GOBERNAR LA INTELIGENCIA ARTIFICIAL

Un modelo de gobierno corporativo para la era de la inteligencia artificial

Governance Actually somos un Think Tank y un observatorio sobre temas de Gobierno Corporativo, Riesgos, Control Interno y Auditoría.

Position Paper

Marzo 2026

ÍNDICE

1. Resumen ejecutivo 3
2. La inteligencia artificial como realidad operativa 3
3. Los riesgos de la IA: argumentos para su gobierno 4
3.1. Riesgos operativos y de información 4
3.2. Riesgos de ciberseguridad 5
3.3. Riesgo de descontrol: la dependencia sin supervisión 5
3.4. Riesgos globales y sistémicos 5
3.5. La IA amplifica lo que encuentra 5
4. Un modelo de gobierno corporativo de la IA 6
4.1. Principios y política de IA 6
4.2. Estrategia e integración en el gobierno corporativo 6
4.3. Roles y responsabilidades 6
4.4. Gestión de riesgos de IA 7
4.5. Controles y procesos 7
4.6. Transparencia y rendición de cuentas 8
4.7. Ética y valores: lo que no se puede programar 8
4.8. Formación y cultura 8
5. Integración en la gobernanza corporativa existente 9
5.1. Alineación con marcos de referencia 9
5.2. Integración, no duplicación 9
5.3. El papel de la auditoría interna 9
5.4. Medición y mejora continua 9
6. Esquema integrado del modelo de gobierno 10
7. Conclusiones 11
Referencias 12

1. Resumen ejecutivo

La inteligencia artificial ha dejado de ser una promesa tecnológica para convertirse en una realidad que transforma organizaciones, mercados y sociedades. Su capacidad para procesar información, aprender de datos y tomar decisiones de forma autónoma plantea oportunidades sin precedentes, pero también riesgos que exigen respuestas estructuradas. **El riesgo principal no es la tecnología en sí misma, sino la ausencia de control sobre ella.**

Este documento propone una visión del gobierno de la IA fundamentada en **tres convicciones**. Primera, que la regulación excesivamente detallada y prescriptiva es contraproducente: llega tarde, frena la innovación y genera deslocalización de capacidades hacia jurisdicciones menos exigentes. Segunda, que el gobierno corporativo —entendido como sistema de dirección, control y rendición de cuentas— constituye el mecanismo más eficaz para garantizar un uso responsable de la IA en las organizaciones. Tercera, que el papel de la regulación debe limitarse a establecer un marco habilitante que obligue a las organizaciones a disponer de un modelo de gobierno de la IA transparente, verificable y sujeto a supervisión; sin imponer soluciones técnicas concretas.

El documento propone un modelo operativo de gobierno de la IA para empresas que integra estrategia, gestión de riesgos, controles internos, ética, formación y rendición de cuentas; todo ello alineado con los marcos de gobierno corporativo existentes. No se trata de crear estructuras paralelas, sino de **incorporar la gobernanza de la IA al sistema de gobierno de la organización**.

2. La inteligencia artificial como realidad operativa

La velocidad de cambio tecnológico se ha acelerado de forma exponencial. Lo que antes tardaba generaciones en transformarse, hoy se modifica en años o meses. La inteligencia artificial representa el último y más profundo de estos saltos: sistemas capaces de gestionar, interpretar y generar información con capacidades que ya superan, en muchos ámbitos, las del ser humano individual y colectivo.

Conviene distinguir entre los distintos estadios de desarrollo de la IA. La IA estrecha o específica, que es la que opera generalizadamente hoy, se especializa en tareas concretas: generación de texto e imágenes, análisis de datos, reconocimiento de patrones, optimización de procesos. La IA general aspiraría a un razonamiento autónomo comparable al humano. Y la súper inteligencia artificial, de carácter teórico, superaría la inteligencia humana en todas sus dimensiones.

Varios elementos convergen para hacer posible la IA actual: redes de comunicación veloces, capacidades masivas de almacenamiento, enormes volúmenes de datos digitalizados, capacidad de cómputo creciente —potenciada por la computación cuántica emergente—, algoritmos evolucionados y lenguajes de programación cada vez más accesibles.

Lo relevante para el propósito de este documento no es el detalle tecnológico, sino sus consecuencias: **la IA ya se integra en procesos críticos de las organizaciones** (toma de decisiones, generación de contenidos, interacción con clientes, optimización operativa) y esta integración se acelera. No estamos ante una tecnología emergente que requiera cautela expectante, sino ante una **realidad operativa que exige gobierno**.

3. Los riesgos de la IA: argumentos para su gobierno

El problema de la inteligencia artificial ya no es la tecnología. El verdadero desafío es humano, organizativo y social. El éxito de la IA no se medirá por sus capacidades técnicas, sino por la capacidad colectiva de mantener autonomía, criterio y control sobre sistemas cada vez más potentes. Cada uno de los riesgos que se describen a continuación constituye, por sí mismo, un argumento para que las organizaciones dispongan de un modelo de gobierno de la IA sólido.

3.1. Riesgos operativos y de información

Los sistemas de IA operan con datos y generan resultados que alimentan decisiones. Si los datos son inexactos, sesgados o incompletos, los resultados lo serán también, con la particularidad de que la IA escala esos errores a una velocidad y un volumen que la revisión humana manual no puede absorber. La generación de contenido falso o engañoso a gran escala, los sesgos derivados de datos de entrenamiento que producen resultados discriminatorios, la publicación no intencional de datos confidenciales o el uso de datos inexactos en decisiones estratégicas son riesgos concretos y crecientes.

Para una organización, esto se traduce en riesgo reputacional, riesgo legal, riesgo financiero y riesgo operativo. Sin un modelo de gobierno que establezca controles sobre la calidad de los datos, la validación de resultados y la trazabilidad de las decisiones automatizadas, estos riesgos quedan sin gestión.

3.2. Riesgos de ciberseguridad

La IA potencia las capacidades de defensa en ciberseguridad, pero también potencia las capacidades de ataque. Los sistemas de IA pueden ser manipulados mediante datos de entrenamiento envenenados, evasión de cortafuegos o alteración de algoritmos. La superficie de ataque se amplía: los sistemas de IA son, en sí mismos, activos críticos que deben protegerse. La seguridad en IA va más allá de la ciberseguridad tradicional e incluye la integridad del propio sistema de aprendizaje.

3.3. Riesgo de descontrol: la dependencia sin supervisión

El riesgo principal, el que subyace a todos los demás, es **no controlar la inteligencia artificial**. Transferir la gestión de información y la toma de decisiones a sistemas más capaces que los humanos, sin mecanismos de supervisión y gobierno adecuados, es el escenario que debe evitarse. La dependencia excesiva de la IA sin intervención humana, especialmente en decisiones con impacto significativo sobre personas, genera un riesgo sistémico que ningún control técnico aislado puede mitigar.

3.4. Riesgos globales y sistémicos

La dimensión global de la IA añade una capa adicional de complejidad. La IA opera sin fronteras claras: los modelos se entrenan con datos globales, los servicios se prestan desde cualquier jurisdicción y los efectos de una decisión automatizada pueden impactar simultáneamente en múltiples mercados y comunidades. Los riesgos de concentración tecnológica en un número reducido de actores, la asimetría de información entre quienes desarrollan la IA y quienes la utilizan, y la velocidad con que la tecnología supera la capacidad de respuesta de reguladores e instituciones son factores que refuerzan la necesidad de que cada organización asuma la responsabilidad de gobernar su propia relación con la IA.

3.5. La IA amplifica lo que encuentra

Una característica crítica de la IA es que amplifica la calidad —o la falta de calidad— de las organizaciones donde se integra. En organizaciones con procesos claros, datos fiables y liderazgo sólido, la IA multiplica capacidades. En organizaciones sin estructuras claras de decisión y control, la IA acelera errores y conflictos. Esta característica amplificadora convierte al gobierno corporativo en la primera línea de defensa: la calidad del gobierno de una organización determina la calidad de los resultados que obtendrá de la IA.

4. Un modelo de gobierno corporativo de la IA

Este capítulo propone un modelo operativo de gobierno de la IA integrado en la gobernanza corporativa existente. No se trata de crear una estructura paralela, sino de incorporar la gobernanza de la IA al sistema de dirección, control y rendición de cuentas de la organización, en línea con los principios del gobierno corporativo generalmente aceptados.

La premisa es que **la IA no requiere un gobierno separado: requiere que el gobierno corporativo se adapte para incorporarla**. Del mismo modo que las organizaciones integraron en su día la gestión de riesgos financieros, la ciberseguridad o el cumplimiento normativo en sus estructuras de gobierno, deben ahora integrar la gobernanza de la IA.

4.1. Principios y política de IA

El punto de partida es la definición de una política de IA aprobada por el máximo órgano de gobierno de la organización. Esta política debe establecer los principios éticos que rigen el desarrollo y uso de la IA en la empresa, los usos permitidos y prohibidos, los criterios para evaluar nuevas aplicaciones de IA, los estándares de transparencia exigidos y el compromiso con la supervisión humana en decisiones críticas.

La política de IA no es un documento decorativo. Es el marco de referencia que orienta todas las decisiones posteriores y debe ser concreta, conocida por toda la organización y revisada periódicamente a medida que evoluciona la tecnología y el contexto.

4.2. Estrategia e integración en el gobierno corporativo

La estrategia de IA debe ser parte integral de la estrategia empresarial, no un apéndice tecnológico. Esto implica definir para qué utiliza la organización la IA, qué objetivos de negocio persigue con ella, qué límites establece y cómo mide el valor generado. La IA no genera valor por sí misma; lo genera cuando se integra en procesos organizativos claros, con datos fiables y bajo liderazgo competente.

El consejo de administración debe supervisar la estrategia de IA como parte de su función de supervisión estratégica, asegurando que los desarrollos de IA estén alineados con el propósito, los valores y los objetivos a largo plazo de la organización. **La IA es hoy un asunto de consejo, no un asunto exclusivamente tecnológico.**

4.3. Roles y responsabilidades

Un modelo de gobierno eficaz requiere asignación clara de roles y responsabilidades en todos los niveles de la organización.

Consejo de administración: Supervisión de la estrategia de IA, aprobación de la política de IA, vigilancia del mapa de riesgos asociado y revisión periódica del modelo de gobierno. El consejo no necesita ser técnico en IA, pero sí debe entender sus implicaciones estratégicas y de riesgo.

Alta dirección: Implementación de la estrategia y política de IA, asignación de recursos, definición de líneas de reporte y supervisión operativa de los desarrollos de IA. Un ejecutivo de máximo nivel debe asumir la responsabilidad última sobre la gobernanza de la IA en la organización.

Comité de IA: Órgano técnico-estratégico con capacidad para evaluar y validar desarrollos de IA, revisar aspectos éticos, supervisar cortafuegos y aprobar despliegues de sistemas de alto riesgo. Puede integrarse como función dentro de comités existentes (riesgos, tecnología) o constituirse como órgano específico según la dimensión y complejidad de la organización.

Responsables operativos: Cada área que desarrolle o utilice sistemas de IA debe tener un responsable identificado que supervise el cumplimiento de la política, la calidad de los datos, la documentación técnica y la gestión de incidencias.

4.4. Gestión de riesgos de IA

La gestión de **riesgos de IA debe integrarse en el sistema de gestión de riesgos corporativo,** no funcionar como un ejercicio aislado. Esto requiere identificar los riesgos específicos que la IA genera o amplifica, evaluarlos en términos de probabilidad e impacto, definir controles y planes de mitigación, asignar responsables y establecer indicadores de seguimiento.

Los riesgos específicos a considerar incluyen, entre otros: sesgo y discriminación en resultados, inexactitud de datos y de las conclusiones generadas, falta de explicabilidad de las decisiones automatizadas, vulnerabilidades de ciberseguridad, dependencia excesiva de sistemas de IA sin supervisión humana, publicación o filtración no intencional de datos confidenciales, riesgos reputacionales derivados de un uso inapropiado y riesgos de cumplimiento normativo.

El sistema de gestión de riesgos debe ser iterativo y continuo, con revisión periódica del mapa de riesgos de IA, alineado con el ciclo de revisión estratégica de la organización. El apetito de riesgo en materia de IA debe definirse de forma explícita y ser aprobado por el consejo.

4.5. Controles y procesos

El modelo de gobierno requiere controles específicos adaptados a la naturaleza de los sistemas de IA. Estos controles deben abordar, al menos, las siguientes dimensiones.

Calidad de datos: Los datos son el combustible de la IA. Controles sobre la procedencia, integridad, actualidad y representatividad de los datos de entrenamiento y operación son fundamentales. Datos sesgados o incompletos producen resultados sesgados o erróneos.

Transparencia y trazabilidad: Los sistemas de IA deben ser documentados de forma que sea posible entender qué hacen, con qué datos operan, cómo se han entrenado y qué lógica aplican. La trazabilidad de las decisiones automatizadas es esencial para la rendición de cuentas.

Supervisión humana: Todo sistema de IA que participe en decisiones con impacto significativo debe contar con mecanismos de supervisión humana. El grado de intervención humana debe ser proporcional al nivel de riesgo de la decisión.

Seguridad: Protección de los sistemas de IA frente a manipulación de datos, alteración de algoritmos, evasión de cortafuegos y accesos no autorizados.

Validación y pruebas: Antes del despliegue de cualquier sistema de IA de riesgo relevante, debe realizarse una evaluación rigurosa que incluya pruebas de funcionamiento, análisis de sesgos, verificación de robustez y evaluación de impacto.

4.6. Transparencia y rendición de cuentas

La transparencia es la esencia del buen gobierno. En materia de IA, esto se traduce en dos dimensiones complementarias. La transparencia interna implica que la dirección y el consejo dispongan de información fiable, relevante y oportuna sobre el estado de los sistemas de IA, los riesgos asociados y la eficacia de los controles. La transparencia externa implica que los grupos de interés —reguladores, inversores, clientes, sociedad— puedan conocer cómo la organización gobierna su uso de la IA.

La rendición de cuentas es la base de esta transparencia. Debe quedar claro quién decide sobre la IA en la organización, con qué criterios, con qué responsabilidad y con qué consecuencias en caso de desviaciones. Los usuarios de servicios que incorporen IA deben conocer la naturaleza, alcance y riesgos de dichos servicios.

4.7. Ética y valores: lo que no se puede programar

Un sistema de IA puede tomar conciencia de sí mismo, puede tener reglas de programación que simulen comportamiento ético, pero **no puede tener ética humana genuina**. La ética es intrínseca a la conciencia humana. Los seres humanos poseen algo que no es programable ni aprendible mediante algoritmos: la capacidad de remordimiento, ese sentido que reformula conclusiones aparentemente correctas según datos, pero que violarían principios profundos.

Esto no es un argumento contra la IA, sino a favor de **mantener la supervisión humana en el centro del modelo de gobierno.** Las decisiones verdaderamente correctas incluyen factores humanamente imprevisibles. La IA puede simular estas decisiones mediante reglas, pero sin genuina conciencia ética. Por eso, humanizar la IA —es decir, garantizar que los valores y la ética humana permanezcan como referencia última de las decisiones— es quizás la dimensión más importante del gobierno de la IA.

En términos prácticos, esto se traduce en incorporar revisiones éticas en los procesos de desarrollo y despliegue de IA, en incluir perfiles no técnicos en los comités de supervisión y en mantener la decisión humana como última instancia en todo sistema de IA con impacto significativo sobre personas.

4.8. Formación y cultura

Ningún modelo de gobierno funciona sin las personas adecuadas. La formación en IA es un pilar central de la gobernanza y, sin embargo, es un aspecto frecuentemente descuidado. No se trata solo de formación técnica, sino cultural y organizativa: saber interpretar resultados de IA, cuestionarlos, tomar decisiones informadas sin delegar automáticamente el criterio humano a una máquina.

Así como la sociedad se esfuerza en que la educación humana sea adecuada, reglada y supervisada, debe aplicarse un rigor similar al aprendizaje de los sistemas de IA. El aprendizaje de la IA es hoy desordenado y falto de supervisión. Mientras esto importa poco en la IA estrecha actual, con la evolución hacia sistemas más autónomos el control del aprendizaje de la IA se convierte en una cuestión crítica.

La organización debe invertir en **alfabetización digital y en IA** a todos los niveles: desde el consejo de administración hasta los usuarios finales. Sin competencias adecuadas, aumenta el riesgo de errores, dependencia excesiva y desinformación.

5. Integración en la gobernanza corporativa existente

El modelo de gobierno de la IA propuesto no opera en el vacío ni requiere crear estructuras organizativas nuevas. Se integra en los marcos de gobernanza corporativa de referencia y se implementa a través de los órganos y funciones que las organizaciones ya tienen.

5.1. Alineación con marcos de referencia

Un marco de gobierno corporativo organizado proporciona una base natural para integrar la gobernanza de la IA. La supervisión de la IA encaja en el componente de supervisión del consejo; la estrategia de IA se integra en el componente estratégico; la formación y la ética conectan con cultura y personas; la transparencia se alinea con comunicación; y la gestión de riesgos de IA refuerza la resiliencia organizativa.

No se trata de adaptar la organización a la IA, sino de adaptar el gobierno de la organización para que la IA quede gobernada. La diferencia es relevante: lo primero implica crear aparatos nuevos; lo segundo, reforzar los existentes.

5.2. Integración, no duplicación

Las organizaciones ya disponen de consejos de administración, comisiones de auditoría, comités de riesgos, funciones de cumplimiento y auditoría interna. La gobernanza de la IA debe aprovechar estos órganos y funciones, ampliando su mandato y reforzando sus capacidades cuando sea necesario, en lugar de crear un aparato organizativo separado que genere fricción, duplicidad y coste.

Esto significa, en términos concretos, que la comisión de auditoría debe incluir la IA en su perímetro de supervisión, que el comité de riesgos debe incorporar riesgos de IA en su mapa, que la función de cumplimiento debe verificar la adhesión a la política de IA y que la auditoría interna debe evaluar la eficacia de los controles sobre sistemas de IA.

5.3. El papel de la auditoría interna

La auditoría interna juega un papel fundamental en el modelo de gobierno de la IA. Su función no es solo verificar controles técnicos, sino evaluar si el modelo de gobierno funciona en la práctica: si la política de IA se aplica, si los roles y responsabilidades están operativos, si los riesgos se gestionan efectivamente, si la información que llega al consejo es fiable y si la rendición de cuentas es real. Para ello, la auditoría interna necesitará reforzar sus capacidades con perfiles que entiendan la tecnología de IA, sin que esto signifique convertirla en una función técnica: su valor reside en la perspectiva independiente sobre la eficacia del gobierno.

5.4. Medición y mejora continua

Como todo sistema de gobierno, el modelo requiere mecanismos de medición y mejora continua. Esto incluye indicadores clave de riesgo específicos para IA, evaluaciones periódicas de la eficacia de los controles, revisiones del mapa de riesgos alineadas con la evolución tecnológica y evaluaciones externas cuando la complejidad lo justifique. El consejo de administración debe recibir información periódica y estructurada sobre el estado del gobierno de la IA, no como un ejercicio de cumplimiento formal, sino como herramienta para la toma de decisiones estratégicas.

6. Esquema integrado del modelo de gobierno

El siguiente esquema sintetiza los componentes del modelo de gobierno propuesto y su relación con la estructura de gobernanza corporativa.

Componente	Contenido y alcance
Política de IA	Principios éticos, usos permitidos/prohibidos, criterios de evaluación, estándares de transparencia. Aprobada por el consejo.
Estrategia	Objetivos de negocio con IA, integración en plan estratégico, medición de valor, límites definidos.
Roles	Consejo (supervisión), alta dirección (implementación), comité de IA (validación técnica), responsables operativos (áreas).
Gestión de riesgos	Mapa de riesgos de IA integrado en ERM, apetito de riesgo definido, controles, indicadores, revisión periódica.
Controles	Calidad de datos, trazabilidad, supervisión humana, seguridad, validación previa al despliegue.
Transparencia	Información interna al consejo, información externa a grupos de interés, rendición de cuentas clara.
Ética	Revisiones éticas, perfiles no técnicos en supervisión, decisión humana como última instancia.
Formación	Alfabetización en IA a todos los niveles, formación continua, cultura de cuestionamiento crítico.
Auditoría	Sistema de calidad del modelo de gobierno, auditoría interna de sistemas de IA, evaluación periódica.

Este modelo proporciona un marco de referencia adaptable a la dimensión, complejidad, sector y nivel de madurez de cada organización en el uso de la IA. Lo esencial es que exista un modelo documentado, conocido, supervisado y sujeto a rendición de cuentas.

7. Conclusiones

La inteligencia artificial es bienvenida por sus capacidades de procesamiento, aprendizaje y ejecución autónoma. Abre horizontes tecnológicos sin precedentes. Pero su potencial solo se materializará de forma positiva si se gobierna adecuadamente.

El riesgo principal de la IA se reduce a una sola palabra: **descontrol**. Y la respuesta a ese riesgo no está en la regulación prescriptiva —que llega tarde, burocratiza y genera asimetrías competitivas— sino en el gobierno corporativo responsable, que actúa desde dentro de las organizaciones con agilidad, conocimiento y rendición de cuentas. La regulación debe limitarse a exigir que las organizaciones dispongan de modelos de gobierno de la IA transparentes y verificables.

Este *paper* propone un modelo fundamentado en las tres convicciones que lo vertebran:

Primera. La regulación excesivamente detallada y prescriptiva es contraproducente. Llega tarde, burocratiza y genera asimetrías competitivas. Los riesgos de la IA —operativos, de ciberseguridad, de descontrol, globales— son reales y crecientes, pero la respuesta eficaz no está en legislar cada supuesto técnico, sino en exigir que quienes desarrollan o utilizan la IA la gobiernen.

Segunda. El gobierno corporativo —entendido como sistema de dirección, control y rendición de cuentas— constituye el mecanismo más eficaz para garantizar un uso responsable de la IA. La gobernanza de la IA no es un tema aislado ni un asunto exclusivamente tecnológico: está presente en el día a día de las compañías y debe integrarse en las estructuras de gobierno existentes, con la misma seriedad que cualquier otro ámbito crítico. El modelo propuesto en este documento —política de IA, estrategia, roles, gestión de riesgos, controles, transparencia, ética y formación— ofrece esa estructura. Y en su centro, lo más importante: humanizar la IA, porque los valores y la ética humana no son programables ni aprendibles por algoritmos.

Tercera. El papel de la regulación debe limitarse a establecer un marco habilitante que obligue a las organizaciones a disponer de un modelo de gobierno de la IA transparente, verificable y sujeto a supervisión, sin imponer soluciones técnicas concretas. La transparencia y la rendición de cuentas son la piedra angular: sin ellas, cualquier estructura de gobierno es una formalidad vacía.

El futuro de la inteligencia artificial no dependerá de lo que sea capaz de hacer, sino de cómo decidamos gobernarla.

Referencias

COSO – NACD Corporate Governance Framework, Public Exposure Draft, mayo 2025. Committee of Sponsoring Organizations of the Treadway Commission / National Association of Corporate Directors.

Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial).

OECD AI Principles. Organización para la Cooperación y el Desarrollo Económicos.

Global Partnership on Artificial Intelligence (GPAI). Iniciativa multilateral de gobiernos, industria y sociedad civil.

World Economic Forum – Global Risks Report 2025. Foro Económico Mundial.

Del Río, R. – Artículos de opinión publicados en Escudo Digital (2025-2026): *Gobernando la Inteligencia Artificial; El cóctel de la IA que contiene computación cuántica; La nueva estrategia de EEUU en IA; El problema de la IA ya no es la tecnología; OSINT: cuando los datos públicos se convierten en inteligencia.*

Del Río, R. – *La Inteligencia Artificial y a dónde nos lleva: riesgos y control*, diciembre 2024. Documento de trabajo.