

Chapter 9 / Capítulo 9

AI, metaverse and sustainability in the new
organizational design
ISBN: 978-9915-9928-8-4
DOI: 10.62486/978-9915-9928-8-4.Ch09
Pages: 54-59



© 2026 The authors. This is an open access
book chapter distributed under the terms of
the Creative Commons Attribution (CC BY)
4.0 License.

Transformation governance: who decides when the technology decides

Gobernanza de la transformación: quién decide cuando decide la tecnología

Jesús Alejandro Vázquez Meza¹ 

Víctor Manuel Martínez García² 

Yennifer Díaz Romero³ 

Manuel Iván Tostado Ramírez⁴ 

Rogelio Estrada Lizárraga⁴ 

José de Jesús Reyes Machain² 

1 Universidad Autónoma de Sinaloa, Unidad Académica Preparatoria Rubén Jaramillo, Vicerrectoría, Sinaloa, México

2 Universidad Autónoma de Sinaloa, Facultad de Arquitectura y Diseño, Vicerrectoría, Sinaloa, México

3 Universidad Autónoma de Sinaloa, Facultad de Ciencias de la Educación, Vicerrectoría, Sinaloa, México

4 Universidad Autónoma de Sinaloa, Facultad de Informática, Vicerrectoría, Sinaloa, México

5 Universidad Autónoma de Sinaloa, Facultad de Informática, Vicerrectoría, Sinaloa, México

ABSTRACT

This chapter examines the governance of artificial intelligence when systems systematically influence organizational decisions. It begins by asking who determines their purposes, controls their operation, and remains responsible for their consequences. The discussion reviews responsibility for automated decisions and regulatory frameworks addressed in the manuscript, together with auditing and verifiable documentation. Data protection is treated as a matter of authority, purpose, and consent as well as technical security. Ethical committees and worker participation are considered as ways of introducing diverse perspectives into oversight. The chapter also discusses trust and the legitimacy of algorithmically mediated decisions. It emphasizes that adopting a tool and governing its consequences are distinct but inseparable processes. Governance requires arrangements for assessing risks, explaining decisions, correcting problems, and preserving identifiable human responsibility. Institutional oversight and deliberation are presented as conditions for sustaining technological transformation with legitimacy over time. Rather than assuming that one framework solves every problem, the chapter brings together the requirements that should guide organizational implementation.

Keywords: Algorithmic governance; Automated decisions; AI auditing; Regulation; Worker participation

RESUMEN

El capítulo examina la gobernanza de la inteligencia artificial cuando los sistemas participan de manera sistemática en decisiones organizacionales. Parte de la necesidad de identificar quién define sus fines, controla su funcionamiento y responde por los efectos que generan. Revisa las responsabilidades vinculadas con decisiones automatizadas y los marcos regulatorios discutidos en la obra, así como la función de la auditoría y de la documentación verificable. La protección de datos se aborda como una cuestión de autoridad, propósito y consentimiento, además de seguridad técnica. El texto considera los comités éticos y la participación de los trabajadores como vías para incorporar perspectivas diversas a la supervisión. También analiza la confianza y la legitimidad de las decisiones mediadas por algoritmos. En conjunto, la discusión destaca que adoptar una herramienta y gobernar sus consecuencias son procesos diferentes pero inseparables. La gobernanza requiere procedimientos para evaluar riesgos, explicar decisiones, corregir problemas y mantener responsabilidades humanas identificables. El capítulo presenta la supervisión y la deliberación institucional como condiciones para que la transformación tecnológica pueda sostenerse en el tiempo con legitimidad. No da por resuelta la eficacia de un marco único, sino que organiza las exigencias que deberían orientar su implementación. El recorrido muestra conexiones entre procedimientos institucionales y derechos de las personas afectadas. Insiste en que la revisión de decisiones, la documentación y la posibilidad de participación forman parte de una misma arquitectura de responsabilidad organizacional.

Palabras clave: Gobernanza algorítmica; Decisiones automatizadas; Auditoría de IA; Regulación; Participación laboral

9.1 GOBERNANZA ALGORÍTMICA

Cuando un algoritmo empieza a influir de forma sistemática en decisiones que antes tomaban personas, como a quién contratar, qué precio ofrecer o qué contenido mostrar, surge una pregunta que ninguna organización puede seguir postergando, la de quién gobierna ese algoritmo. El planteamiento de esta pregunta se hace desde un sentido político, sin pensar en el aspecto técnico de quien lo mantiene funcionando y este se refiere mas que nada a quién decide lo que se codifica, sus límites y ante quién rinde cuentas.

La UNESCO (2021) sitúa esta pregunta en el centro de su *Recomendación sobre la Ética de la Inteligencia Artificial*, en donde la supervisión humana significativa, lejos de ser un accesorio opcional, se entiende como una condición necesaria; a partir de ello se puede pensar que, sin dicha supervisión, ningún sistema de IA podría considerarse legítimo dentro de una organización o una sociedad. Entendida de esta manera, la gobernanza algorítmica es el conjunto de estructuras, procesos y personas responsables de que esa supervisión se ejerza en la práctica, más allá de su declaración formal, lo que implica que exista alguien con autoridad efectiva para cuestionar, ajustar o incluso desactivar un sistema cuando sus resultados dejan de ser aceptables.

Si bien la brecha entre esta aspiración y la práctica real se está cerrando con rapidez, todavía es considerable. Al respecto, McKinsey & Company (2025) encontró que solo el 28 % de las organizaciones tiene un director ejecutivo que asume responsabilidad directa sobre la gobernanza de la IA, y que apenas un 17 % reporta que esa responsabilidad recae en otras áreas. En otras palabras, la existencia de un comité no garantiza que haya una autoridad real y visible detrás de él, distinción que separa a la gobernanza algorítmica auténtica de la simple formalidad organizacional.

9.2 RESPONSABILIDAD ANTE DECISIONES AUTOMATIZADAS

La pregunta de quién responde cuando una decisión automatizada causa daño, discutida ya en el capítulo 7 desde la perspectiva ética, adquiere aquí una dimensión más estructural, esto porque, además de determinar la responsabilidad después de que algo salió mal, es necesario que desde el inicio de las operaciones existan mecanismos que permitan asignar la responsabilidad de forma clara.

Loayza Villanueva (2025) analiza este vacío desde la perspectiva del derecho laboral y documenta cómo la incorporación acelerada de la inteligencia artificial a las relaciones de trabajo genera una necesidad urgente de reconfigurar los marcos legales existentes, mismos que en su mayoría fueron diseñados pensando en decisiones tomadas exclusivamente por personas. De acuerdo con su análisis, sin una regulación equitativa y responsable que distribuya con claridad esta responsabilidad, las organizaciones quedan operando en una zona gris en donde la ausencia de reglas explícitas termina beneficiando, casi siempre, a quien tiene más poder dentro de la relación, que por lo general es el empleador que decide implementar el sistema y no la persona afectada por sus decisiones.

Por su parte, Castillo Martínez et al. (2024), analizados con más detalle en el Capítulo 3, documentan una dimensión adicional de este vacío que rara vez se discute junto con la responsabilidad legal y que tiene que ver con la responsabilidad sobre la autenticidad y la propiedad intelectual del contenido que un sistema de IA genera dentro de una organización. Cuando un producto de trabajo combina contribución humana y generación algorítmica en proporciones difíciles de deslindar, el atribuir responsabilidad, tanto sobre errores como sobre créditos y reconocimiento, exige criterios que la mayoría de las organizaciones todavía no ha definido de forma explícita. Se trata de la misma zona gris que documenta Loayza Villanueva (2025), vista ahora desde el ángulo de la autoría, que se suma al del daño, lo que muestra que la responsabilidad ante decisiones automatizadas va más allá de quién responde cuando algo sale mal e incluye, también, quién responde, y quién recibe crédito, cuando todo camina de la mejor forma en los resultados de la implementación de la misma.

9.3 REGULACIÓN

En los últimos años, el panorama regulatorio sobre inteligencia artificial dejó de ser un terreno mayoritariamente voluntario para convertirse en un campo con obligaciones legales concretas y exigibles. Tal es el caso de la Unión Europea (2024), que a través del Reglamento (UE) 2024/1689, conocido como la Ley de Inteligencia Artificial, implementó el primer marco legal

integral del mundo sobre esta tecnología, mismo que se basa en lo niveles de riesgos, en donde, cuanto mayor es el riesgo potencial de un sistema de IA para los derechos y la seguridad de las personas, mayores son las obligaciones que debe cumplir quien lo desarrolla o lo utiliza.

Cabe destacar que la atención al riesgo no es exclusiva de la regulación europea. La OCDE (2019), en su *Recomendación del Consejo sobre Inteligencia Artificial*, adoptada por más de 40 países y actualizada en 2024 para incorporar los desafíos específicos de la IA generativa, estableció principios que hoy sirven de referencia normativa para buena parte de las regulaciones nacionales del mundo, un trabajo en paralelo con la Ley de IA europea. Sus cinco principios centrales, los cuales son el desarrollo sostenible, los valores democráticos, con la equidad y la privacidad y la seguridad en el sentido de la protección, funcionan como un lenguaje común que permite a las organizaciones que operan en múltiples países entender qué se espera de ellas, incluso cuando las leyes locales específicas todavía están en desarrollo.

Por su parte, Loayza Villanueva (2025) documenta cómo este panorama regulatorio internacional se traduce, o no, en marcos nacionales concretos dentro del ámbito laboral, y en su análisis encuentra que, si bien la Ley de IA europea ofrece principios generales sólidos, algunos países todavía carecen de una legislación laboral específica que traduzca esos principios en obligaciones concretas para el empleador que decide usar IA en procesos de contratación, evaluación o supervisión de su personal. Esta brecha entre la regulación general y la regulación laboral específica hace que muchas organizaciones sigan operando bajo un criterio de autorregulación que depende enteramente de su propia voluntad de actuar con responsabilidad, situación que, como advierte la propia autora, tiende a beneficiar de forma desproporcionada a quien ya tiene más poder dentro de la relación laboral.

9.4 AUDITORÍA

Ninguna regulación, por bien diseñada que esté, tiene efecto real sin mecanismos que verifiquen su cumplimiento efectivo. La auditoría de sistemas de inteligencia artificial, la cual consiste en revisar con metodología y de forma periódica si un sistema opera dentro de los límites éticos y legales que se supone debe respetar, se ha convertido en una función organizacional cada vez más indispensable y cada vez menos opcional.

Para los sistemas de IA clasificados como de alto riesgo, la Unión Europea (2024) exige documentación técnica detallada y mecanismos de apelación disponibles para las personas afectadas por sus decisiones, lo cual presupone la existencia de procesos de auditoría capaces de generar y verificar esa documentación. Una organización que despliega un sistema de este tipo sin haber establecido antes cómo lo auditará incumple una formalidad legal y, sobre todo, opera sin ninguna manera confiable de saber si el sistema hace lo que se supone que debe hacer.

El *AI Index Report* de Stanford (2025) muestra que la infraestructura de auditoría todavía va muy por detrás de la velocidad de despliegue de estos sistemas, esto al documentar que evaluar sistemas de IA con criterios de responsabilidad sigue siendo poco común, aunque empiezan a surgir puntos de referencia estandarizados diseñados para cerrar esta brecha, como los que evalúan la factualidad y la seguridad de forma sistemática. Mientras tanto, prácticamente todos los desarrolladores de modelos de frontera reportan resultados detallados sobre pruebas de capacidad técnica, es decir, sobre qué tan bien resuelve problemas un modelo, los principales desarrolladores prueban sus modelos con los mismos puntos de referencia sobre sus riesgos éticos y sociales. Esa asimetría entre cuánto se audita la capacidad y cuánto se audita la responsabilidad es el vacío que las organizaciones tienen que llenar por su cuenta cuando ninguna autoridad externa las obliga todavía a hacerlo.

9.5 PROTECCIÓN DE DATOS

La protección de datos, ya discutida en el capítulo anterior desde la perspectiva de la ciberseguridad, reaparece aquí como una pregunta de gobernanza que, junto con la forma en que se protegen técnicamente los datos, abarca quién decide qué datos se recolectan, con qué propósito y bajo qué mecanismo de consentimiento real de las personas involucradas.

La Unión Europea (2024) refuerza este punto desde el propio texto de la Ley de Inteligencia Artificial, que exige a los proveedores de sistemas de alto riesgo demostrar, con documentación verificable, cómo se recolectaron y gobernaron los datos usados para entrenar y operar sus sistemas. Dicha exigencia conecta de forma directa con los principios de la OCDE (2019) sobre

transparencia y explicabilidad, puesto que una organización no puede afirmar de forma creíble que gobierna bien sus datos si no puede explicar, con suficiente detalle, de dónde provienen y quién tiene acceso a ellos en cada etapa del proceso. Cuando se mira desde la gobernanza, además de hacerlo desde la ciberseguridad técnica, la protección de datos es, en esencia, una cuestión de poder, es decir, de quién decide, quién controla y quién rinde cuentas sobre la información más sensible que una organización posee.

9.6 COMITÉS ÉTICOS

Frente a la complejidad creciente de estas decisiones, cada vez más organizaciones están adoptando una estructura de gobernanza que hasta hace poco parecía reservada casi exclusivamente al ámbito de la investigación médica, como lo es el comité de ética, ahora aplicado a la supervisión de sistemas de inteligencia artificial.

Estos comités, integrados idealmente por personas de distintas áreas de la organización, además de las del departamento técnico que desarrolla o implementa el sistema, cumplen una función que ningún individuo aislado puede replicar bien, la cual consiste en someter las decisiones sobre IA a un proceso deliberativo real antes de que se conviertan en política operativa. La UNESCO (2021) respalda este tipo de mecanismos como parte de una gobernanza de IA responsable.

9.7 PARTICIPACIÓN DE LOS TRABAJADORES

Un vacío recurrente en la gobernanza algorítmica es la ausencia casi total de las personas directamente afectadas por estos sistemas, es decir, de los propios trabajadores, en las decisiones sobre cómo se diseñan, implementan y supervisan. Loayza Villanueva (2025) documenta que esta exclusión, que no es accidental, obedece a marcos legales que todavía no contemplan mecanismos claros de consulta o participación laboral diseñados para el contexto de la inteligencia artificial en el trabajo.

Corregir este vacío es una cuestión de justicia procedimental, aunque no se limita a ella. Las personas que usan un sistema todos los días, o que son evaluadas por él, suelen detectar problemas prácticos, sesgos sutiles o efectos no anticipados mucho antes que cualquier auditoría externa formal. Una organización que excluye de manera sistemática a sus trabajadores de las decisiones sobre gobernanza algorítmica genera desconfianza y, al mismo tiempo, renuncia a una fuente valiosa de información temprana sobre cuándo un sistema empieza a fallar.

El caso documentado por Ruiz Sevilla (2026) en el ámbito universitario ilustra este punto desde un ángulo distinto pero complementario. En su investigación se encontró que los propios estudiantes, quienes en ese contexto equivalen a los “trabajadores” cuyo comportamiento cotidiano determina buena parte de la seguridad institucional, identificaban con precisión los riesgos técnicos de ingeniería social, aunque nadie dentro de la institución había diseñado un mecanismo para canalizar ese conocimiento hacia mejoras concretas en la política de seguridad. Si bien el conocimiento estaba ahí, distribuido entre quienes vivían el problema todos los días, hacía falta un canal formal para que influyera en las decisiones de gobernanza. Se trata del mismo vacío que Loayza Villanueva (2025) documenta en el terreno laboral, en el sentido de que, cuando no hay participación, el conocimiento útil sigue existiendo, solo que no tiene ningún lugar formal adonde llegar.

9.8 CONFIANZA Y LEGITIMIDAD

Todos los elementos discutidos en este capítulo, es decir, la gobernanza algorítmica, la responsabilidad, la regulación, la auditoría, la protección de datos, los comités éticos y la participación laboral, convergen, en última instancia, en una sola pregunta sobre si las personas confían en que la organización está usando estas tecnologías de manera legítima.

Con base en la Recomendación de la UNESCO (2021), se puede plantear que ningún sistema de inteligencia artificial, por técnicamente sofisticado que sea, puede sostenerse a largo plazo si carece de la confianza de quienes lo usan o son afectados por él. Dicha confianza, que no se decreta ni se compra con una campaña de comunicación bien diseñada, se construye de forma acumulativa a través de decisiones consistentes de gobernanza que demuestren, con hechos verificables, más allá de las declaraciones de intención, que la organización prioriza la

responsabilidad sobre la velocidad de implementación.

Los datos más recientes muestran que esta confianza es un recurso frágil y desigualmente distribuido. Tal es el caso del *AI Index Report* de Stanford (2025), en el cual se documentan diferencias considerables en el optimismo público hacia la inteligencia artificial según el país. Existe una amplia mayoría de la población en países como China, Indonesia y Tailandia que percibe que los productos y servicios de IA son de beneficio mas que de perjuicios, ese optimismo no es igual en países como Canadá, Estados Unidos y los Países Bajos, donde apenas cuatro de cada diez personas comparten esa misma percepción. Esta brecha de confianza, que es algo más que un simple dato cultural curioso, significa para cualquier organización que opera en múltiples países que la legitimidad de su uso de la inteligencia artificial no puede asumirse como uniforme y tiene que construirse de manera deliberada y adaptada a cada contexto, dado que esa confianza parte, de entrada, de niveles muy distintos.

En este sentido, la organización que atiende estos elementos tiende a tratar la gobernanza de su transformación tecnológica como la base sobre la que se sostiene, a largo plazo, su legitimidad para seguir operando, en lugar de verla como un obstáculo burocrático que hay que sortear.