyes sectoriales y estatales le aplican, adoptando buenas prácticas de seguridad y transparencia incluso donde la normativa sea menos exigente.

En Europa, el enfoque es más homogéneo y se basa en el Reglamento General de Protección de Datos (GDPR). Este reglamento se apoya en principios como licitud, transparencia, limitación del propósito, minimización, calidad y seguridad, y añade la idea de responsabilidad proactiva: no basta con cumplir, hay que poder demostrar que se cumple. En un sistema de riesgo de reincidencia, la startup tendría que justificar la base jurídica del tratamiento (por ejemplo, interés público en el ámbito penal), hacer una evaluación de impacto en protección de datos y aplicar medidas técnicas y organizativas robustas.

Además, el producto se encuadraría en la categoría de alto riesgo del Reglamento Europeo de IA, al utilizarse en contextos de justicia y decisiones que afectan a derechos fundamentales. Esto implica requisitos adicionales sobre calidad de los datos, gestión de riesgos, documentación, supervisión humana y transparencia hacia las personas afectadas. En el caso español, la Estrategia de IA 2024 refuerza esta línea, subrayando la importancia de una IA "confiable" y de autoridades específicas de supervisión como la AESIA. En resumen, la startup debe diseñar su producto pensando desde el inicio en cumplir tanto el mosaico normativo estadounidense como el marco mucho más estructurado de la Unión Europea.

2.3. Síntesis

En un producto de predicción de reincidencia criminal, la ventaja competitiva de la startup no debería estar en "exprimir" más datos, sino en demostrar que sabe gobernarlos bien. Esto pasa por integrar la privacidad y la seguridad en el diseño del sistema, aplicar principios éticos y de equidad y ofrecer modelos explicables y auditables. Al mismo tiempo, debe ser capaz de operar en dos marcos regulatorios muy distintos: uno estadounidense, disperso y sectorial, y otro europeo, centrado en derechos fundamentales, el GDPR y el AI Act. Si la empresa consigue cumplir el estándar más exigente y hacerlo visible para jueces, administraciones y ciudadanía, la protección de derechos y el buen gobierno algorítmico se convierten en su principal argumento frente a otros competidores.

Referencias utilizadas: [1, 2, 3, 7, 8, 9, 5, 10, 11, 6].

Pregunta 3

Para la implantación de la iniciativa de la startup y dentro del contexto de la GDPR y la AI Act:

3.1. ¿Cómo relacionarías los conceptos de privacidad, seguridad y cumplimiento legal?

En los materiales de la asignatura, privacidad, seguridad y cumplimiento legal se presentan como tres piezas que rodean al dato y que tienen que ir coordinadas. Yo lo entiendo así: la privacidad marca los límites sobre cómo se pueden tratar los datos personales (principios del RGPD: licitud, minimización, limitación de la finalidad, etc.); la seguridad se ocupa de aplicar medidas técnicas y organizativas para proteger confidencialidad, integridad y disponibilidad; y el cumplimiento legal se asegura de que todo esto esté alineado con el RGPD, la AI Act y la normativa sectorial que toque en cada contexto.

En una startup que ofrece un sistema de predicción de riesgo de reincidencia, esta relación se vuelve crítica. No basta con tener "el dato protegido": hay que justificar por qué se usan ciertos datos penales o sociales (privacidad), demostrar que sólo accede quien debe y que el sistema está disponible cuando se necesita (seguridad), y documentar todo el ciclo de vida del dato y del modelo para poder superar auditorías y evaluaciones de impacto (cumplimiento). El caso COMPAS en Estados Unidos, analizado en el material de Loomis, muestra lo que pasa cuando un modelo de riesgo es opaco: aunque haya medidas de seguridad, si la herramienta no es transparente ni explicable puede chocar con derechos básicos como la defensa o la igualdad ante la ley. Algo parecido ocurre con los agentes de IA en justicia que describe Ercilla: si no se conectan bien las tres capas (privacidad, seguridad y cumplimiento), el riesgo jurídico y reputacional es muy alto.

3.2. ¿Qué principios aplicarías en tu programa de seguridad de datos?

El programa de seguridad de datos de la startup debería seguir los principios que se explican en el módulo de seguridad. En primer lugar, garantizar la confidencialidad y el cumplimiento normativo durante todo el ciclo de vida del dato: recoger sólo los datos necesarios para el modelo de reincidencia, definir bien la base jurídica del tratamiento (por ejemplo, interés público en el ámbito penal) y fijar plazos de conservación ligados a la finalidad y no indefinidos.

En segundo lugar, minimizar el riesgo de acceso no autorizado o uso indebido. Esto implica controles de acceso basados en roles (jueces, equipos técnicos, administración del sistema), autenticación fuerte y separación clara entre entornos de desarrollo, pruebas y producción, usando datos anonimizados o pseudonimizados fuera de producción.

El tercer principio es reducir el impacto de una posible brecha. Aquí entrarían el cifrado fuerte de bases de datos y copias de seguridad, un plan de respuesta a

incidentes y procedimientos claros para notificar a autoridades y organizaciones usuarias si algo falla.

Por último, documentar los controles y poder demostrar su efectividad (accountability). Esto significa tener registros de actividades de tratamiento, informes de evaluación de impacto (DPIA) para el sistema de riesgo de reincidencia y auditorías periódicas donde se revisen tanto la seguridad técnica como el comportamiento del modelo (precisión, posibles sesgos, estabilidad).

3.3. ¿Qué mejores prácticas aplicarías en el ámbito de la privacidad y seguridad de los datos?

Las mejores prácticas que se recogen en el módulo se pueden adaptar bastante bien al caso de la startup. Lo primero sería construir un inventario y una clasificación de activos de información: qué datos personales se tratan (historial penal, datos de contexto social, informes médicos o psicológicos, etc.), de qué fuentes vienen y en qué sistemas se almacenan. Esa clasificación permite aplicar controles más fuertes justo donde más se necesitan.

Otra práctica clave es aplicar privacidad por diseño y por defecto. Esto implica pensar desde el inicio qué variables son realmente necesarias para predecir la reincidencia, configurar por defecto los accesos de forma restrictiva, y documentar los riesgos mediante una DPIA antes de desplegar el sistema en un entorno real.

También es importante diferenciar la protección según la criticidad: las bases de datos con expedientes y las versiones en producción del modelo deberían tener el nivel máximo de protección (cifrado, monitorización reforzada, copias de seguridad frecuentes), mientras que otros datos menos sensibles pueden gestionarse con controles más ligeros.

Finalmente, el módulo insiste mucho en el factor humano. En una startup pequeña, un error de configuración en la nube o el uso de servicios externos sin supervisión puede abrir la puerta a fugas de datos. Por eso la formación del equipo en RGPD, seguridad básica y riesgos específicos de los modelos de IA no es un "extra", sino una parte central de la estrategia de seguridad.

3.4. ¿Qué elementos clave de la seguridad y la privacidad destacarías?

Los elementos clave de la seguridad que se explican en la asignatura giran en torno al modelo CIA: confidencialidad, integridad y disponibilidad. La confidencialidad exige que sólo personas y sistemas autorizados puedan ver los datos; la integridad, que los datos y las salidas del modelo no se alteren sin control; y la disponibilidad, que el sistema funcione cuando jueces o técnicos lo necesiten. A esto se suman herramientas como el cifrado, la gestión de identidades y accesos, los logs de actividad y los planes de continuidad de negocio.

En privacidad, el eje es la privacidad por diseño y por defecto: integrar la protección de datos en las decisiones de arquitectura (qué datos se almacenan, cómo se anonimiza o pseudonimiza, qué se muestra en las interfaces) y configurar de entrada la opción más protectora para la persona presa o en libertad

condicional. El gobierno del dato y la auditoría completan el cuadro: políticas claras de quién puede hacer qué con los datos, revisiones periódicas del modelo y de sus efectos, y capacidad para reconstruir cómo se generó una determinada recomendación en caso de recurso o conflicto.

La transparencia y la explicabilidad se sitúan a caballo entre ética, privacidad y seguridad. El ejemplo de Loomis muestra que un modelo cerrado puede ser legalmente problemático incluso si "cumple" en otros aspectos. Por eso, en un sistema de riesgo de reincidencia es fundamental poder explicar qué variables han influido en cada puntuación y cuáles son las limitaciones del modelo, algo que también exige la AI Act para sistemas de alto riesgo.

3.5. ¿Qué marco tecnológico aplicarías?

A partir del marco tecnológico que se presenta en el módulo, yo entiendo que la startup debería pensar su arquitectura de seguridad y privacidad por capas. En la base estaría la infraestructura: desplegar el sistema en centros de datos o nubes que permitan cumplir RGPD y AI Act (por ejemplo, nube europea o infraestructuras controladas por la administración de justicia), con segmentación clara de entornos y protección de redes y sistemas.

Encima de esa base estaría la capa de gobierno del dato y MLOps: herramientas para descubrir y clasificar datos personales, flujos reproducibles de limpieza y preparación, registro de versiones de modelos y de sus datos de entrenamiento, y monitorización del rendimiento y del posible "drift" del modelo en el tiempo. Todo ello con logs detallados que permitan reconstruir decisiones importantes.

Una tercera capa sería la de cumplimiento y gobierno de la IA: catálogo de sistemas de IA de la startup, evaluación del riesgo de cada uno, documentación tipo "model card" explicando finalidad y límites del sistema de reincidencia, y evaluaciones de impacto en protección de datos y derechos fundamentales.

Por último, una capa de interacción humana y transparencia: interfaces para jueces y profesionales que no sólo muestren una puntuación, sino también una explicación comprensible, avisos cuando la calidad de los datos no sea suficiente, y un canal para registrar desacuerdos o correcciones humanas que luego se puedan usar para mejorar el sistema.

En conjunto, este marco tecnológico permite que la parte técnica, la parte jurídica y la parte ética de la solución vayan alineadas, que es justo lo que piden el RGPD, la AI Act y los materiales de la asignatura para sistemas de IA en contextos tan sensibles como la justicia penal.

Referencias utilizadas: [2, 3, 12, 4, 13, 6, 14].

Pregunta 4

Para la implementación de la iniciativa de la startup, que ha de cumplir con la legislación vigente, la creación de un nuevo marco ético cobra relevancia.

4.1. ¿Qué principios orientadores de la acción de un buen profesional crees que debería seguir?

En esta iniciativa, la startup quiere usar agentes de IA para apoyar decisiones muy sensibles (por ejemplo, en justicia penal). Yo plantearía el marco ético combinando principios del profesional de datos y de la organización, siguiendo las líneas del material de Ética en ciencia de datos y de Ética y big data.

En primer lugar, pondría el respeto a los derechos fundamentales y a la dignidad por encima de cualquier objetivo técnico. Si el sistema influye en decisiones sobre libertad, clasificaciones penitenciarias o acceso a recursos, no se puede tratar a las personas como "registros", sino como sujetos de derechos. La regla práctica sería: si una decisión no podría justificarse sin el algoritmo, tampoco es aceptable esconderse detrás del algoritmo para tomarla.

También tendría como principios centrales la equidad y la no discriminación. Esto implica revisar qué variables se usan, analizar cómo se comporta el modelo por grupos y estar dispuesto a ajustar el modelo (aunque pierda algo de precisión) si se detectan diferencias injustificadas en falsos positivos o falsos negativos para ciertos colectivos. En el ámbito penal, pequeñas diferencias técnicas pueden traducirse en impactos muy serios sobre personas reales.

Otro eje sería la transparencia y la explicabilidad. El modelo no puede ser una caja negra inapelable: hay que poder explicar qué datos se han utilizado, qué transformaciones se han hecho y qué factores han pesado más en cada recomendación. Esa explicación debe ser comprensible para jueces y equipos técnicos, pero también para la persona afectada, en línea con las exigencias de IA confiable que recogen tanto la AI Act como las estrategias de ética "by design".

Finalmente, incluiría de forma explícita la privacidad (minimización de datos, uso legítimo, límites a la reutilización), la responsabilidad y la rendición de cuentas (dejar rastro de las decisiones técnicas, definir responsables humanos claros) y la integridad profesional. El científico de datos no solo construye modelos: también debe saber decir "no" cuando un uso de los datos entra en conflicto con estos principios, aunque sea legal o técnicamente posible.

4.2. ¿Qué puntos de decisión ética definirías?

En Ética y big data se propone trabajar con puntos de decisión ética a lo largo del ciclo de vida del dato. Aplicado a la startup, yo estructuraría al menos cinco momentos clave donde detenerse y revisar el proyecto con gafas éticas.

El primer punto sería antes de empezar el proyecto (justificación). Aquí la pregunta es si realmente tiene sentido construir este sistema: qué problema quiere resolver, si existen alternativas menos intrusivas y si el beneficio esperado compensa los riesgos para los derechos. Este paso conecta con la idea de "ética

by design": no se trata solo de cumplir después, sino de decidir si el proyecto merece la pena desde el inicio.

El segundo punto estaría en la captura y el mantenimiento de los datos. En esta fase habría que revisar cuál es la base jurídica del tratamiento, si se están recogiendo solo los datos necesarios o se acumulan otros "por si acaso", y qué sesgos arrastran los datos históricos (por ejemplo, decisiones pasadas más duras sobre determinados colectivos). Aquí se combinan la mirada ética y la de protección de datos: no todo lo que está almacenado es legítimo usarlo.

Un tercer punto de decisión aparece en el diseño del modelo y en la elección de variables. En este momento hay que descartar variables o combinaciones que puedan funcionar como proxies de características sensibles, decidir el equilibrio entre precisión, explicabilidad y equidad, y planificar cómo se explicará el modelo a los distintos tipos de usuarios. Es donde los valores declarados (equidad, transparencia) se traducen en decisiones técnicas concretas (métricas de fairness, restricciones, documentación).

El cuarto punto estaría justo antes del despliegue general. La organización debería revisar los resultados de pilotos controlados, analizar efectos por grupos, simular escenarios adversos y decidir con qué condiciones mínimas se permite el uso real del sistema. También es un momento para implicar a perfiles no técnicos (juristas, responsables de cumplimiento, incluso representantes de personas afectadas) y completar las evaluaciones de impacto en protección de datos y derechos.

Por último, fijaría un punto de decisión cada vez que se quieran reutilizar datos o modelos, y en el cierre del sistema. Reusar datos de justicia penal para otros fines, o prolongar el uso de un modelo más allá de lo razonable, no debería hacerse de forma automática. Hay que volver a plantearse si el nuevo uso encaja con las finalidades originales, qué impacto puede tener y si es preferible anonimizar o eliminar determinadas cohortes.

De esta forma, el marco ético de la startup no se queda en una declaración general, sino que se integra en momentos concretos del ciclo de vida del dato y del modelo, conectando con las ideas de ética "by design", IA confiable y responsabilidad profesional que recogen tanto los materiales de la asignatura como las estrategias europeas de IA.

Referencias utilizadas: [3, 12, 10, 6, 15, 11, 14].

Pregunta 5

Destaca dos de las intervenciones del segundo debate:

5.1. Intervención de Andrés Miguel Iriarte Padilla

Me ha parecido interesante la intervención de Andrés porque conecta dos marcos de referencia que hemos visto también en la asignatura: los seis principios de la Comisión Europea para sistemas de IA y los tres principios de IBM sobre ética de la IA. Él muestra que los principios de IBM (respeto por las personas, beneficencia, justicia) pueden entenderse como incluidos dentro del marco más amplio de la Comisión (agencia humana, privacidad y gobernanza de datos, equidad, bienestar, transparencia, responsabilidad y supervisión). Esto encaja bastante bien con la idea de "Ethics by design": no inventar un código nuevo para cada proyecto, sino apoyarse en marcos que ya están consolidados a nivel europeo.

Además, valoro que dé importancia a la privacidad y a la gobernanza de datos como parte del marco ético, y no solo a la típica pareja "privacidad + no discriminación". Que incluya también el bienestar individual, social y ambiental va en la línea de lo que hemos trabajado en la asignatura: los proyectos de ciencia de datos y de agentes inteligentes no solo deben ser técnicamente correctos, sino también aportar valor a la sociedad y reducir posibles daños.

Aun así, echo en falta un paso más hacia la práctica. Andrés se queda en la comparación de marcos de principios, pero no llega a concretar qué implicaría, por ejemplo, "respetar la agencia humana" cuando un agente de IA ayuda a evaluar el riesgo de reincidencia. En el contexto de la PEC, habría sido útil que enlazara esos principios con decisiones concretas: qué variables usar, cómo explicar el modelo a las personas afectadas, o qué límites poner al uso del sistema en justicia penal. En resumen, comparto su enfoque general, pero creo que podría haber bajado algo más al terreno práctico, que es donde suelen aparecer los dilemas reales.

5.2. Intervención de Aide Pizano Escalante

La intervención de Aide se centra en otra dimensión del problema: no tanto qué principios deben guiar a los agentes inteligentes, sino quién debería regularlos y cómo. Ella defiende la creación de un organismo regulador específico para agentes de IA, similar a las agencias de protección de datos, y señala bien los ámbitos donde estos sistemas ya tienen impacto (finanzas, trabajo, salud, educación). Me parece acertado que no se quede solo en la idea de "más regulación", sino que concrete tres funciones: normativa (criterios de transparencia y equidad), supervisora (revisión de algoritmos y detección de sesgos) y educativa (formación y buenas prácticas). Esta combinación encaja bastante con lo que hemos visto como "gobierno de la IA": no es solo cumplir la ley, sino tener procesos y roles claros para usar la IA de forma responsable.

También es positivo que conecte su propuesta con el contexto europeo actual, mencionando la AI Act y la necesidad de coordinación internacional para evitar lagunas legales. Esto recuerda al debate sobre la responsabilidad en IA que hemos visto en los materiales: aunque existan principios y reglamentos, si no hay una arquitectura institucional clara, el control efectivo de los sistemas se queda corto.

Como matiz, creo que su idea de un organismo "global" puede ser difícil de llevar a la práctica. En la realidad, las autoridades de protección de datos funcionan por países o regiones, coordinadas entre sí, y probablemente con los agentes de IA pase algo parecido (autoridades nacionales y europeas, más que una única agencia mundial). También echo de menos algo más de énfasis en el papel de las propias organizaciones: un regulador fuerte es importante, pero si las empresas y administraciones no tienen su propio marco de gobierno de la IA (comités éticos, responsables internos, políticas de documentación y auditoría), el regulador siempre llegará tarde. Aun así, su intervención aporta una visión clara sobre la necesidad de supervisión independiente, que complementa bien la parte más "de principios" del debate.

5.3. Síntesis personal

Comparando ambas intervenciones, yo las veo como dos piezas del mismo puzzle. Andrés pone el foco en el nivel de principios: agencia humana, equidad, privacidad, bienestar, transparencia y responsabilidad. Aide se centra más en la capa institucional: quién vigila a los agentes inteligentes y en qué ámbitos debe actuar esa supervisión. Desde mi punto de vista como estudiante de ciencia de datos, las dos son necesarias para hablar en serio de gobernanza de agentes IA, sobre todo en contextos como la justicia penal o la evaluación del riesgo de reincidencia.

Como punto crítico común, sí creo que las dos intervenciones podrían haber conectado algo más con el caso concreto que trabajamos en la PEC (sistemas de puntuación de riesgo, decisiones sobre libertad o medidas alternativas, posibles sesgos en colectivos vulnerables). Aun así, me han servido para estructurar mi propia posición: los agentes inteligentes solo deberían desplegarse en estos ámbitos cuando haya, por un lado, principios éticos claros que orienten su diseño, y por otro, estructuras de gobierno internas y externas capaces de supervisar su uso, detectar sesgos y corregir desviaciones a tiempo. Sin esa doble capa (principios + instituciones), el riesgo de que la IA consolide desigualdades o erosione derechos es demasiado alto.

Referencias utilizadas: [3, 12, 15, 14, 16, 6].

Referencias

[1] M. Pérez González. *El ciclo de vida del dato*. Universitat Oberta de Catalunya (UOC).

[2] D. Cabanillas Barbacil. *Seguridad y privacidad de datos*. Universitat Oberta de Catalunya (UOC).

[3] A. Pita Lozano. *Ética en la ciencia de datos*. Universitat Oberta de Catalunya (UOC), 2025.

[4] J. Ercilla. *Agentes de Inteligencia Artificial en la Administración de Justicia española: Marco normativo y potencial de implementación*. Laleynext.es, 28 de nov. de 2024.

[5] T. Mucci y C. Stryker. *¿Qué es el gobierno de la IA?* IBM, 8 de abr. de 2025.

[6] Parlamento Europeo y Consejo de la Unión Europea. *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo de 13 de junio de 2024 relativo a la inteligencia artificial y por el que se modifican determinados actos legislativos de la Unión (Ley de Inteligencia Artificial o AI Act)*. Diario Oficial de la Unión Europea, 13 de jun. de 2024.

[7] M. Korolov. *Cinco casos de uso "top" para los agentes de IA en la empresa*. CIO, 19 de mar. de 2025.

[8] M. Padrón. *Agentes de IA: qué son y cómo implementarlos*. KPMG Tendencias, 24 de jul. de 2025.

[9] D. Moody y P. Walsh. *Measuring the Value of Information: An Asset Valuation Approach*. 1997.

[10] Gobierno de España. *España Digital 2026. Estrategia de Inteligencia Artificial 2024*. Gobierno de España, 2024.

[11] World Economic Forum. *Ethics by Design: An organizational approach to responsible use of technology*. World Economic Forum, 2020.

[12] A. Valencia. *Ética y big data*. Universitat Oberta de Catalunya (UOC).

[13] E. Linde, P. Solar Calvo y P. Lacal Cuenca. *Ley de Inteligencia Artificial y Derecho Penal. El caso State v. Loomis como ejemplo*. LegalToday.com, 26 de dic. de 2023.

[14] IBM. *¿Qué es la ética de la IA?* IBM.

[15] European Commission. *Ethics by Design and Ethics of Use Approaches for Artificial Intelligence*. European Commission, 2021.

[16] A. Gomstyn y A. Jonker. *¿Nuevos riesgos éticos por cortesía de los agentes de IA? Los investigadores están con el caso*. IBM, 22 de abr. de 2025.