

El Consejo de Europa fija el estándar global en IA y derechos humanos: lo que toda organización debe saber ahora

Por Iván Martínez, CEO de Intedya

El Manual del CDDH, publicado en abril de 2026, traduce el Convenio Marco sobre IA en obligaciones concretas para los sectores público y privado. Una hoja de ruta que los profesionales de compliance no pueden ignorar. En un momento en que la inteligencia artificial ya no es una tecnología del futuro sino una realidad operativa cotidiana en juzgados, hospitales, servicios sociales y fronteras, el Consejo de Europa acaba de publicar el documento que marca el antes y el después de la regulación de la IA en el ámbito de los derechos humanos. El Handbook on Human Rights and Artificial Intelligence, elaborado por el Comité Director de Derechos Humanos (CDDH) y publicado en abril de 2026, no es un texto académico: es una herramienta de gobernanza práctica dirigida a funcionarios, responsables políticos y —por extensión directa— a todas las organizaciones que desarrollen o utilicen sistemas de IA que afecten a personas bajo jurisdicción europea. Para los profesionales del compliance y la integridad corporativa, este manual es, en esencia, el mapa de riesgos legales y reputacionales de la IA para los próximos años.

El marco normativo: tres capas que toda organización debe dominar

El Manual articula la gobernanza de la IA sobre tres instrumentos vinculantes que se superponen y refuerzan mutuamente. El primero y más conocido es el Convenio Europeo de Derechos Humanos (CEDH), que sigue siendo plenamente aplicable a cualquier uso de IA por parte de autoridades públicas o privadas que actúen en su nombre. El segundo es la Carta Social Europea (CSE), el gran olvidado del debate sobre la IA, que garantiza derechos económicos y sociales fundamentales que resultan directamente afectados por los sistemas de decisión automatizada en áreas como el empleo, los servicios sociales o la sanidad. El tercero, y más novedoso, es el Convenio Marco sobre IA y Derechos Humanos, Democracia y Estado de Derecho, el primer tratado internacional vinculante sobre IA, que se convierte así en la columna vertebral de todo el sistema.

Los 7 principios del Convenio Marco — referencia de compliance

1. Dignidad humana y autonomía individual
2. Transparencia y supervisión
3. Responsabilidad y rendición de cuentas
4. Igualdad y no discriminación
5. Respeto a la privacidad y protección de datos
6. Fiabilidad
7. Innovación segura

Toda evaluación de riesgo IA debe verificar el cumplimiento de estos siete ejes a lo largo del ciclo de vida completo del sistema. El Convenio Marco introduce una noción que los equipos de compliance deben integrar de inmediato en sus metodologías: la obligación de realizar evaluaciones de riesgo e impacto sobre derechos humanos, democracia y Estado de Derecho en todo el ciclo de vida del sistema de IA. No solo en el momento del despliegue, sino desde la planificación y el diseño hasta la retirada del sistema. Esto amplía sustancialmente el perímetro de la diligencia debida tecnológica. "Cualquier Estado que reclame un papel pionero en el desarrollo de nuevas tecnologías asume una responsabilidad especial en establecer el equilibrio adecuado." — TEDH, S. y Marper c. Reino Unido, citado en el Manual

El "efecto caja negra": el riesgo más subestimado en compliance de IA

Uno de los desarrollos conceptuales más relevantes del Manual para los profesionales de compliance es la distinción operativa entre transparencia, explicabilidad e interpretabilidad. El documento dedica una sección técnica rigurosa a clarificar por qué estos tres conceptos no son intercambiables y por qué su confusión genera brechas de cumplimiento. La transparencia se refiere a la apertura en la gobernanza del ciclo de vida del sistema. La explicabilidad —definida como la capacidad de ofrecer explicaciones suficientemente

comprensibles sobre por qué un sistema produce un resultado determinado— es el estándar exigible cuando ese resultado afecta a derechos de las personas. La interpretabilidad va más allá: implica hacer accesibles los mecanismos internos de decisión tanto a expertos como a no expertos. El Manual alerta específicamente sobre el efecto "caja negra" de los sistemas de aprendizaje automático: la opacidad no es solo un problema técnico, sino un riesgo jurídico directo. Oculta sesgos, inexactitudes y alucinaciones, y compromete el derecho de las personas afectadas a comprender y contestar las decisiones que les atañen. Para los responsables de compliance, esto se traduce en una exigencia clara: cualquier sistema de IA utilizado para decisiones que afecten a derechos debe poder ofrecer explicaciones inteligibles, o expone a la organización a responsabilidad bajo el Artículo 6 del CEDH (derecho a un proceso justo) y el Artículo 13 (derecho a un recurso efectivo).

Análisis sectorial: donde el riesgo de cumplimiento es más alto

Administración de justicia

El Manual identifica la predicción de reincidencia, la evaluación de riesgos para la libertad provisional y los sistemas de puntuación judicial como los usos de IA de mayor exposición bajo el Artículo 6 del CEDH. La exigencia de "verificabilidad y objetividad" en las decisiones sobre detención y privación de libertad choca frontalmente con el efecto caja negra de los modelos predictivos. Las organizaciones que proveen tecnología al sector justicia deben incluir en sus contratos garantías explícitas de explicabilidad y mecanismos de auditoría accesibles a las partes procesales.

Fuerzas y cuerpos de seguridad

La sección dedicada a seguridad pública es, probablemente, la más técnicamente exigente del Manual. El reconocimiento facial en tiempo real se sitúa en el nivel de mayor riesgo: requiere salvaguardas reforzadas contra el abuso y los efectos de disuasión sobre la libertad de expresión y reunión. El Manual documenta casos reales de sesgo algorítmico que han llevado a la encarcelación de personas inocentes, y establece que los sistemas biométricos entrenados con datos sesgados pueden constituir discriminación estructural bajo el Artículo 14 del CEDH. Para los proveedores tecnológicos que trabajan con administraciones públicas de seguridad, esto supone una responsabilidad de diligencia debida que va más allá del cumplimiento técnico del Reglamento de IA de la UE.

Servicios sociales y bienestar

El caso holandés del sistema SyRi —un algoritmo de detección de fraude en prestaciones sociales declarado contrario al Artículo 8 del CEDH por el Tribunal de La Haya— se convierte en referencia central del Manual. La lección de compliance es inequívoca: la agregación de datos sensibles para construir perfiles de riesgo individuales, sin base legal suficientemente precisa y sin mecanismos de supervisión efectivos, vulnera el derecho a la vida privada. El escándalo holandés de las ayudas por cuidado infantil, en el que funcionarios aplicaban sesgos raciales al revisar manualmente los casos con mayor puntuación de riesgo, demuestra además que la supervisión humana no neutraliza por sí sola el riesgo de discriminación si los operadores no reciben información sobre los motivos del sistema.

Salud

El Manual documenta cómo los modelos de IA médica entrenados predominantemente con datos de determinadas poblaciones infradiagnostican condiciones en grupos subrepresentados —mujeres, minorías étnicas, personas mayores, personas con discapacidad—. Cita estudios que identifican tasas de error diferenciales en el diagnóstico de condiciones dermatológicas según el tono de piel, y casos de sesgos en sistemas de priorización de trasplantes. La exigencia de consentimiento informado libre bajo el Artículo 5 del Convenio de Oviedo implica que los pacientes deben recibir información suficientemente detallada sobre el alcance del uso de herramientas de IA en su tratamiento, lo que plantea desafíos prácticos cuando el sistema no puede explicar su propio funcionamiento.

Empleo y procesos democráticos

Dos sectores de riesgo emergente que el Manual aborda con especial atención. En el ámbito laboral, los sistemas de monitorización de empleados, evaluación del desempeño y toma de decisiones sobre contratación o despido presentan exposición simultánea bajo múltiples artículos del CEDH y la CSE. En el ámbito de los procesos democráticos, la generación automatizada de desinformación, la microfocalización electoral y el uso de IA en la gestión de censos electorales configuran un panorama de riesgo sistémico que el Manual aborda con referencia explícita a la jurisprudencia de la Comisión de Venecia.

Responsabilidad empresarial: las obligaciones positivas del Estado alcanzan al sector privado

Una de las aportaciones más relevantes del Manual para el compliance corporativo es su desarrollo de las obligaciones positivas del Estado respecto a actores privados. El Manual es explícito: los Estados no solo deben abstenerse de violar derechos mediante IA; tienen la obligación activa de regular y supervisar las actividades de las empresas que desarrollan o despliegan sistemas de IA que puedan afectar a derechos humanos. Esto incluye la obligación de garantizar recursos efectivos frente a las actuaciones de actores privados.

El Manual referencia los Principios Rectores de Naciones Unidas sobre Empresas y Derechos Humanos como marco de referencia para la responsabilidad corporativa. La diligencia debida en materia de derechos humanos —el proceso mediante el cual las empresas identifican, previenen, mitigan y dan cuenta de los impactos sobre derechos humanos derivados de sus actividades de IA— deja de ser una práctica voluntaria de ESG para convertirse en el estándar de cumplimiento esperado por los órganos supervisores europeos.

Lista de verificación para responsables de compliance

1. ¿Existe una evaluación de impacto en derechos humanos para cada sistema de IA en producción?
2. ¿Pueden los afectados recibir una explicación comprensible de las decisiones automatizadas?
3. ¿Los datos de entrenamiento han sido auditados para detectar sesgos sistemáticos?
4. ¿Existen mecanismos de recurso efectivos accesibles para las personas afectadas?
5. ¿El sistema mantiene logs de auditoría inalterables accesibles a supervisores independientes?

6. ¿Se ha evaluado el riesgo de efecto de disuasión ("chilling effect") sobre libertades fundamentales?

La herramienta HUDERIA: el puente entre el Convenio Marco y la práctica

Para hacer operativas las obligaciones del Convenio Marco, el Consejo de Europa ha desarrollado la herramienta HUDERIA (Risk and Impact Assessment of AI Systems from the point of view of Human Rights, Democracy and the Rule of Law). Aunque no es jurídicamente vinculante, constituye la guía metodológica de referencia para implementar las evaluaciones de riesgo e impacto exigidas por el Convenio. El Manual recomienda explícitamente su uso como complemento de los marcos técnicos de gestión de riesgos existentes. Para los equipos de compliance, HUDERIA es el equivalente funcional de lo que representó la ISO 19600 para los sistemas de gestión de compliance hace una década: el estándar metodológico de facto que los reguladores utilizarán como referencia en sus supervisiones, aunque no sea obligatorio por ley.

Conclusión: de la declaración de principios a la gestión de riesgo operativa

El Manual del CDDH no es un texto filosófico sobre los valores de la tecnología. Es un documento de trabajo que mapea, sector por sector, los artículos del CEDH y la CSE que se activan ante determinados usos de IA, cita la jurisprudencia relevante del Tribunal Europeo de Derechos Humanos, e identifica los puntos de fallo más documentados en implementaciones reales. Su nivel de detalle —101 páginas con referencias cruzadas a la legislación vigente y a precedentes judiciales concretos— lo convierte en una referencia técnica de primer orden para cualquier organización que tome en serio su gestión de riesgos en el ámbito de la inteligencia artificial.

El mensaje central es claro: el marco regulatorio de la IA en Europa no se limita al Reglamento de IA de la UE. Se superpone sobre un sistema de derechos humanos vinculante, interpretado dinámicamente por tribunales activos, que alcanza a actores privados a través de las obligaciones positivas de los Estados. Las organizaciones que gestionen su compliance de IA mirando solo a Bruselas tienen un punto ciego estratégico de proporciones considerables.

La pregunta relevante para los consejos de administración y los comités de compliance no es si su organización usará sistemas de IA. Ya los usa. La pregunta es si tiene un marco de gobernanza capaz de responder ante un juez europeo de derechos humanos.