centro i
PARA LA SOCIEDAD DEL FUTURO

MANUAL · 2026

GUÍA PRÁCTICA

Regulación ágil para la Gobernanza de la Inteligencia Artificial

Selección e implementación de modelos sistémicos de gobernanza y regulación ágil.

CRÉDITOS

Regulación ágil para la Gobernanza de la Inteligencia Artificial

AUTORÍA

Leonardo Martínez Flores

Director de investigación de Centro-i

DIRECCIÓN

Elena Estavillo Flores

Directora General de Centro-i

EQUIPO DE INVESTIGACIÓN

Luis Adrián Salazar · Jhonatan López · Arzu Martínez

FECHA DE PUBLICACIÓN

Septiembre de 2026

NÚMERO DE DOCUMENTO

R-2026-1

AGRADECIMIENTOS

Como parte del proceso de elaboración de este documento, el Centro-i organizó un taller con especialistas en diferentes temas para revisar planteamientos iniciales, discutir sobre la conveniencia de utilizar diferentes instrumentos para la regulación ágil de la inteligencia artificial y trazar las líneas argumentativas a desarrollar en el documento. Las recomendaciones hechas por los participantes cubren un amplio espectro de temas, como el hecho de que el uso de una regulación responsable se centre en el abanico de riesgos enfrentados; la importancia de darle al documento una orientación operativa, de aplicación práctica; que se ofrezcan ejemplos concretos de cómo adoptar la inteligencia artificial en distintos contextos; considerar las relaciones entre los instrumentos propuestos y el orden jurídico existente; usar métricas adecuadas que no se limiten a la eficiencia, sino que también capturen la inclusión y el impacto en los derechos de las personas afectadas; que los resultados de la aplicación de los instrumentos sean comparables y verificables; y sobre la conveniencia de seguir la tendencia regulatoria internacional que se inclina cada vez más hacia un enfoque de verticales, es decir, hacia la regulación diferenciada por sectores.

Apreciamos estas recomendaciones, que fueron tomadas en cuenta y por lo tanto su inclusión se entrelaza con otros elementos a lo largo del texto. Los participantes en el taller, citados en orden alfabético fueron: Mario Berganza, Claudia Del Pozo, Armando Guio, Adriana Labardini, Daniel Legaspi, Jhonatan López, Hugo Martínez McNaught, Lourdes Martínez, Lina Munar Guevara, Alejandro Pisanty Baruch, Manuel Pliego y Patricia Villa.

¿Para quién es este manual?

Este manual es una guía práctica para la selección e implementación de modelos sistémicos de gobernanza y regulación ágil. Su uso es recomendable para:

01 **Empresas privadas**

Interesadas en actualizar sus políticas de gobernanza, en realizar proyectos de innovación de productos o servicios, en llevar a cabo procedimientos de cumplimiento regulatorio, o en realizar auditorías y procesos de certificación relacionados con inteligencia artificial (IA) y tecnologías digitales en rápida evolución.

02 **Universidades**

Escuelas y centros de investigación que enfrentan la necesidad de contar con protocolos para la gestión de proyectos de IA y tecnologías digitales o para la actualización de planes de estudio, de realizar programas de colaboración con empresas privadas y públicas o de diseñar protocolos de evaluación de alumnos y docentes usando IA.

03 **Sociedad civil**

Y organizaciones no gubernamentales que requieren diseñar y aplicar programas e intervenciones específicas o proponer y monitorear proyectos y políticas para poblaciones marginadas, o para reducir desigualdades y brechas de género haciendo uso de IA y tecnologías digitales.

04 **Inversionistas**

Y fondos de inversión que requieren evaluar riesgos regulatorios, identificar buenas prácticas de gobernanza y fomentar la inversión responsable.

05 **Gobiernos**

Y organismos públicos que requieren crear marcos legales y regulatorios, desarrollar normas técnicas, supervisar sistemas de alto riesgo, elaborar políticas públicas, protocolos y estrategias para el desarrollo y adaptación de sistemas y modelos de IA.

¿Es este manual adecuado para tu proyecto?

Contesta las siguientes preguntas
para saber si este manual es adecuado para tu proyecto

PREGUNTA 01

¿Tu organización desarrolla, adquiere, despliega o regula sistemas de IA?

Incluye uso de IA de terceros para tomar decisiones

SÍ

NO

PREGUNTA 02

¿La organización considera prioritario al menos uno de los puntos siguientes?

- Crear entorno de aprendizaje continuo y modificar temporalmente estructuras, reglas e incentivos.
- Considerar efectos sistémicos, no sólo los efectos directos.

NO

RESULTADO

El manual no es aplicable directamente

Puede ser útil como referencia futura

SÍ

PREGUNTA 03

¿Tienes ya un marco de gobernanza o regulación de IA en funcionamiento?

Política, comité o proceso sistemático

NO

SÍ

RUTA A

Sigue los pasos indicados en el manual para construir una primera versión ligera de tu sistema de gobernanza y/o regulación ágil. Para contar con una versión detallada, tendrás que consultar los aspectos metodológicos y los ejemplos desarrollados en los Anexos de este documento.

RUTA B

Usa el manual para actualizar o ampliar tu marco existente

Contenido

Créditos 02

Agradecimientos 03

¿Para quién es este manual? 04

¿Es este manual adecuado para tu proyecto? 05

Presentación 07

Contexto actual 08

01 ¿Cómo se interpretan los conceptos de gobernanza y regulación de la IA en este manual? 11

1.1 Las definiciones básicas 11

1.2 La visión tradicional no es suficiente, hay que agilizarla 13

1.3 Como lo ágil a secas tampoco es suficiente: la concepción ecosistémica 14

1.4 Pensamiento lineal vs. pensamiento sistémico 15

1.5 Principios elementales de la gobernanza y la regulación de la IA 16

02 Tipología y comparación de modelos de regulación ágil 27

Catálogo de instrumentos de gobernanza y regulación ágil 27

Tabla 2.1 Comparación de modelos de regulación ágil para la IA 28

03 Guía para la selección e implementación de modelos sistémicos 33

3.1 Visión y fundamento: el mapa sistémico 33

3.2 Clasificación de riesgos 35

3.3 Matriz impacto-certidumbre (cuadrantes A–D) 34

3.3.1–3.3.4 Cuadrantes A, B, C y D 36

3.4 Las capas de ajuste transversales 45

3.5 Nota sobre las combinaciones de modelos 46

3.6 Metodología genérica para justificar y poner en marcha modelos 47

3.7 Metodología para diseñar y ejecutar un sandbox regulatorio 52

04 Conclusiones y recomendaciones 59

Anexos · disponibles en la página del Centro-i

Anexo 1. Concepción ecosistémica de la gobernanza y la regulación ágil de la IA

Anexo 2. Evaluación del impacto de los sistemas de IA

Anexo 3. Ejemplos de aplicación de la metodología desarrollada en este manual

Anexo 4. Panorama de los sandboxes en Latinoamérica

Anexo 5. Casos de sandboxes para la innovación de la IA en el mundo

Anexo 6. Fichas de referencia para imprimir

Presentación

La velocidad inédita con la que la inteligencia artificial (IA) se está desarrollando, adoptando y transformando industrias, instituciones públicas, entornos educativos y en general la vida cotidiana de las personas, ha puesto en jaque a las estructuras organizacionales vigentes y a sus procesos de operación y gestión, tanto en los gobiernos como en el sector privado.

En particular, los esquemas de gobernanza y los marcos jurídicos y regulatorios tradicionales que fueron diseñados para responder a velocidades mucho más lentas de cambio tecnológico, han entrado súbitamente en una fase de inesperada obsolescencia.

La inteligencia artificial, con todos los beneficios que promete, no debe considerarse como un fin en sí misma, sino como una herramienta sumamente poderosa para solucionar problemas y aumentar la frontera de posibilidades de empresas, individuos y sociedades. Incluso, se concibe cada vez más como infraestructura. Aunados a los impactos positivos, también comporta riesgos de diferente índole y gravedad, de allí que sea tan importante asegurar su gobernanza, es decir la capacidad de conducirla a los fines deseados, gestionando sus riesgos.

Frente a esta necesidad, nos encontramos con que la inteligencia artificial evoluciona continuamente, incrementa sus capacidades y se introduce en nuevos campos. De allí que la arquitectura de reglas, procesos y mecanismos que sostiene su gobernanza necesita ser flexible y poder adaptarse, incorporando nuevos aprendizajes continuamente. Es decir, requiere agilidad. Así, una pregunta central es: ¿cuál puede ser entonces esa nueva generación de esquemas de gobernanza y regulación que nos permita hacerle frente a la revolución de la IA?

Para dar algunas respuestas a esa pregunta central, este manual ofrece un conjunto de principios, metodologías e instrumentos que pueden ser aplicados por entidades públicas y privadas para concebir, diseñar y aplicar marcos sistémicos de gobernanza cimentados en instrumentos de regulación ágil, más adecuados y mejor preparados que los tradicionales para mantener alineados los procesos de innovación, despliegue y adaptación de la IA hacia los objetivos definidos, gestionando sus riesgos.

Contexto actual

En los últimos años, la gobernanza de la inteligencia artificial ha transitado de declaraciones de principios de carácter voluntario hacia marcos normativos progresivamente vinculantes, aunque de naturaleza y alcance muy heterogéneos entre países y jurisdicciones. Instrumentos fundacionales como los Principios de IA de la OCDE[1], la Recomendación de la UNESCO sobre la Ética de la IA[2] y el Convenio Marco del Consejo de Europa sobre IA, Derechos Humanos, Democracia y Estado de Derecho (CETS 225)[3] sentaron un lenguaje común basado en riesgos, transparencia, rendición de cuentas y protección de derechos fundamentales. Sobre esa base, la Unión Europea consolidó con la Ley de Competencia —AI Act— el primer régimen integral y horizontal de regulación por niveles de riesgo, convirtiéndose en referencia obligada —y con frecuencia en punto de comparación o incluso de transferencia normativa directa— para procesos regulatorios en curso en América Latina, incluyendo México. El CETS 225, en su Art. 13, habilita expresamente entornos de pruebas seguros (sandboxes) para innovación. El Convenio fue adoptado en mayo de 2024, abierto a firma en septiembre de 2024 y firmado por 16 países; si bien México participó en la negociación y redacción del Convenio, a la fecha de la publicación de este manual ningún país latinoamericano lo ha firmado aún.

1 OCDE (2019), Principios de Inteligencia Artificial Confiable. https://oecd.ai/en/ai-principles

2 UNESCO (2021), Recomendación de la Ética de la Inteligencia Artificial. https://unesdoc.unesco.org/ark:/48223/pf0000381137

3 Consejo de Europa (2024), Convenio Marco sobre Inteligencia Artificial. Este convenio fue redactado por los 46 Estados Miembros más los países observadores Canadá, Japón, México, la Santa Sede y Estados Unidos, además de otros países no miembros: Australia, Argentina, Costa Rica, Israel, Perú y Uruguay, si bien no ha sido firmado por todos ellos. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

En la Evaluación del estado de preparación de la inteligencia artificial en México de la UNESCO[4], publicado en el 2024, se hacen varias recomendaciones específicas sobre estos temas. Se sugiere innovar en el diseño de instrumentos regulatorios ágiles que respondan a la dinámica del avance tecnológico y contemplen mecanismos para el aprendizaje y la mejora continua de la regulación, entre los cuales están los sandboxes, lineamientos y guías revisables modularmente a medida que sea necesario, así como definición de espacios seguros para la compartición de experiencias. También se propone incluir espacios de innovación y experimentación regulatoria, esto es de regulación ágil, que enriquezcan los procesos de aprendizaje continuo e iterativo a partir de prueba y error para las diferentes etapas del ciclo de vida de la IA, siempre con protecciones a los derechos humanos. Se insiste también en la necesidad de innovar en la gestión de riesgos y en la formación de equipos incluyentes, multidisciplinarios y diversos.

En paralelo, distintos países han optado por aproximaciones sectoriales, principios no vinculantes o marcos de autorregulación (como en el caso de Estados Unidos), configurando un mosaico regulatorio internacional todavía en proceso de consolidación y sin una taxonomía unificada, incluso respecto a instrumentos ya relativamente maduros como los propios sandboxes regulatorios.

Frente a la dificultad de anticipar los efectos sistémicos de tecnologías digitales que evolucionan en ciclos de meses y no de años, han ido surgiendo algunas propuestas metodológicas de regulación ágil, principalmente de organismos internacionales como la UIT y la OCDE, que han promovido explícitamente el concepto de "regulación basada en evidencia", que integra evaluación ex ante, monitoreo continuo y evaluación ex post, y que encuentra en los espacios de experimentación regulatoria (sandboxes, laboratorios de política pública, entre otros) su instrumento operativo más visible. Este giro metodológico reconoce que, en contextos de alta incertidumbre técnica y normativa, la calidad regulatoria depende menos de la sofisticación formal del marco jurídico y más de la capacidad de los agentes involucrados para generar conocimiento acumulativo y abrirse al aprendizaje continuo y la adaptación, sin sacrificar salvaguardas ni legitimidad democrática u organizacional en el proceso.

4 UNESCO (2024), Reporte RAM sobre el Estadío de Preparación de México en Inteligencia Artificial. https://www.unesco.org/es/articles/mexico-evaluacion-del-estadio-de-preparacion-de-la-inteligencia-artificial

La UIT, con apoyo de la Unión Europea (Global Gateway), publicó en 2026 el documento Designing Regulatory Experimentation: A practical Guide for the Public Sector. Approaches and Criteria for Designing Regulatory Experimentation and Innovation Spaces[5], cuyo objetivo es dotar a reguladores y autoridades públicas de un marco conceptual y operativo para seleccionar e implementar espacios de experimentación regulatoria frente a la aceleración de tecnologías digitales como la IA, 5G, blockchain e IoT. El documento insiste en que la experimentación no busca desregular, sino mejorar la regulación a través del aprendizaje práctico, y menciona que el sandbox regulatorio no debe convertirse en la herramienta por defecto, pues en contextos institucionales con capacidades limitadas o marcos legales rígidos conviene priorizar instrumentos más simples y proporcionales para cada caso. El documento Regulatory Sandbox Toolkit de la OCDE[6] coincide con los propósitos del de la UIT aunque su propuesta metodológica es distinta y sugiere una prueba más detallada para determinar si el sandbox es el instrumento adecuado.

Además de los documentos mencionados, se han publicado otras propuestas metodológicas como la presentada por Datasphere Sandboxes for AI: Tool for a New Frontier[7], que han proporcionado elementos útiles para ir elaborando un marco básico de la regulación ágil de la IA, todo lo cual ha servido para nutrir de alguna u otra manera este manual.

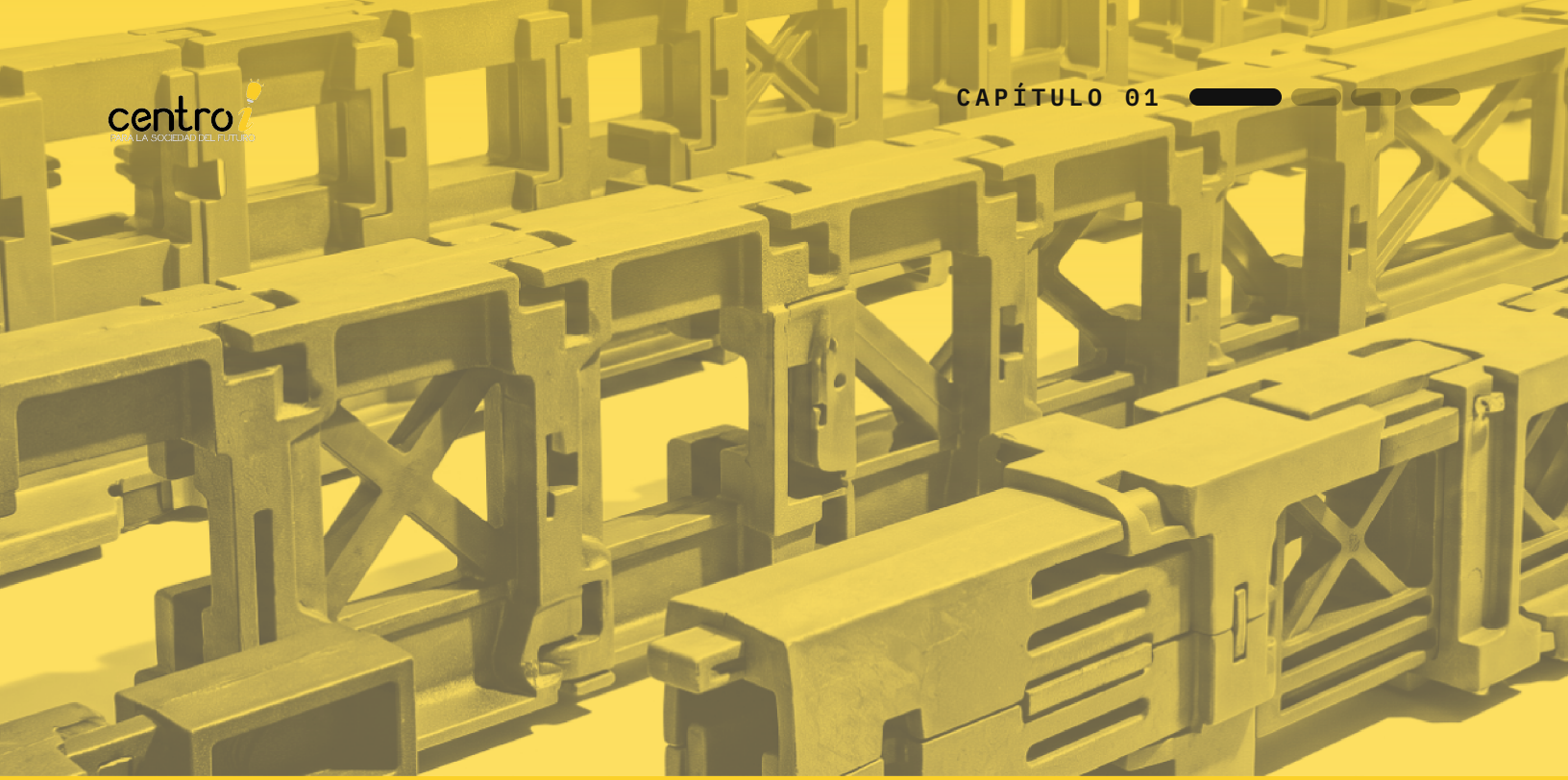
Sin embargo, hay dos aspectos centrales que distinguen este documento de los demás. Uno, que la metodología desarrollada se ha concebido para que pueda ser aplicada tanto por entidades públicas como privadas. Esto implica que el rol del regulador no necesariamente corresponde a una instancia gubernamental, sino según el caso puede ser un área o comité dentro de una empresa o una unidad académica en una universidad que formen parte del esquema de gobernanza diseñado para tal propósito. Y dos, se desarrolla una concepción ecosistémica de la gobernanza y la regulación que aprovecha la robustez del pensamiento sistémico para ampliar la eficiencia y los alcances de los esquemas de gobernanza y regulación.

> Acompañan a este documento varios anexos técnicos disponibles en la página del Centro-i y un paquete de fichas técnicas imprimibles pensadas para facilitar la aplicación de la metodología desarrollada en sesiones colaborativas, que pueden consultarse aquí:
> centroi.org/project/regulacion-agil-para-la-gobernanza-de-la-inteligencia-artificial

5 UIT (2026). https://www.itu.int/hub/publication/d-pref-guide-2026/

6 OCDE (2025), Regulatory Sandbox Toolkit. https://www.oecd.org/en/publications/regulatory-sandbox-toolkit_de36fa62-en.html

7 The Datasphere (2024), Sandboxes for AI Report. https://www.thedatasphere.org/wp-content/uploads/2025/02/Report-Sandboxes-for-AI-2025.pdf



01

¿Cómo se interpretan los conceptos de gobernanza y regulación de la IA en este manual?

1.1 Las definiciones básicas

1.2 La visión tradicional no es suficiente, hay que agilizarla

1.3 Como lo ágil a secas tampoco es suficiente: la concepción ecosistémica

1.4 Pensamiento lineal vs. pensamiento sistémico

1.5 Principios elementales de la gobernanza y la regulación de la IA

1.1 SECCIÓN 1 DE 5

Las definiciones básicas

La gobernanza se refiere a la capacidad de ejercer un control razonable sobre los procesos y toma de decisiones en una organización, que lleven a obtener los fines que se persiguen. Tener mejor gobernanza significa tener mayor capacidad de establecer, mantener y hacer funcionar un diseño institucional cuyas estructuras y procesos estén alineados con los objetivos planteados para cada caso. En concordancia con el enfoque sistémico de este manual, la gobernanza toma la forma de un sistema cuyo propósito es ejercer y mantener control sobre los elementos del mismo para reforzar los impactos positivos y contener los impactos negativos de la IA.[8]

En cuanto al concepto de regulación, lo más común es que se le dé una connotación predominantemente normativa y desde lo público: evoca un marco legal, instrumentos como leyes, decretos y reglamentos, vigilancia, sanciones y la actuación de autoridades administrativas y judiciales. Generalmente, sus objetivos se justifican por la necesidad de corregir problemas o distorsiones provocadas por la actuación de agentes económicos o sociales con el propósito de proteger y salvaguardar el interés público.

Pero como se explica en este documento, la regulación no es un concepto exclusivo de la administración pública pues en última instancia se trata de un sistema de reglas vinculantes que sirven para definir procesos, procedimientos y alinear incentivos entre actores participantes en todo tipo de ámbitos. Inclusive en este ámbito, puede tenerse una visión restringida del concepto para abarcar reglas de hacer o no hacer o que imponen obligaciones y requisitos. El concepto que se adopta en este documento es más amplio, pues incluye diversos instrumentos que de forma directa o indirecta modifican incentivos para provocar comportamientos o decisiones congruentes con los objetivos que se buscan.

En los de la empresa privada, la educación y los temas ciudadanos, por ejemplo, existen dimensiones culturales y vocabularios propios que en lugar de hablar de regulación hablan de gobernanza, de ética y gestión responsable, de pilares de responsabilidad social, de ética en la investigación y de desarrollo o de validación técnica. La tabla siguiente muestra algunas de las equivalencias convencionales del término regulación en esos otros ámbitos.

8 Para un desarrollo más formal de los conceptos de gobernanza y regulación, ver: ¿Cómo construir la gobernanza de la Inteligencia artificial? Recomendaciones para México. Centro-i para la Sociedad del Futuro, 2025.

1.1 SECCIÓN 1 DE 5

TABLA 1.1

Equivalencias convencionales del término “regulación” en diferentes ámbitos

ACTOR / ÁMBITO	TÉRMINO EQUIVALENTE A «REGULACIÓN»	INSTRUMENTOS TÍPICOS	EJEMPLOS CONCRETOS
Gobierno, autoridad pública	Regulación, legislación, normativa, reglamentación	Ley, reglamento, norma, decreto, resolución administrativa	EU AI Act, Ley Fintech, decreto que determina la IA de alto riesgo
Empresa privada	Gobernanza, gestión responsable, cumplimiento, marco ético corporativo	Política interna, código de conducta, estándar corporativo, certificaciones, normas voluntarias	Responsible AI Standard (Microsoft), AI Ethics Board (IBM), ISO 42001, ODS (objetivos de desarrollo sostenible).
Universidad, centro de investigación	Ética de la investigación, integridad científica, protocolos de investigación responsable	Protocolo de comité de ética, reglamento de investigación, directriz académica	Política de uso, protocolo de IRB (Institutional Review Board) extendido.
Organizaciones civiles	Supervisión ciudadana, rendición de cuentas, estándares de derechos humanos	Informe de auditoría ciudadana, estándar de rendición de cuentas, observación ciudadana	Evaluaciones de AI Now Institute, herramientas de evaluación de impacto de IA de Women at the Table
Sector financiero	Cumplimiento regulatorio, gestión de riesgos, controles internos	Marco de riesgo operacional	Marco de riesgo de modelos de IA (BIS/FSB), directrices FCA.[9]
Sector salud	Buenas prácticas clínicas, validación de dispositivos, gestión de evidencia	Protocolo de validación clínica, registro de dispositivo médico con IA	FDA AI/ML Action Plan, norma IEC 62304 para software médico.[10]

Fuente: Elaboración propia.

9 BIS: Bank for International Settlements (Banco de Pagos Internacionales); FSB: Financial Stability Board (Consejo de Estabilidad Financiera); FCA: Financial Conduct Authority, es el regulador de conducta financiera del Reino Unido. Los tres producen los documentos más influyentes sobre gobernanza de modelos de IA en el sector financiero.

10 FDA AI/ML Action Plan: la FDA es la Food and Drug Administration, el regulador federal de alimentos, medicamentos y dispositivos médicos de Estados Unidos. Su AI/ML Action Plan es un documento de política publicado en 2021 que establece el marco regulatorio para sistemas de inteligencia artificial y aprendizaje automático incorporados en dispositivos médicos. IEC 62304: IEC es la International Electrotechnical Commission, el organismo internacional que publica normas técnicas para tecnologías eléctricas y electrónicas. La norma 62304 es el estándar internacional para el ciclo de vida del software en dispositivos médicos.

1.2 SECCIÓN 2 DE 5

La visión tradicional no es suficiente, hay que agilizarla

Es necesario tener claro desde el inicio cuáles son las diferencias entre los conceptos de gobernanza y regulación tradicionales, y los que definen la gobernanza y la regulación ágiles. Estos enfoques divergen profundamente en su lógica de operación, su relación con el tiempo y las actitudes que se le exigen a los actores que forman parte del ecosistema de interés.

No se trata de una dicotomía absoluta: en la práctica, los marcos regulatorios modernos combinan elementos de ambos enfoques según el nivel de riesgo y la madurez de cada sector.

¿Por qué regulación ágil y no un esquema de regulación tradicional?

1. Porque las velocidades de desarrollo, adopción y transformación de la IA vuelven inoperantes los ciclos legislativos, regulatorios y de políticas internas tradicionales.
2. Porque los sistemas de IA son inherentemente contextuales: el mismo modelo puede ser de bajo riesgo en un sector y de alto riesgo en otro.
3. Porque los ciclos de la regulación tradicional la destinan a permanecer obsoleta por largos lapsos, manteniendo una carga de obligaciones innecesarias que pueden inhibir la innovación y las inversiones.
4. Porque la discusión internacional mejor informada al día de hoy indica que es imprescindible contar con reglas y protocolos flexibles y adaptables de aprendizaje continuo.

1.3 SECCIÓN 3 DE 5

Como lo ágil a secas tampoco es suficiente, introducimos la concepción ecosistémica de la gobernanza y la regulación

Queda claro que la efectividad de los esquemas ágiles de gobernanza y regulación de la IA es claramente mayor que la de los esquemas tradicionales, pero este manual incorpora un elemento más a la ecuación, que hace que la eficiencia y la efectividad de los modelos presentados crezca sustancialmente: se trata de la concepción ecosistémica de la gobernanza y de la regulación.

Si abordamos el tema utilizando pensamiento sistémico, la regulación es en sí un sistema de reglas e instrumentos aplicables, simultánea o secuencialmente, a diferentes elementos del ecosistema objetivo con el propósito de incidir en la conducta y en las decisiones de las personas clave para su funcionamiento, de manera que los efectos directos e indirectos de las acciones emprendidas contribuyan al logro de los objetivos planteados.

A la suma de todos los efectos directos e indirectos causados por la aplicación de un sistema específico de reglas e instrumentos regulatorios le llamamos impacto regulatorio. Es importante notar que en muchas ocasiones se generan efectos perversos —en el sentido económico— o encontrados, por lo que el impacto regulatorio es el resultado neto de sumar y restar todos los efectos identificados en un caso dado.

Desde la misma óptica del pensamiento sistémico, y en concordancia con lo que se mencionó en el apartado sobre definiciones básicas, en este manual se concibe a la gobernanza de la IA como un sistema diseñado para poder ejercer y mantener un cierto control sobre el desarrollo y los usos de la IA.[11]

11 Para una explicación detallada de la concepción ecosistémica de la gobernanza y la regulación ágil de la IA, así como de su aplicación en ejemplos concretos, consultar los anexos de este documento.

1.4 SECCIÓN 4 DE 5

La gobernanza y la regulación ágil de la IA en la arena del pensamiento lineal vs. el pensamiento sistémico

Cuando hablamos de la concepción ecosistémica de la gobernanza y la regulación de la IA es oportuno empezar por explicar las diferencias básicas entre el pensamiento lineal y el pensamiento sistémico, lo cual aprovechamos para presentar los argumentos que ayudan a construir la concepción ecosistémica de la gobernanza y la regulación de la IA.

El pensamiento lineal es el método tradicional de solución de problemas enseñado en las escuelas, en las universidades y aplicado repetitivamente en todos los ámbitos profesionales. Es un método que descontextualiza el problema a resolver y busca soluciones mediante la aplicación de recetas probadas, lo cual en la práctica significa que permanece ciego ante el hecho de que los mecanismos objeto de la regulación están siempre anidados en un ecosistema, y que las acciones regulatorias pueden generar efectos directos e indirectos (a menudo identificados como de segundo orden), tanto temporal como espacialmente hablando, y que operan a diferentes velocidades.

De manera radicalmente distinta, el pensamiento sistémico empieza por contextualizar el problema que se quiere resolver, lo cual implica asumir que éste habita en un ecosistema específico que hay que tomar en cuenta. Las ventajas de este enfoque son muy amplias porque surgen de reconocer que aquello que se quiere regular es de naturaleza dinámica, que tanto su entorno como la adopción de la IA seguirán cambiando a diferentes velocidades y que sus efectos espaciales y temporales continuarán reforzando o compensando acciones y procesos de toda índole. La concepción ecosistémica de la gobernanza y de la regulación ágiles permite identificar y estimar mejor los riesgos, las capacidades y los requerimientos intertemporales de recursos de cada uno de los actores participantes, entender mejor las relaciones y los balances de poder entre actores, ya sean sectores, organizaciones o elementos de una misma organización, y por lo tanto diseñar estrategias de solución de naturaleza sistémica que reconocen que el tema no se termina con la aplicación de un instrumento regulatorio durante el periodo acordado, sino que continúa y que habrá que seguir evaluando, aprendiendo y tomando decisiones de muchos tipos.

> Para una explicación más detallada de las diferencias entre el análisis lineal y el pensamiento sistémico, así como de las implicaciones que este último enfoque tiene para los modelos de gobernanza y de regulación ágil, consultar el Anexo 1.

1.5 SECCIÓN 5 DE 5

Principios elementales de la gobernanza y la regulación de la IA

La Recomendación sobre la Ética de la Inteligencia Artificial[12], adoptada por la 41.ª Conferencia General de la UNESCO en noviembre de 2021, constituye el primer instrumento normativo de alcance mundial en esta materia, aplicable a los 194 Estados miembros de la organización. La recomendación sugiere diez principios concretos para la formulación de políticas, leyes y directrices basadas en la ética de la inteligencia artificial, y son los siguientes:

1. **Proporcionalidad e inocuidad.** Los procesos de las tecnologías de la IA no deben rebasar lo necesario para alcanzar objetivos o fines legítimos, y debe recurrirse a una evaluación de riesgos para prevenir daños derivados de dichos usos.
2. **Seguridad y protección.** Deben evitarse riesgos innecesarios y solucionarse cuando los sistemas de IA se utilicen conforme a un fin determinado o explícito, así como en caso de uso secundario, atendiendo a su seguridad y protección a lo largo de todo su ciclo de vida.
3. **Equidad y no discriminación.** Los actores de la IA deben promover la justicia social y salvaguardar la equidad, luchando contra los sesgos y evitando la discriminación de cualquier tipo, conforme al derecho internacional.
4. **Sostenibilidad.** Deben evaluarse constantemente los efectos de las tecnologías de la IA en la sostenibilidad, entendida en sentido amplio, ambiental, social y económico, a la luz de los Objetivos de Desarrollo Sostenible.
5. **Derecho a la intimidad y protección de datos.** La privacidad, derecho esencial para la protección de la dignidad humana, la autonomía y la agencia, debe respetarse, protegerse y promoverse a lo largo de todo el ciclo de vida de los sistemas de IA, junto con la protección de datos.
6. **Supervisión y decisión humanas.** Los Estados miembros deben velar por que los sistemas de IA no menoscaben la posibilidad última de los seres humanos de ejercer la supervisión y de tomar decisiones, y que dicha supervisión pueda materializarse en distintas etapas del ciclo de vida del sistema.
7. **Transparencia y explicabilidad.** El grado de transparencia y explicabilidad de los sistemas de IA debe ser apropiado al contexto, con miras a lograr un equilibrio con otros principios como la privacidad, la seguridad y la protección.
8. **Responsabilidad y rendición de cuentas.** Los actores de la IA y los Estados miembros deben respetar y atender la ética y los derechos humanos en todas las etapas del ciclo de vida de los sistemas de IA, debiendo poderse atribuir en último término la responsabilidad y la rendición de cuentas a personas físicas o entidades jurídicas.
9. **Sensibilización y educación.** Debe promoverse mediante una educación abierta, accesible y de calidad, la alfabetización mediática e informacional y el desarrollo de las competencias digitales, con miras a empoderar a las personas y reducir la brecha digital.
10. **Gobernanza y colaboración adaptativas y de múltiples partes interesadas.** La cooperación internacional y nacional en materia de datos y de IA es necesaria para la aplicación eficaz de los derechos humanos y las libertades fundamentales, requiriendo la participación de múltiples partes interesadas y una gobernanza que se adapte a los rápidos cambios tecnológicos.

12 Documento oficial (versión en español): https://unesdoc.unesco.org/ark:/48223/pf0000380455_spa

1.5 SECCIÓN 5 DE 5

En cuanto a los derechos fundamentales que deben ser protegidos por la gobernanza y la regulación de la IA, partimos de la lista contenida en la Carta Iberoamericana de Principios y Derechos en los Entornos Digitales, aprobada en la XXVIII Cumbre Iberoamericana. En la tabla siguiente se complementan los principios de la Carta Iberoamericana y se identifican algunos de los riesgos de los usos de IA en relación con los derechos considerados; también se mencionan algunas salvaguardas regulatorias y se incluyen algunas de las referencias documentales y legales pertinentes. Este es un marco general aplicable a cualquiera que sea el ámbito en el que se desea construir un esquema de gobernanza y o de regulación ágil.

> La Tabla 1.2 se despliega en las páginas siguientes y se lee con desplazamiento lateral en pantalla; al imprimir se despliega completa, repitiendo el encabezado en cada página.

1.5 SECCIÓN 5 DE 5

TABLA 1.2

Derechos fundamentales que deben ser protegidos por la gobernanza y la regulación de la IA

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
1. CENTRALIDAD DE LA PERSONA – DERECHOS Y DEBERES EN ENTORNOS DIGITALES			
Protección integral de derechos en entornos digitales	Transferencia de sesgos históricos y estructurales del mundo físico al digital, que afectan el ejercicio de derechos humanos, como por ejemplo acceso a la salud, al trabajo, a la autodeterminación. Ausencia de mecanismos de tutela efectiva en plataformas.	Obligación y mecanismos para garantizar, respetar y proteger en entornos digitales todos los derechos reconocidos en la Constitución Política de los Estados Unidos Mexicanos (CPEUM) y demás leyes y reglamentos.	· Art. 1º CPEUM (derechos humanos y su garantía por el Estado). · Art. 1º, párr. 3º CPEUM (obligación de todas las autoridades de promover, respetar, proteger y garantizar los derechos humanos).
No discriminación	Reproducción y/o amplificación algorítmica de discriminación histórica por conceptos de género, etnia, edad, estrato socioeconómico, religión o preferencias políticas.	Políticas de inclusión digital; Garantizar inclusión por diseño, mecanismos para evaluar, detectar y corregir decisiones y dinámicas discriminatorias.	· Art. 1º CPEUM (prohibición de discriminación). · Art. 6º CPEUM (acceso a TIC sin discriminación). · LFTR, Art. 2 (prohibida toda discriminación).
2. INCLUSIÓN DIGITAL Y CONECTIVIDAD			
Derecho de acceso a internet y banda ancha	· Exclusión por barreras de acceso, de uso o de apropiación. · Concentración de infraestructura en zonas urbanas.	· Planes de conectividad y servicio universales · Programas de alfabetización digital generales y específicos. · Estrategias integrales de inclusión. · Incentivos a la inversión en infraestructura e impulso de la competencia.	· Art. 6º CPEUM (acceso a TIC sin discriminación). · LFTR, Art. 2 (prohibida toda discriminación).

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Reducción de desigualdades y brechas digitales (género, etaria, territorial)	Replicación de desigualdades históricas y sistémicas por el diseño y el entrenamiento de la IA a partir de datos sesgados y subrrepresentación de grupos históricamente marginados. Reproducción en los espacios digitales de brechas históricas económicas y sociales.	· Considerar desigualdades interseccionales en diseño y entrenamiento de modelos IA. · Aplicar principios de diversidad e inclusión en todas las fases del ciclo de vida de la IA. · Acciones de inclusión tecnológica para grupos con desventajas sociales y económicas.	· Art. 1º CPEUM (igualdad y no discriminación). · Art. 4º CPEUM (igualdad entre hombre y mujer).
3. PRIVACIDAD, CONFIANZA, SEGURIDAD DE DATOS Y CIBERSEGURIDAD			
Privacidad y protección de datos personales	· Uso y procesamiento masivo de datos sin consentimiento informado o para fines y por actores no autorizados. · Perfilado con IA · Reconocimiento facial y biometría sin consentimiento. · Fallas en ciberseguridad.	· Marcos de protección de datos (privacidad por diseño, privacidad por defecto) · Evaluaciones de impacto. Reguladores autónomos. · Actualización de mecanismos de protección de datos para responder a avances tecnológicos.	· Art. 16º CPEUM, párr. 2º: derecho a la protección de datos personales (ARCO). · Art. 6º, Apt. A, fracc. II CPEUM. · LFPDPPP (2010) y LGPDPPSO (2017).
Soberanía digital personal	Manipulación algorítmica de decisiones y comportamientos sin consentimiento informado (patrones oscuros)*	· Las personas deben conservar el control sobre su identidad biológica y digital, decidir quién puede capturarla, para qué fines, por cuánto tiempo, con qué límites y bajo qué controles. · Obligación de revelar el uso de algoritmos · Prohibición de técnicas subliminales · Derechos ARCO-Plus*. · Salvaguardas frente a las neurotecnologías	· Tesis I.20o.A.56 A (12a.) de la SCJN, 2026. · Art. 16º CPEUM (derechos ARCO). · LFPDPPP, Arts. 22-27 (derechos de acceso, rectificación, cancelación y oposición).

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Seguridad digital y ciberseguridad	· Uso de IA para realizar ciberataques. · Uso de software malicioso (ransomware)*	· Seguridad por diseño (security by design) · Estrategias nacionales y sectoriales de ciberseguridad. · Cooperación iberoamericana.	· Art. 16º CPEUM. párr. 12 (inviolabilidad de las comunicaciones). · LFTR, Arts. 189-192 (obligaciones de seguridad de concesionarios). · Estrategia Nacional de Ciberseguridad 2017.
4. ACCESO PLENO A LA EDUCACIÓN, LA CULTURA Y LA SALUD EN ENTORNOS DIGITALES INCLUSIVOS Y SEGUROS			
Derecho a la educación digital y alfabetización tecnológica	· Brechas en el acceso tecnológico de la población y dentro del sistema educativo. · Falta de preparación de escuelas, empresas y universidades en IA y educación.	· Protocolos para uso de IA en distintos campos de la educación (desarrollo de materiales, uso por alumnos, docentes, directivos y autoridades educativas). · Actualización de currícula, modernización de estrategias pedagógicas y sistemas de evaluación de conocimientos.	· Art. 3º CPEUM (derecho a la educación de calidad). · Ley General de Educación, Art. 9, 17, 24, 52, 84, 85, 86 (uso de TIC y tecnología digital en educación).
Salud mental en entornos digitales	· Daños cognitivos o comportamientos adictivos por sobredependencia de la IA. · Explotación intencionada de heurísticas por la IA.* · Resignación y pérdida de agencia como consecuencia del servilismo algorítmico*	· Regulación de diseños adictivos. · Protección especial para menores y grupos vulnerables. · Enseñanza de pensamiento crítico, pensamiento sistémico, y herramientas de metacognición.*	· Art. 4º CPEUM (derecho a la protección de la salud). · NOM-035-STPS-2018 (factores de riesgo psicosocial, aplicable a entornos digitales laborales).

13 Estos son los términos utilizados en la legislación mexicana, si bien reflejan sesgos y son revictimizantes, por lo que hay una corriente mundial hacia su sustitución.

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Acceso pleno a la cultura en entornos digitales	· Geobloqueo. · Pérdidas culturales por exclusión algorítmica. · Modelos con sesgos culturales, lingüísticos y geográficos debido a que los datos de entrenamiento provienen de culturas hegemónicas.	· Regulación basada en principios de inclusión y no discriminación. · Promoción de contenidos en lenguas nacionales y originarias. · Neutralidad de red.	· Art. 4º CPEUM (acceso a la cultura). · Ley Federal del Derecho de Autor (transmisión de derechos patrimoniales; medidas tecnológicas de protección, información, gestión de derechos y proveedores de servicios de internet)
5. ESPECIAL ATENCIÓN A NIÑAS, NIÑOS Y ADOLESCENTES			
Protección de niñas, niños y adolescentes (NNA) en entornos digitales	· Explotación maliciosa de vulnerabilidades psicoemocionales, socioeconómicas y culturales. · Perfilado con fines perversos, comerciales, o delictivos. · Exposición a contenidos violentos o de odio · Reclutamiento para realizar acciones violentas o polarizantes · Grooming y ciberacoso · Generación y distribución de imágenes sintéticas de índole sexual	· Protección de datos de NNA · Regulación de publicidad dirigida a NNA · Plataformas seguras por diseño · Agilización de mecanismos de prevención, detección y denuncia	· Art. 4º CPEUM (derechos de la niñez; interés superior del menor). · Ley General de los Derechos de Niñas, Niños y Adolescentes. · LFPDPPP (tratamiento datos de menores). · Código Penal Federal (pornografía infantil y corrupción de menores[13]). · Ley Olimpia (violencia digital). · Declaración Conjunta sobre Inteligencia Artificial y Derechos del Niño (UIT, UNICEF, UNESCO, OIT)

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
6. PARTICIPACIÓN SOCIAL, ECONÓMICA Y POLÍTICA EN ENTORNOS DIGITALES			
Libertad de expresión e información en entornos digitales	Manipulación algorítmica de mensajes y opiniones. Desinformación amplificada por IA. Censura/polarización disimuladas por plataformas.	· Transparencia en criterios de moderación. · Mecanismos de apelación. Supervisión independiente. Prohibición de censura previa.	· Art. 6º CPEUM (manifestación de ideas; acceso a información). · Art. 7º CPEUM (libertad de difusión de opiniones e ideas por cualquier medio).
Acceso a la información pública y transparencia	· Opacidad de algoritmos gubernamentales. · Uso de IA en toma de decisiones públicas sin rendición de cuentas.	· Obligación de publicación proactiva de información pública · Auditorías transparentes de algoritmos del sector público · Datos abiertos, gobernanza de datos.	· Art. 6º, Apt. A CPEUM (derecho de acceso a la información pública). · Ley General de Transparencia y Acceso a la Información Pública (2015).
Participación digital en política y democracia	Uso de IA para microperfilamiento y personalización con propósitos de desinformación político-electoral y erosión de la democracia	· Regulación de publicidad política digital. · Regulación de bots. · Moderación para la pluralidad informativa. · Investigación y sanciones a las campañas de desinformación. · Combate a la violencia digital y discurso de odio.	· Art. 6º CPEUM (derecho a la información pública). · Art. 35 CPEUM (derechos políticos). · Ley General de Instituciones y Procedimientos Electorales (propaganda en medios digitales).

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Sostenibilidad ambiental	Muy alto consumo energético y de agua en centros de datos y de cómputo.	· Estrategias que Incentiven el uso de tecnologías sostenibles. · Educación sobre el consumismo tecnológico. · Estándares de eficiencia energética para infraestructura digital. · Políticas que incentiven la economía circular.	· Art. 4º CPEUM (derecho al medio ambiente sano). · Ley General del Equilibrio Ecológico y la Protección al Ambiente (LGEEPA).
7. ADMINISTRACIÓN PÚBLICA DIGITAL			
Debido proceso en decisiones algorítmicas del Estado	· Decisiones automatizadas en justicia penal o crédito social. · Otorgamiento automatizado de ayudas sociales y servicios públicos sin revisión humana significativa, ni explicabilidad, ni apelación.	· Derecho a explicación de decisiones automatizadas; revisión humana obligatoria en contextos de alto impacto; mecanismos de impugnación. · Prohibición o limitación de uso de sistemas de IA en casos de riesgos elevados para los derechos humanos.	Art. 16º CPEUM (fundamentación y motivación de actos de autoridad).

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Identidad digital e inclusión	· Costos crecientes para ciudadanos excluidos del ecosistema digital. · Fraude de identidad.	· Trámites digitales inclusivos, sencillos e interoperables. · Identidad digital cibersegura con opción de canal presencial alternativo. · Sistemas de identidad digital ciberseguros, fiables, inclusivos e interoperables. · Seguridad por diseño, minimización, altos umbrales de justificación para uso de datos biométricos.	Ley General de Población, CURP biométrica.
8. ECONOMÍA DIGITAL JUSTA, INCLUSIVA Y SEGURA			
Acceso a servicios esenciales (crédito, seguro, empleo) sin discriminación algorítmica	· Exclusión por sistemas opacos de puntuación social o decisiones opacas automatizadas en créditos, seguros o empleos. · Uso de sistemas automatizados sin evaluación de sesgos.	· Obligación de corrección de sesgos estructurales y sistémicos en bases de datos · Principios obligatorios de no discriminación y de eliminación de sesgos por diseño. · Evaluación de impacto ético de sistemas. · Mecanismos funcionales de apelación y denuncia	· Art. 1º CPEUM (no discriminación). · Ley para Regular las Instituciones de Tecnología Financiera – Ley Fintech (2018). · CNBV – regulación de modelos de crédito.
Derechos laborales en la economía digital y teletrabajo	· Microperfilamiento para vigilancia laboral abusiva. · Precarización del trabajo de etiquetado de contenidos.	· Actualización de marcos laborales para teletrabajo. · Derecho a la desconexión digital. · Protección de datos de trabajadores.	· Art. 123 CPEUM (derechos laborales). · LFT, Art. 311 Bis (teletrabajo, reforma 2021). · NOM-037-STPS-2023 (teletrabajo).

TABLA 1.2 · CONTINUACIÓN

DERECHO / PRINCIPIO	ALGUNOS RIESGOS DE LA IA EN RELACIÓN CON EL DERECHO	SALVAGUARDA REGULATORIA	REFERENCIAS DOCUMENTALES Y LEGALES
Protección al consumidor digital	· Prácticas abusivas en comercio electrónico Manipulación de precios dinámicos con IA · Microperfilamiento de las personas consumidoras.	· Nivel de protección al consumidor digital no inferior al comercio físico. · Transparencia en algoritmos de recomendación y de precios. · Prohibición de explotación de los datos de las personas con fines de personalización abusiva de de las ofertas comerciales y la publicidad.	· Art. 28 CPEUM (protección al consumidor). · Ley Federal de Protección al Consumidor.
9. TECNOLOGÍAS EMERGENTES CON CENTRALIDAD DE LAS PERSONAS			
Uso ético y responsable de la inteligencia artificial	· Inexistencia, y en su caso, sistemas de gobernanza y regulación de la IA mal diseñados. · Opacidad algorítmica. · Uso altamente riesgoso de IA, por ejemplo, en armamento autónomo o para vigilancia masiva	Marcos de gobernanza de IA basados en derechos humanos; evaluaciones de impacto; certificación de sistemas de alto riesgo; órganos de supervisión.	· Recomendación UNESCO sobre la Ética de la IA. · Principios de OCDE para la IA (México es miembro).
Neurotecnologías y protección de datos neuronales	· Lectura y modificación de actividad cerebral sin consentimiento informado. · Perfilado cognitivo con IA. Neuromarketing invasivo.	· Actualización de marcos de protección de datos para incluir datos neuronales. Evaluación ética de neurotecnologías. · Moratoria en usos no médicos.	· Art. 16º CPEUM (datos personales sensibles; inviolabilidad de comunicaciones). · LFPDPPP.

Fuente: elaboración propia con base en la Carta Iberoamericana de Principios y Derechos en los Entornos Digitales.

1.5 SECCIÓN 5 DE 5

Enfoque diferenciado para grupos vulnerables

Los sistemas de IA pueden tener impactos mayores en personas y comunidades que ya enfrentan barreras y desigualdades estructurales. La gobernanza y la regulación éticas deben incorporar análisis diferenciados que lo consideren.

Principio de inclusión en el proceso regulatorio

- Los grupos potencialmente afectados o excluidos deben participar en el diseño de la gobernanza y la regulación: no basta con que formen parte de procesos de consulta, sino que estén representados en los equipos que diseñan, ejecutan y evalúan los mecanismos.
- Los procesos de consulta deben estar disponibles en formatos accesibles y en los idiomas pertinentes.
- Las evaluaciones de impacto deben incluir análisis desagregados por género, etnia, edad, nivel socioeconómico y otras categorías.
- El monitoreo postdespliegue debe medir impactos diferenciados, no sólo rendimiento agregado.

02

Tipología y comparación de modelos de regulación ágil

Doce instrumentos que pueden ser usados de manera complementaria, en función de las necesidades del caso que se aborde.

CATÁLOGO

La necesidad de contar con marcos de gobernanza y regulación más adecuados para hacer frente a la IA ha generado un abanico cada vez más amplio de instrumentos regulatorios. En este manual el abanico incluye doce instrumentos que pueden ser usados de manera complementaria, en función de las necesidades del caso que se aborde.

DIAGRAMA 2.1

Catálogo de instrumentos de gobernanza y regulación ágil

01 **Sandboxes regulatorios** LIMITADA	**02** **Pruebas piloto supervisadas** LIMITADA, CON CLÁUSULAS DE SALIDA	**03** **Marco de incentivos (soft law) con revisión periódica** PERMANENTE CON REVISIONES	**04** **Marcos basados en principios** PERMANENTE
05 **Protocolos adaptativos y modulares** CONTINUA	**06** **Autorregulación, corregulación y certificación** PERMANENTE CON EVALUACIONES	**07** **Regulación basada en resultados** PERMANENTE	**08** **Gobernanza anticipatoria** PERMANENTE Y CONTINUA
09 **Tecnología regulatoria y de supervisión** CONTINUA, EN TIEMPO REAL	**10** **Hubs de innovación regulatoria** PERMANENTE COMO ESTRUCTURA	**11** **Puerto seguro con divulgación** PERMANENTE COMO RÉGIMEN	**12** **Cláusulas de revisión, entrada en vigor y expiración** INSCRITA EN LA NORMA

VINCULATORIEDAD ALTA · VARIABLE · NO VINCULANTE · VINCULATORIEDAD MEDIA

Fuente: Elaboración propia.

La Tabla 2.1 compara los doce instrumentos en términos de su locus de aplicación, grado de vinculatoriedad, temporalidad, objetivos, fortalezas, riesgos, capacidades institucionales requeridas y su relación con los otros modelos. El análisis comparativo de estos instrumentos de regulación ágil se puede encontrar sistematizado en fichas esquemáticas en los anexos de este documento: centroi.org/regulacion-agil-ia/anexo-6-fichas-para-imprimir

TABLA 2.1

Comparación de modelos de regulación ágil para la IA

DIMENSIÓN	1 · SANDBOXES REGULATORIOS	2 · PRUEBAS PILOTO SUPERVISADAS
Definición	· Sistema experimental que permite probar productos/ servicios innovadores con excepciones normativas temporales y supervisión intensiva · Hay dos tipos genéricos: · el sandbox solicitado por el postulante de una innovación, y el sandbox postulado por el gobierno para probar una nueva regulación	Experimentos en entornos reales para probar nuevos esquemas de regulación, servicios o formas de aplicación normativa
Locus posible de aplicación	Sectores público y privado	Sectores público y privado
Grado de vinculatoriedad	Alto dentro del sandbox (condiciones específicas)	Variable (puede incluir obligaciones legales)
Temporalidad	Limitada pero puede dar lugar a un marco o acción regulatoria permanente	Limitada, con cláusulas de salida, pero puede dar lugar a un marco o acción regulatoria permanente
Objetivos principales	· Aprendizaje regulatorio en tiempo real · Reducir barreras a innovadores · Generar evidencia para reformas	· Probar intervenciones públicas antes de generalizar · Obtener evidencia robusta · Diseño participativo
Fortalezas	· Libertad para experimentar de manera controlada · Aprendizaje conjunto de todas las partes	· Evidencia robusta · Permite ajustes predespliegue · Más participación ciudadana
Riesgos y limitaciones	· Captura regulatoria · Fragmentación normativa · Posible exposición de usuarios a riesgos · Altos costos · Limitada escalabilidad · Necesario renovarla con cada innovación	· Requiere alta capacidad institucional · Altos costos · Limitada escalabilidad · Necesario renovarlo con cada innovación
Capacidades institucionales requeridas	· Base jurídica clara · Diseño experimental · Monitoreo continuo · Recolección sistemática de datos · Evaluación ex post	· Base jurídica clara · Monitoreo continuo · Diseño experimental · Recolección sistemática de datos · Evaluación ex post
Ejemplos de referencia	· Sandboxes FCA (Reino Unido)[14] · Sandboxes de IA (OCDE)[15] · Sandbox de redes eléctricas (CNMC)[16] · Sandbox de IA (MTDFP y AESIA de España)[17] · Sandboxes de temas diversos (ICO, Reino Unido)[18]	Pilotos de redes inteligentes (UE)[19]
Relación con otros modelos	Complementa pilotos; alimenta marco de incentivos y legislación	Complementa sandboxes; alimenta marco de incentivos y legislación

14 fca.org.uk/firms/innovation/regulatory-sandbox **15** oecd.ai/en/sandboxes **16** cnmc.es/prensa/sandbox-redes-electricas-20250128 **17** digital.gob.es · primer sandbox de inteligencia artificial
18 ico.org.uk/for-organisations/advice-and-services/regulatory-sandbox **19** bridge-smart-grid-storage-systems-digital-projects.ec.europa.eu

Fuente: Elaboración propia.

CONTINÚA EN LA HOJA SIGUIENTE

TABLA 2.1

Comparación de modelos de regulación ágil para la IA

HOJA 2 DE 5

INSTRUMENTOS 3 A 5 DE 12 01 02 03 04 05 06 07 08 09 10 11 12

DIMENSIÓN	3 · MARCO DE INCENTIVOS (SOFT LAW) CON REVISIÓN PERIÓDICA	4 · MARCOS BASADOS EN PRINCIPIOS	5 · PROTOCOLOS ADAPTATIVOS Y MODULARES
Definición	· Instrumentos no vinculantes (recomendaciones, principios, códigos, incentivos conductuales, guías, sellos de confianza, distintivos) y ciclos de actualización · Puede basarse en diseños específicos de economía conductual (reforzado) o en buenas prácticas (ligero).	Enunciados amplios sobre resultados deseados que dejan margen a los regulados para diseñar los mecanismos para su cumplimiento	Conjuntos de controles organizados en módulos (privacidad, seguridad, transparencia) que se actualizan según nueva evidencia
Locus posible de aplicación	Organismos internacionales, proyectos con múltiples actores	Sectores público y privado	Sectores público y privado
Grado de vinculatoriedad	Bajo (no exigible judicialmente, pero influyente)	Medio-alto (principios obligatorios, implementación flexible)	Variable (voluntario a obligatorio según sector)
Temporalidad	Permanente con revisiones periódicas	Permanente (principios estables)	Continua (actualizaciones modulares)
Objetivos principales	· Puente temporal hacia legislación o complemento a la regulación · Coordinación internacional · Experimentación normativa de conceptos · Flexibilidad	· Flexibilidad ante la innovación · Énfasis en cultura organizacional y resultados · Reducir rigidez prescriptiva	· Sincronizar regulación con cambio tecnológico · Actualizar sin reformas legislativas · Integrar estándares
Fortalezas	· Flexibilidad de actualización · Facilita la coordinación internacional · Incidir en los incentivos permite disminuir la necesidad de prescribir conductas	· Adaptable a nuevos modelos de negocio · Fomenta la responsabilidad organizacional · Evita obsolescencia	· Permite actualizaciones parciales · Integrable con certificaciones · Escalable por capas
Riesgos y limitaciones	· Falta de exigibilidad · Fragmentación por proliferación de mecanismos · Riesgo de simulación (AI-washing)	· Incertidumbre jurídica si los principios son vagos · Requiere reguladores muy capacitados · Puede volver a lógica de reglas	· La capacidad técnica dificulta la vigilancia ciudadana · Exige gobernanza de datos robusta · Complejidad de coordinación
Capacidades institucionales requeridas	· Procesos de consulta multiparte · Transparencia y amplia difusión · Mecanismos de actualización	· Supervisión cualitativa · Análisis de gobernanza corporativa · Evaluación basada en riesgos para el diseño de mecanismos de cumplimiento	· Gobernanza de datos avanzada · Capacidad técnica para definir módulos · Coordinación con estándares
Ejemplos de referencia	· Recomendación UNESCO sobre ética de IA · Principios OCDE sobre IA · Principios G20	· Principios de FSA/FCA Reino Unido ('tratar clientes de forma justa') · Marcos OCDE, UNESCO, G7 para IA	· NIST AI Risk Management Framework · Arquitecturas de cumplimiento modulares · ISO/IEC 42001[20]
Relación con otros modelos	· Refuerza las conductas deseadas, puede disminuir la necesidad de cierta regulación · Base de conocimiento para marcos por principios y corregulación	Orienta otros modelos de regulación al establecer objetivos	Operacionaliza principios en protocolos técnicos

20 iso.org/es/norma/42001

Fuente: Elaboración propia.

CONTINÚA EN LA HOJA SIGUIENTE

TABLA 2.1

Comparación de modelos de regulación ágil para la IA

HOJA 3 DE 5

INSTRUMENTOS 6 A 8 DE 12 01 02 03 04 05 06 07 08 09 10 11 12

DIMENSIÓN	6 · AUTORREGULACIÓN, CORREGULACIÓN Y CERTIFICACIÓN	7 · REGULACIÓN BASADA EN RESULTADOS	8 · GOBERNANZA ANTICIPATORIA
Definición	Variaciones desde la autoimposición autónoma de estándares y conductas, a otros esquemas donde colaboran los reguladores con autoridades y organismos certificadores o de normalización para aplicar estándares técnicos y certificaciones verificables	El regulador define objetivos o resultados que deben alcanzarse, sin estipular los medios técnicos para lograrlo	Uso de análisis prospectivo y de señales tempranas para iniciar diseño de gobernanza y regulación antes de la materialización de ciertos umbrales
Locus posible de aplicación	Sectores público y privado	Sectores público y privado	Sectores público y privado
Grado de vinculatoriedad	Medio (corregulación con respaldo legal y auto rregulación con norma de reserva)	Alto en los resultados; flexible en los medios	Variable, pero sus alertas pueden activar instrumentos obligatorios
Temporalidad	Permanente con evaluaciones periódicas	Permanente; revisión de indicadores con mayor frecuencia que la norma prescriptiva	Permanente y continua
Objetivos principales	· Reacción rápida a cambios · Aprovechar el conocimiento experto · Reducir la carga estatal	· Fomentar innovación en medios de cumplimiento · Evitar obsolescencia técnica de la norma · Reducir carga administrativa	· Anticipar riesgos antes de su materialización · Informar el diseño regulatorio con evidencia prospectiva
Fortalezas	· Velocidad de respuesta · Estándares técnicamente realistas · Combina conocimiento público-privado	· Estabilidad regulatoria a largo plazo · Incentiva innovación en métodos de cumplimiento · Más fácil de actualizar	· Proactividad: actúa antes del daño · Mejora la calidad del diseño normativo · Fomenta deliberación pública sobre futuros tecnológicos
Riesgos y limitaciones	· Sesgo proindustria · Falta de transparencia · Sin salvaguarda legal, riesgo de simulación (AI-washing)	· Dificultad para definir y medir resultados robustos · Riesgo de manipulación de indicadores · Mayor exigencia de capacidad supervisora	· Costoso en recursos de inteligencia estratégica · Resultados inciertos (el futuro es indeterminado) · Riesgo de parálisis por análisis
Capacidades institucionales requeridas	· Supervisión de cumplimiento de códigos de conducta · Auditoría de certificaciones · Capacidad de activar salvaguardas legales (regulación contingente)	· Sistema robusto de métricas de resultado · Capacidad de auditoría técnica para verificar evidencia presentada	· Unidad de prospectiva o inteligencia tecnológica · Metodologías de señales de futuro · Articulación con el proceso legislativo
Ejemplos de referencia	· NOM mexicanas que son definidas por Comités Consultivos Nacionales de Normalización · Certificaciones ISO · Consejo de Autorregulación y Ética Publicitaria de México	· Libro Blanco de IA del Reino Unido[21] · Las reglas de FCA son la aplicación práctica de los principios del libro blanco	· OECD AI Policy Observatory[22] · Unidades de prospectiva y AI Safety Institutes de distintos países como de Australia, Canadá, Estonia, Finlandia · AI Watch (Comisión Europea)
Relación con otros modelos	Concreta principios y compromisos en mecanismos y estándares verificables	Complementa la regulación basada en principios añadiendo métricas verificables; habilita las tecnologías de supervisión para supervisar resultados	Nutre todos los demás modelos con inteligencia estratégica; activa sandboxes y protocolos adaptativos cuando detecta señales de riesgo

21 gov.uk/government/publications/ai-regulation-a-pro-innovation-approach **22** oecd.ai

Fuente: Elaboración propia.

CONTINÚA EN LA HOJA SIGUIENTE

TABLA 2.1

Comparación de modelos de regulación ágil para la IA

HOJA 4 DE 5

INSTRUMENTOS 9 A 10 DE 12 01 02 03 04 05 06 07 08 09 10 11 12

DIMENSIÓN	9 · TECNOLOGÍA REGULATORIA Y TECNOLOGÍA DE SUPERVISIÓN	10 · HUBS DE INNOVACIÓN REGULATORIA
Definición	Uso de tecnología (IA, big data, API) para automatizar el cumplimiento normativo (RegTech) y la supervisión regulatoria (SupTech)	Espacio institucional de orientación y diálogo entre reguladores e innovadores para clarificar marcos aplicables antes de aplicar un modelo de regulación ágil
Locus posible de aplicación	Sectores público y privado	Sectores público y privado
Grado de vinculatoriedad	Variable: puede ir de voluntario a obligatorio	No vinculante: orientación informal sin sanciones directas
Temporalidad	Continua y en tiempo real (monitoreo automatizado)	Permanente como estructura; interacciones puntuales según demanda
Objetivos principales	· Reducir costos de cumplimiento · Escalar la supervisión regulatoria sin aumentar proporcionalmente el personal	· Reducir incertidumbre jurídica para innovadores · Facilitar el acceso al sistema regulatorio · Identificar candidatos a sandbox o pruebas piloto
Fortalezas	· Supervisión continua y escalable · Reduce asimetrías de información entre regulador y supervisado · Menor costo marginal por entidad supervisada	· Claridad temprana para innovadores · Reduce costos de entrada al mercado regulado · Fuente de inteligencia para el regulador
Riesgos y limitaciones	· Dependencia tecnológica de terceros proveedores · Concentración de mercado en RegTech · Errores, omisiones y sesgos por gobernanza inadecuada y desplazamiento del juicio humano	· No vinculante: riesgo de ignorarse · Puede convertirse en cuello de botella si está mal dotado · Riesgo de captura por los mejor relacionados
Capacidades institucionales requeridas	· Infraestructura de datos compartida · APIs seguras con supervisión · Equipo técnico en el regulador (científicos de datos) · Confianza institucional en el uso de los datos, seguridad y privacidad	· Personal dedicado con conocimiento técnico-jurídico · Protocolo de orientación estandarizado · Coordinación entre reguladores
Ejemplos de referencia	· Plataformas SupTech del BIS y su Financial Stability Institute · Sistemas de supervisión usados por bancos centrales[23] · Uso de IA por SAT para obligaciones fiscales (México)	· AI and Digital Hub (Reino Unido)[24] · Innovation Hub FCA (Reino Unido)[25] · Plataforma GRIP, Global Regulatory Innovation Platform (WEF/UAE, 2025)[26]
Relación con otros modelos	Herramienta transversal: puede implementar cualquier otro modelo (monitorear sandboxes, verificar principios, auditar protocolos)	Puerta de entrada al ecosistema regulatorio; filtra qué casos van a sandbox vs. marco de incentivos vs. autorregulación y corregulación

23 centralbanking.com · RegTech y SupTech en bancos centrales (2026) **24** drcf.org.uk/ai-and-digital-hub **25** fca.org.uk/firms/innovation **26** initiatives.weforum.org/global-regulatory-innovation-platform

Fuente: Elaboración propia.

CONTINÚA EN LA HOJA SIGUIENTE

TABLA 2.1

Comparación de modelos de regulación ágil para la IA

HOJA 5 DE 5

INSTRUMENTOS 11 A 12 DE 12 01 02 03 04 05 06 07 08 09 10 11 12

DIMENSIÓN	11 · PUERTO SEGURO CON DIVULGACIÓN	12 · CLÁUSULAS DE REVISIÓN, ENTRADA EN VIGOR Y EXPIRACIÓN
Definición	Protección jurídica parcial siempre que no se rebasen umbrales o se cumpla con mínimos	· Disposiciones que programan automáticamente la entrada en vigor (sunrise) o expiración (sunset) de obligaciones, vinculadas a hitos o condiciones de madurez tecnológica · Obligaciones de revisión periódica de marcos regulatorios
Locus posible de aplicación	Sectores público y privado	· Sectores público y privado · Procesos legislativos o reglamentarios
Grado de vinculatoriedad	Voluntario con incentivo fuerte: protección frente a sanciones a cumplimiento de mínimos	Alto: la norma se activa o expira automáticamente al cumplirse el hito o la condición o bien está activa continuamente hasta la siguiente revisión
Temporalidad	Permanente como régimen; se activa ante cada hito, condición o incidente reportado	La norma lleva inscrita su propia temporalidad
Objetivos principales	· Invertir los incentivos: premiar la divulgación honesta, el cumplimiento de mínimos o el avance, en vez de sancionar el error · Generar aprendizaje colectivo	· Evitar la regulación prematura o perpetua · Programar la caducidad del régimen cuando la incertidumbre se resuelva · Abrir espacio de evaluación y actualización
Fortalezas	· Fomenta la cultura de transparencia y cumplimiento · Genera datos de incidentes para el ecosistema · Reduce la subnotificación por temor a sanciones	· Seguridad jurídica sobre el cambio · Obliga a una revisión activa · Evita la perpetuación de normas obsoletas
Riesgos y limitaciones	· Requiere una definición precisa de condiciones de elegibilidad · Puede dejar descubiertos a terceros por daños causados · Riesgo de uso estratégico para eludir sanciones	· Complejidad de diseño de las condiciones detonantes · Riesgo de activación prematura o tardía · Requiere capacidad de rediseño rápido · Los procesos de revisión son costosos para los reguladores, aunque pueden ser eficientes para el ecosistema
Capacidades institucionales requeridas	· Registro centralizado de incidentes y evolución de condiciones · Protocolo de evaluación de reportes · Equipo de análisis de causas raíz	· Expertise en redacción normativa condicional · Sistema de monitoreo de condiciones detonantes y evaluación de efectividad · Proceso eficiente de revisión
Ejemplos de referencia	· Sistemas de reporte de incidentes de aviación civil (OACI) · Divulgación responsable de vulnerabilidades en ciberseguridad[27] · Propuesta WEF/GRIP (enero 2026)[28]	· Revisiones periódicas de evaluación de la AI Act de UE · Revisión bienal de regulación de preponderancia de telecomunicaciones y radiodifusión (México)
Relación con otros modelos	Complementa sandboxes y pilotos; reduce la subnotificación que debilita el aprendizaje regulatorio de nuevos modelos	· Marco temporal que puede envolver cualquier otro modelo: un sandbox puede tener cláusula de ocaso; un protocolo adaptativo puede tener cláusula de amanecer · Cualquier regulación puede revisarse periódicamente

27 weforum.org/publications/global-cybersecurity-outlook-2026 **28** initiatives.weforum.org/global-regulatory-innovation-platform

Fuente: Elaboración propia.

FIN DE LA TABLA 2.1

03

Guía para la selección e implementación de modelos sistémicos

3.1 Visión y fundamento: el mapa sistémico

3.2 Clasificación de riesgos

3.3 Matriz impacto-certidumbre

3.4 Las capas de ajuste transversales

3.5 Nota sobre las combinaciones de modelos

3.6 Metodología genérica

3.7 Metodología para un sandbox regulatorio

3.1 SECCIÓN 1 DE 7

Visión y fundamento: el mapa sistémico del proyecto de gobernanza o regulación

Para poder diseñar un esquema de gobernanza y/o mecanismos de regulación, el primer paso es entender de la mejor manera posible el funcionamiento del caso objetivo. En concordancia con la argumentación desarrollada en las secciones previas, el primer paso para entender a fondo el funcionamiento del caso objetivo es mapear el sistema que lo representa, y ello implica abordar el caso de interés desde la perspectiva del pensamiento sistémico.

El mapa sistémico es una herramienta muy poderosa porque la información que contiene va mucho más allá de la visualización de los actores y de los elementos que interactúan para dar forma al proyecto que nos interesa gobernar y/o regular. El mapa nos muestra que hay actores de diferentes jerarquías funcionales, esto es, que hay actores y elementos críticos sobre los que conviene incidir para influir sobre el funcionamiento del sistema; nos ayuda a identificar las capacidades que tienen y los recursos que requiere cada actor; nos facilita la comprensión del juego de poder siempre existente entre actores, sean entidades públicas, divisiones académicas o direcciones y departamentos dentro de una empresa privada; nos muestra los ciclos de compensación o de refuerzo de los procesos, así como las diferentes velocidades que definen sus efectos temporales o espaciales; nos enseña con toda claridad los efectos de primer y segundo orden, es decir, los efectos e impactos que generan en cadena las decisiones y acciones del sistema; y, entre muchas otras ventajas, nos permite entender el origen de los sesgos y complicaciones que enlodan el proyecto, así como la paleta de estrategias sistémicas con las que podemos resolverlo. El mapa nos muestra de manera transparente por qué las causas de los problemas que queremos resolver hoy, son las decisiones y acciones que tomamos anteriormente.

3.3 SECCIÓN 3 DE 7

GRÁFICA 3.1

Matriz impacto-certidumbre para la selección de instrumentos de regulación ágil para IA

CERTIDUMBRE SOBRE LOS EFECTOS: ALTA / BAJA

A · Impacto bajo + efectos bien conocidos

Foco: fomentar buenas prácticas, actualizar conforme cambien las condiciones

INSTRUMENTOS PRINCIPALES

- Marco de incentivos con revisión periódica
- Marcos basados en principios
- Cláusulas de entrada/expiración
- Cumplimiento voluntario de buenas prácticas
- Hubs de innovación (orientación)

B · Impacto alto + efectos bien conocidos

Foco: garantizar cumplimiento verificable y rendición de cuentas robusta

INSTRUMENTOS PRINCIPALES

- Corregulación y certificación
- Regulación basada en resultados
- SupTech y auditoría obligatoria
- Puerto seguro con divulgación
- Protocolos adaptativos
- RegTech de monitoreo continuo
- Evaluación de impacto (EIDF)

C · Impacto bajo + efectos inciertos

Foco: aprender antes de formalizar, generar evidencia bajo control

INSTRUMENTOS PRINCIPALES

- Piloto supervisado ligero
- Sandbox exploratorio
- Gobernanza anticipatoria
- Marco de incentivos experimental
- Hub de innovación regulatoria
- Metodologías de señales de futuro
- Marcos de principios exploratorios

D · Impacto alto + efectos inciertos

Foco: proteger mientras se genera evidencia bajo supervisión estricta

INSTRUMENTOS PRINCIPALES

- Sandbox estricto con supervisión activa
- Gobernanza anticipatoria intensiva
- Piloto supervisado con criterios de cese
- Protocolos adaptativos con condiciones detonantes
- SupTech en tiempo real
- Evaluación de impacto (EIDF) previa
- Moratoria temporal si el riesgo es extremo

BAJO — GRADO DE IMPACTO — ALTO

Fuente: Elaboración propia. Los distintos tipos de impacto se pueden observar directamente en las personas y la protección de sus derechos individuales, pero también en el capital institucional como la democracia, la transparencia, la igualdad, la confianza, etc.

¿Por qué estos dos ejes y no otros?

— El nivel de riesgo determina la intensidad de la respuesta regulatoria: a mayor impacto potencial, mayor exigencia de salvaguardas.

— La certidumbre sobre los efectos determina si es viable regular directamente o si primero es necesario generar evidencia.

— La intersección de ambas dimensiones define no sólo qué instrumentos son adecuados, sino también su secuencia y combinación.

— Las demás dimensiones (tiempo, capacidad, alcance) modulan los instrumentos del cuadrante pero no cambian la familia de modelos base.

3.2 SECCIÓN 2 DE 7

Clasificación de riesgos

TABLA 3.1

Marco de clasificación de riesgos de sistemas de IA

INACEPTABLE NIVEL 4	Sistemas que vulneran derechos fundamentales de manera irreversible o que instrumentalizan a las personas en contra de sus propios intereses. Su daño no es accidental, es inherente al propósito del sistema. Ejemplos: sistemas de crédito social, predicción de delitos por individuos	MEDIDA REGULATORIA **Prohibición absoluta**
ALTO NIVEL 3	Sistemas que pueden tener impactos significativos en derechos fundamentales, seguridad pública o acceso a servicios esenciales. El daño no es inevitable, pero el riesgo es suficientemente alto como para exigir salvaguardas robustas antes, durante y posteriormente al despliegue. Ejemplos: gestión de infraestructura crítica, gestión de recursos humanos, servicios de salud, control migratorio, impartición de justicia, asignación de beneficios sociales	MEDIDA REGULATORIA **Evaluación de impacto, auditoría (sesgos, robustez), gobernanza de datos, transparencia, supervisión humana**
LIMITADO NIVEL 2	Sistemas que presentan riesgo principalmente de engaño, desinformación o falta de conocimiento informado por parte del usuario. El daño potencial no deriva de la decisión del sistema sino de la opacidad sobre su naturaleza artificial. Ejemplos: usos generales de IA generativa	MEDIDA REGULATORIA **Evaluación de impacto recomendable, transparencia, revelación a la persona usuaria, características de diseño (no adicción, sesgos, salvaguardas en la personalización, etc.)**
MÍNIMO NIVEL 1	Sistemas cuyos impactos son manejables en derechos, seguridad o autonomía de las personas. Corresponden a aplicaciones ampliamente establecidas y cuyo comportamiento erróneo tiene consecuencias menores y fácilmente reversibles. Ejemplos: filtros de spam, traductores, toma de minutas	MEDIDA REGULATORIA **Cumplimiento voluntario de buenas prácticas**

Fuente: Elaboración propia con base en el AI Act de la Unión Europea: https://artificialintelligenceact.eu/es/

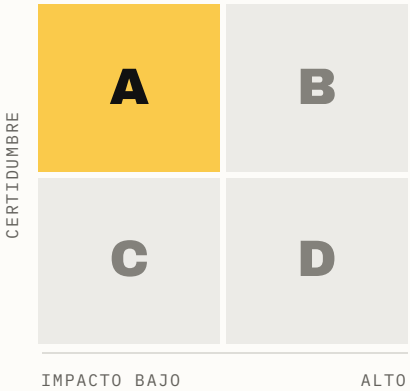
> Vale la pena recordar la diferencia entre riesgo e incertidumbre. Hablamos de riesgo cuando conocemos el conjunto de eventos posibles y sus respectivas probabilidades de ocurrencia; hablamos de incertidumbre cuando no conocemos esas probabilidades.

3.3 SECCIÓN 3 DE 7

3.3.1 · CUADRANTE A

Impacto bajo + efectos bien conocidos

Cuando el sistema de IA tiene bajo impacto directo y sus efectos se conocen bien gracias a precedentes o evidencia acumulada, no está justificado imponer instrumentos de supervisión intensiva. Cuando hablamos de efectos bien conocidos nos referimos también a que ese conocimiento conlleva capacidad de gestión y que se conoce la naturaleza de los mecanismos que los producen. El cuadrante A es el espacio de la exploración e innovación de bajo perfil regulatorio, donde el objetivo principal es fomentar buenas prácticas y actualizar el marco conforme cambien las condiciones, sin crear cargas de cumplimiento desproporcionadas.



Localizador. Ver Gráfica 3.1 para la matriz completa con sus instrumentos.

FICHA DE CUADRANTE	A
INSTRUMENTOS CENTRALES **Marco de incentivos con revisión periódica** **Marcos basados en principios** **Tecnología regulatoria voluntaria** **Normas voluntarias** **Cumplimiento voluntario de buenas prácticas**	INSTRUMENTOS COMPLEMENTARIOS **Cláusulas de expiración para normas que pueden quedar obsoletas** **Hubs de innovación para orientación temprana**
FOCO ESTRATÉGICO Generar adhesión voluntaria mediante procesos de coconstrucción y marcos de incentivos; penalizar incumplimientos graves sin sobrecargar el cumplimiento ordinario	SEÑAL DE ALERTA Si aparecen evidencias de daño imprevisto o si la escala de despliegue aumenta significativamente, reevaluar y decidir si el sistema debe migrar al cuadrante C o B

El riesgo real del cuadrante A no es el daño directo sino la **complacencia regulatoria**: la ausencia de supervisión puede llevar a que sistemas aparentemente inocuos acumulen efectos sistémicos no detectados a medida que se generalizan. No debe asumirse que un sistema permanecerá en el cuadrante A indefinidamente. Las cláusulas de revisión periódica son la salvaguarda fundamental contra ese riesgo.

EJEMPLO ILUSTRATIVO · 3.3.1

Cuadrante A

Marco de incentivos + cláusula de revisión

Innovación responsable
para un impacto positivo

1 Caso de uso

Una empresa de recursos humanos incorpora un corrector gramatical con IA a su plataforma interna de redacción de vacantes.

2 Impacto en las personas

Bajo

BAJO — ALTO

El sistema sugiere cambios de estilo, pero la decisión final siempre la toma una persona.

3 Efectos conocidos

Sí

INCIERTO — CONOCIDO

Hay precedentes claros de este tipo de herramientas y evidencia suficiente sobre sus limitaciones, sin riesgos nuevos o desconocidos en este contexto.

4 Respuesta recomendada

Adoptar el marco de buenas prácticas de la empresa alineado con ISO/IEC 42001, NIST u otro, con cláusula de revisión anual y registro de incidentes voluntario.

- Marco de buenas prácticas
- Cláusula de revisión anual
- Registro de incidentes voluntario

5 Criterio de migración

Si la empresa escala la herramienta para redactar perfiles, convocatorias o filtrar candidatos automáticamente, el sistema pasa al cuadrante B y requiere evaluación de impacto obligatoria y auditoría.

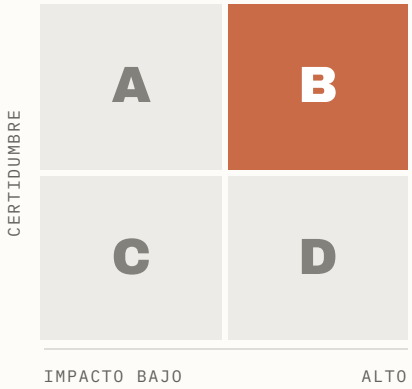
A → B **Evaluación de impacto obligatoria · Auditoría**

3.3 SECCIÓN 3 DE 7

3.3.2 · CUADRANTE B

Impacto alto + efectos bien conocidos

Cuando el sistema tiene alto impacto y sus efectos se conocen con suficiente precisión para regular de forma exigente, la regulación puede —y debe— ser directa y verificable. No se necesita experimentar: hay suficiente evidencia sobre qué puede salir mal y cómo prevenirlo. El cuadrante B es el espacio de la regulación más estructurada dentro del repertorio ágil, compatible con —y a menudo complementario— con la legislación y la normalización formal vinculante.



Localizador. Ver Gráfica 3.1 para la matriz completa con sus instrumentos.

FICHA DE CUADRANTE	B
INSTRUMENTOS CENTRALES **Corregulación y certificación/normalización** **Regulación basada en resultados** **Tecnología de supervisión y auditoría obligatorias** **Evaluación de impacto en derechos fundamentales (EIDF)**	INSTRUMENTOS COMPLEMENTARIOS **Puerto seguro con divulgación** **Protocolos adaptativos para actualización periódica** **Tecnología regulatoria de monitoreo continuo**
FOCO ESTRATÉGICO Garantizar cumplimiento verificable, rendición de cuentas clara y mecanismos de reclamación, reparación y remediación accesibles.	SEÑAL DE ALERTA Si la tecnología evoluciona de forma que los efectos dejan de ser bien conocidos (nueva versión del modelo, nuevo contexto de despliegue, nuevas dinámicas sociales), revisar si se debe migrar al cuadrante D.

La regulación basada en resultados es especialmente valiosa en el cuadrante B porque permite que las organizaciones innoven en los métodos de cumplimiento —aprovechando los avances técnicos más recientes y el conocimiento propio de las organizaciones reguladas— sin que el regulador tenga que actualizar constantemente los requisitos técnicos. Lo que se fija son los resultados; los medios son responsabilidad de la organización.

EJEMPLO ILUSTRATIVO · 3.3.2

Cuadrante B

Corregulación + regulación basada en resultados

Mayor impacto, efectos conocidos.
Reglas claras, supervisión activa.

1 Caso de uso

Un banco desarrolla un sistema de IA para evaluación crediticia que determina el acceso a préstamos personales.

2 Impacto en las personas

Alto

BAJO — ALTO

Afecta el acceso a servicios financieros esenciales, con efectos conocidos sobre grupos históricamente marginados.

3 Efectos conocidos

Sí

INCIERTO — CONOCIDO

Hay amplia evidencia sobre sesgos en modelos de evaluación y marcos sectoriales establecidos (Basilea, directrices FSB del BIS).[29]

4 Respuesta recomendada

Corregulación con el regulador financiero (modelo de gobernanza de riesgos), certificación externa periódica, tecnología de supervisión con acceso del regulador a métricas de equidad en tiempo real, y puerto seguro con divulgación para reportar desviaciones.

- Corregulación con el regulador
- Certificación externa periódica
- Tecnología de supervisión
- Puerto seguro con divulgación

5 Criterio de migración

Si se incorpora un nuevo sistema o se produce una nueva versión, cuyo comportamiento en la evaluación es incierto, revisar si migra temporalmente al cuadrante D para prueba supervisada antes de despliegue en producción.

B → **D** Prueba supervisada antes del despliegue

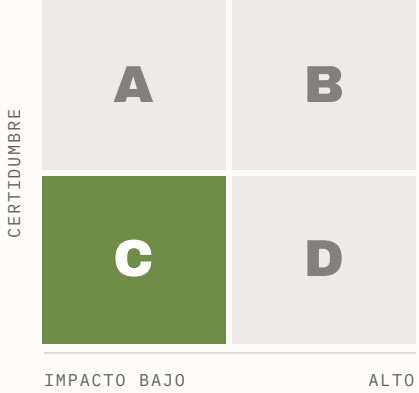
29 Marco de Basilea: https://www.bis.org/bcbs/basel3.htm · «Technology is neutral, governance is not: AI adoption in the banking sector» (2026): https://www.bankingsupervision.europa.eu/press/speeches/date/2026/html/ssm.sp260224~6c5b64a77a.en.html · «Regulating AI in the financial sector: recent developments and main challenges» (2024): https://www.bis.org/fsi/publ/insights63.htm

3.3 SECCIÓN 3 DE 7

3.3.3 · CUADRANTE C

Impacto bajo + efectos inciertos

El cuadrante C combina bajo impacto directo con incertidumbre significativa sobre los efectos. Es el espacio más propicio para la innovación con aprendizaje institucional: no hay urgencia protectora extrema como en el cuadrante D, ni certeza suficiente para regular directamente como en el cuadrante A. Lo que hay es una oportunidad para que el proceso regulatorio aprenda al mismo tiempo que el sistema se desarrolla, generando evidencia que eventualmente alimentará normas más sólidas para toda esa categoría de tecnología.



Localizador. Ver Gráfica 3.1 para la matriz completa con sus instrumentos.

FICHA DE CUADRANTE	C
INSTRUMENTOS CENTRALES **Piloto supervisado ligero** **Sandbox exploratorio** **Gobernanza anticipatoria** **Marco de incentivos experimental** **Hub de innovación regulatoria**	INSTRUMENTOS COMPLEMENTARIOS **Análisis de señales tempranas (horizon scanning)** **Marcos de principios exploratorios** **Cláusulas detonantes condicionadas a evidencia**
FOCO ESTRATÉGICO Aprender antes de formalizar; generar evidencia confiable con el menor costo posible para todos los actores.	SEÑAL DE ALERTA Si los efectos resultan más graves de lo anticipado, activar protocolo de migración al cuadrante D y aplicar salvaguardas más estrictas de forma inmediata.

El cuadrante C es el hábitat natural del sandbox exploratorio

En el cuadrante C, el sandbox no necesita la infraestructura jurídica completa que exige un sistema de alto riesgo. Lo que se busca es un espacio de observación estructurada con tres características básicas:

3.3 SECCIÓN 3 DE 7

3.3.3 · CUADRANTE C · CONTINUACIÓN

- Condiciones de entrada bajas. No requiere evaluación de impacto completa ni certificación previa. Basta con un acuerdo de participación que defina qué se observa, cómo se mide, durante cuánto tiempo y qué datos se comparten con la autoridad o instancia de supervisión.
- Supervisión asimétrica. El regulador no interviene activamente en el desarrollo del sistema, aunque diseña el espacio experimental, observa y registra. La carga principal recae en los responsables de ejecutar el proyecto, que deben reportar métricas en la forma y periodicidad acordadas ex post, no en tiempo real.
- Salida hacia el régimen ordinario. Al terminar el sandbox, la organización o los responsables del proyecto migran al régimen general de su sector si los efectos resultaron manejables, o escala hacia un sandbox más estricto si los hallazgos revelaron riesgos no anticipados.

El piloto supervisado ligero en el cuadrante C

El piloto supervisado es el equivalente del sandbox que no parte de una innovación privada, sino de la respuesta institucional de una autoridad reguladora que quiere poner a prueba un esquema de gobernanza o un mecanismo regulatorio nuevos, antes de generalizarlo. La organización se sujeta al marco normativo en condiciones normales de operación, en un ámbito geográfico, temporal o poblacional acotado, con un protocolo de monitoreo activo.

La diferencia clave con el sandbox es que en el piloto no hay exención normativa: la organización cumple todas las reglas aplicables desde el primer día. Lo que se acota es el alcance del despliegue, no el marco regulatorio. Para el cuadrante C, el piloto supervisado ligero tiene cuatro elementos mínimos:

1. Delimitación clara del alcance: se define con precisión quiénes serán los sujetos a las regulaciones del piloto, en qué contexto opera y durante cuánto tiempo. Esto permite que los hallazgos sean transferibles a la fase de expansión.
2. Métricas de éxito predefinidas: antes de lanzar, se acuerdan los indicadores que determinarán si el piloto fue exitoso, si debe continuar con ajustes o si debe suspenderse.
3. Canal de retroalimentación de usuarios: mecanismo sencillo para que los sujetos al piloto reporten experiencias inesperadas, facilidad de cumplimiento y claridad. Especialmente valioso cuando los efectos son inciertos.
4. Protocolo de escalamiento: si aparecen señales de daño o de riesgo más alto del anticipado, existe un procedimiento claro para pausar, ajustar o trasladar el sistema al cuadrante D.

3.3.3 · CUADRANTE C · CONTINUACIÓN

Combinación secuencial: sandbox seguido de piloto

El sandbox y el piloto no son excluyentes en el cuadrante C; suelen usarse en secuencia. El sandbox exploratorio opera primero, en un entorno deliberadamente controlado donde se despliega el sistema innovador, para identificar los efectos más inesperados antes de tocar usuarios reales en condiciones ordinarias. Una vez que hay evidencia preliminar que justifique el diseño de un marco de gobernanza específico, el piloto supervisado se utiliza para probarlo en condiciones reales con el aprendizaje del sandbox ya incorporado, pero en un ámbito acotado antes de pasar a su generalización.

La gobernanza anticipatoria actúa como capa de inteligencia transversal: las organizaciones, públicas o privadas, analizan señales tempranas para identificar riesgos o efectos no anticipados o bien riesgos conocidos en nuevos contextos. El hub de innovación regulatoria complementa la investigación con la exploración de nuevos mecanismos para gestionar los riesgos identificados.

El riesgo principal del cuadrante C

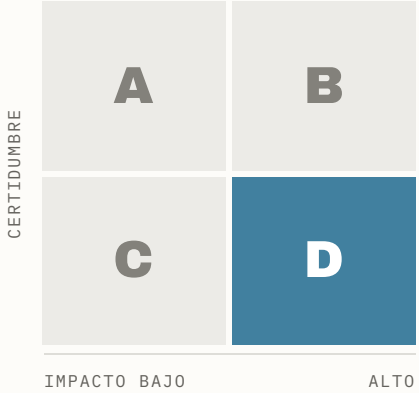
- No es el daño directo —que es bajo por definición— sino la inacción regulatoria por exceso de cautela.
- Organizaciones que, ante la incertidumbre, optan por no hacer nada y despliegan sin ningún protocolo de observación.
- Resultado: se pierde la oportunidad de aprender de forma ordenada y se generan precedentes normativos débiles.
- La regulación ágil invierte esta lógica: prefiere el ensayo controlado a la inacción, y convierte la incertidumbre en fuente de aprendizaje institucional.

3.3 SECCIÓN 3 DE 7

3.3.4 · CUADRANTE D

Impacto alto + efectos inciertos

El cuadrante D es la combinación de mayor exigencia: alto impacto potencial —que también se refiere a cuando no tenemos capacidad de controlar los efectos— y alta incertidumbre sobre cómo se materializarán esos impactos. Es el espacio donde la experimentación regulatoria no es una opción sino una necesidad, y donde las salvaguardas deben ser proporcionales a la gravedad de los daños posibles.



Localizador. Ver Gráfica 3.1 para la matriz completa con sus instrumentos.

FICHA DE CUADRANTE	D
INSTRUMENTOS CENTRALES **Sandbox estricto con supervisión activa y criterios de terminación** **Gobernanza anticipatoria intensiva** **Piloto supervisado con protocolos de intervención inmediata**	INSTRUMENTOS COMPLEMENTARIOS **Protocolos adaptativos con cláusulas detonantes automáticas** **Tecnología de supervisión en tiempo real** **EIDF previa al despliegue** **Moratoria temporal si el riesgo es extremo**
FOCO ESTRATÉGICO Proteger a las personas y capital institucional mientras se genera evidencia robusta bajo supervisión estricta; el aprendizaje no puede comprometer la seguridad.	SEÑAL DE ALERTA Si los efectos resultan bien conocidos y manejables, el sistema puede migrar al cuadrante B y aplicar instrumentos más estables; si resultan inaceptables, activar prohibición o moratoria.

La diferencia central entre los cuadrantes C y D no está en la incertidumbre —que es alta en ambos— sino en las consecuencias de equivocarse. En C, un error genera aprendizaje a bajo costo. En D, un error puede generar daños graves e irreversibles. Esa asimetría justifica exigencias radicalmente distintas a pesar de la similitud superficial en la dimensión de certidumbre. Usando como referencia la clasificación de riesgos de la Unión Europea, un caso extremo del cuadrante D llevaría a la prohibición llana del sistema.

EJEMPLO ILUSTRATIVO · 3.3.4

Cuadrante D

Sandbox estricto + gobernanza anticipatoria

Innovación con alta incertidumbre.
Máxima precaución.

1 Caso de uso

Una empresa tecnológica desarrolla un sistema de IA de apoyo a decisiones judiciales sobre medidas cautelares en materia penal.

2 Impacto en las personas

Alto

Afecta la libertad de personas imputadas; un error puede tener consecuencias irreversibles.

3 Efectos conocidos

Muy inciertos

No hay precedentes suficientes en el contexto jurídico local; los sesgos del modelo ante distintos grupos poblacionales son desconocidos o no se conocen los mecanismos para su prevención o corrección.

4 Respuesta recomendada

Sandbox estricto con participación del Poder Judicial, el regulador tecnológico y organizaciones de defensa de derechos humanos; evaluación de impacto previo obligatoria; criterios de terminación explícitos (si la tasa de error específico supera la humana o X%, suspensión inmediata); tecnología de supervisión con acceso del regulador en tiempo real; gobernanza anticipatoria para vigilar casos y evolución internacional.

- Participación de múltiples actores
- Evaluación de impacto previa obligatoria
- Criterios de terminación explícitos
- Supervisión en tiempo real
- Gobernanza anticipatoria

5 Condición indispensable

El sistema solo puede operar como herramienta de apoyo, sin automatización de la decisión judicial.

6 Criterio de migración

Si tras el sandbox la evidencia muestra comportamiento equitativo y predecible, el sistema puede migrar al cuadrante B y operar bajo corregulación con auditoría periódica, siempre como herramienta de apoyo y no de toma de decisión.

D → **B** Corregulación con auditoría periódica

3.4 SECCIÓN 4 DE 7

Las capas de ajuste transversales

Una vez identificado el cuadrante base del que partimos, cuatro capas de ajuste refinan la combinación de modelos según las condiciones específicas del proyecto. Las capas no reemplazan el cuadrante; modulan la intensidad, velocidad o alcance de los instrumentos seleccionados dentro de él.

CAPA	CUÁNDO SE ACTIVA	EFECTO SOBRE EL CUADRANTE BASE
Capacidad institucional limitada	La organización tiene equipo pequeño, recursos escasos o experiencia regulatoria incipiente.	Simplifica los instrumentos del cuadrante: sustituye sandboxes complejos por pilotos ligeros; sustituye corregulación por marco de incentivos basados en buenas prácticas; prioriza el hub de innovación como primer paso.
Alcance multijurisdiccional	El sistema opera o tiene efectos en más de un país o región con marcos regulatorios distintos.	Añade encima del cuadrante: cooperación multijurisdiccional, reconocimiento mutuo de certificaciones, protocolos de notificación multilateral de incidentes y sandbox cruzado. Este último permite que dos o más autoridades regulatorias (de distintos países o jurisdicciones) coordinen un mismo piloto o ejercicio de experimentación regulatoria, en lugar de que cada una corra su propio sandbox aislado.
Cambio probable de condiciones	Se reciben señales de cambios importantes en el entorno tecnológico, político o de mercado.	Añade encima del cuadrante: cláusulas detonantes, protocolos adaptativos con revisión automática, gobernanza anticipatoria como vigilancia continua.
Urgencia temporal	El tiempo disponible es menor al típicamente necesario para implementar los instrumentos del cuadrante.	Sustituye instrumentos por equivalentes rápidos del mismo cuadrante: marcos de incentivos basados en buenas prácticas en vez de corregulación; principios en vez de sandbox; hub de orientación en vez de piloto formal.

Las capas son acumulativas: un proyecto puede activar varias simultáneamente. Una organización con capacidad limitada que opera en múltiples jurisdicciones y en un entorno muy cambiante puede activar las cuatro capas al mismo tiempo sobre su cuadrante base. La combinación resultante no es más complicada que la suma de sus partes: cada capa añade un conjunto específico de instrumentos o ajusta los ya seleccionados.

3.5 SECCIÓN 5 DE 7

Nota sobre las combinaciones de modelos

La gobernanza y la regulación ágil rara vez se implementan con un solo instrumento. Lo habitual es que un marco de gobernanza requiera uno o más instrumentos regulatorios, además de instancias de decisión, mecanismos de colaboración, de información y definición de responsabilidades. Así, un proyecto puede activar una combinación de dos modelos de regulación que operan en distintas capas temporales y funcionales simultáneamente.

PERFIL	COMBINACIÓN RECOMENDADA
Combinación más frecuente en A	Marco de incentivos como base normativa + hub de innovación para orientación temprana + cláusulas de revisión periódica como salvaguarda adaptativa.
Combinación más frecuente en B	Corregulación para el diseño de estándares + regulación basada en resultados como exigencia de cumplimiento + SupTech para la supervisión + puerto seguro para incentivar el cumplimiento y la divulgación de incidentes.
Combinación más frecuente en C	Hub de innovación como puerta de entrada + sandbox exploratorio para la fase inicial + piloto supervisado para la fase de validación + gobernanza anticipatoria como vigilancia continua.
Combinación más frecuente en D	Sandbox estricto como condición de acceso al mercado + protocolos adaptativos con cláusulas detonantes + SupTech en tiempo real + evaluación de impacto antes del despliegue. Algunos casos extremos en este cuadrante pueden justificar la suspensión del uso o desarrollo de un sistema, al menos para ciertos casos.
Con capa multijurisdiccional	Cualquier combinación base + cooperación multijurisdiccional + reconocimiento mutuo de certificaciones + protocolos de notificación cruzada de incidentes.

Las combinaciones no son rígidas. Lo que este marco proporciona es una lógica de selección —no un algoritmo—: orienta la decisión señalando qué familias de instrumentos son coherentes con el perfil del proyecto y cuáles añadirían cargas innecesarias o serían insuficientes. La experiencia sectorial, el diálogo con los actores afectados y el criterio de quien toma la decisión son siempre componentes indispensables del resultado final.

3.6 SECCIÓN 6 DE 7

Metodología genérica para justificar y poner en marcha modelos de regulación ágil

A lo largo del manual se recomienda la aplicación de varios instrumentos de regulación ágil para construir modelos de gobernanza basados en principios éticos e incluyentes, sin embargo, no se ha hablado de cuáles son los pasos a seguir para planear, diseñar y aplicar esos instrumentos. En esta sección se explica una metodología compacta y replicable que sigue una secuencia de siete fases para diseñar y poner en marcha un modelo de regulación de forma congruente con los principios desarrollados en este manual.

Esta es una metodología genérica que sigue siete fases. En cada fase se describe su objetivo, sus actividades clave, su entregable, el agente responsable, el umbral o disparador que activa una decisión, y la consecuencia asociada. Como en el resto del manual, los plazos funcionan como límites máximos, no como estándares, y la activación real proviene de umbrales cuantificables, no de fechas calendarizadas.

Fase 1 JUSTIFICACIÓN

OBJETIVO

Determinar cuál, o cuáles, son los instrumentos de regulación ágil adecuados para el caso.

ACTIVIDADES CLAVE

- Verificar que existe incertidumbre regulatoria genuina o un limbo normativo.
- Confirmar que el sistema de IA objetivo no opera en un marco adecuado para responder a los cambios rápidos del entorno o de la tecnología.
- Determinar el potencial de escalamiento del caso aplicando el método V-R-I.
- Comparar instrumentos y seleccionar la combinación más conveniente.

ENTREGABLE

Nota de justificación (1-2 páginas) con el criterio V-R-I aplicado y la combinación de instrumentos seleccionada.

AGENTE RESPONSABLE

Unidad regulatoria, administrativa o de innovación que propone el o los instrumentos seleccionados.

UMBRAL / DISPARADOR

Si no existe incertidumbre regulatoria genuina ni potencial de escalamiento, el proceso termina aquí.

CONSECUENCIA

No se activa un instrumento ágil para lo que ya tiene una respuesta regulatoria funcional, ni un instrumento más intrusivo de lo necesario.

FICHAS DE FASE · CONTINUACIÓN

Fase 2 PLANEACIÓN

OBJETIVO

Definir el alcance, los límites y las reglas de operación del instrumento antes de ponerlo en marcha.

ACTIVIDADES CLAVE

- Identificar a los actores, sistemas de IA, sectores y población participante o afectada.
- Definir criterios de elegibilidad y un proceso de selección transparente.
- Establecer qué se flexibiliza, suspende o acompaña temporalmente, y qué permanece vigente sin excepción.
- Fijar la duración máxima del ejercicio y las salvaguardas para las personas usuarias o afectadas.

ENTREGABLE

Términos de referencia del instrumento (participantes, elegibilidad, alcances, salvaguardas, duración).

AGENTE RESPONSABLE

Regulador sectorial o unidad administrativa correspondiente, con validación de la unidad jurídica.

UMBRAL / DISPARADOR

Activación formal del instrumento incluyendo explícitamente las salvaguardas no negociables.

CONSECUENCIA

Sin salvaguardas explícitas, el instrumento no se activa.

Fase 3 DISEÑO

OBJETIVO

Construir el protocolo operativo del instrumento y las métricas que se van a observar.

ACTIVIDADES CLAVE

- Definir las hipótesis o resultados esperados del ejercicio.
- Establecer los indicadores de monitoreo y sus umbrales de alerta; cada indicador debe disparar una acción específica.
- Diseñar el mecanismo de reporte (frecuencia, formato, responsable) entre quienes ejecutan el instrumento y el agente regulador.
- Definir el comité o mesa de evaluación responsable de la gobernanza del ejercicio, bajo criterios de inclusión y diversidad.

ENTREGABLE

Protocolo operativo con hipótesis, indicadores, umbrales y mecanismo de reporte.

AGENTE RESPONSABLE

Equipo técnico conjunto (agente regulador + actores involucrados).

UMBRAL / DISPARADOR

El protocolo no se aprueba si algún indicador carece de umbral de alerta cuantificable.

CONSECUENCIA

Indicador sin umbral cuantificable → protocolo regresa a rediseño.

FICHAS DE FASE · CONTINUACIÓN

Fase 4 DESPLIEGUE / EJECUCIÓN

OBJETIVO	ACTIVIDADES CLAVE
Poner en marcha el instrumento con las condiciones acordadas.	· Activar el instrumento con los actores seleccionados, en los términos acordados en la planeación. · Mantener un canal de comunicación abierto para incidentes no previstos en el protocolo. · Registrar evidencia de forma continua, incluso si el ejercicio se interrumpe antes de tiempo.
ENTREGABLE	**AGENTE RESPONSABLE**
Bitácora de ejecución y registro de incidentes.	Actores involucrados (generan evidencia) y agente regulador (supervisa).
UMBRAL / DISPARADOR	**CONSECUENCIA**
Cualquier incidente que cruce el umbral de riesgo definido en el protocolo.	Cruce de umbral → pausa inmediata del instrumento, sin esperar al reporte periódico.

Fase 5 MONITOREO

OBJETIVO	ACTIVIDADES CLAVE
Verificar en tiempo real que el ejercicio se mantenga dentro de los parámetros definidos.	· Revisar los indicadores contra los umbrales fijados en diseño, con la periodicidad acordada (límite máximo, no estándar: si un umbral se cruza antes, se actúa antes). · Documentar desviaciones entre lo observado y la hipótesis original. · Mantener trazabilidad de las decisiones tomadas durante el ejercicio: quién decidió qué, con base en qué evidencia.
ENTREGABLE	**AGENTE RESPONSABLE**
Reportes de monitoreo periódicos y registro de desviaciones.	Comité de evaluación definido en la fase de diseño.
UMBRAL / DISPARADOR	**CONSECUENCIA**
Cruce de umbral de riesgo, o cruce de umbral positivo de desempeño.	Umbral de riesgo → ajuste del protocolo o terminación anticipada. Umbral positivo → posible ampliación del alcance del instrumento.

FICHAS DE FASE · CONTINUACIÓN

Fase 6 RETROALIMENTACIÓN

OBJETIVO

Convertir la evidencia generada en aprendizaje aplicable, dentro y fuera del ejercicio.

ACTIVIDADES CLAVE

- Sistematizar los resultados frente a las hipótesis planteadas: ¿se confirmaron, se refutaron, quedaron sin resolver?
- Identificar si la evidencia sugiere escalar a un instrumento más permanente, ajustar el instrumento actual, o descartarlo.
- Consultar a las partes afectadas, no sólo a los actores involucrados, sobre su experiencia durante el ejercicio, cuando el caso lo permita.

ENTREGABLE

Informe de resultados con recomendación de ruta: escalar, ajustar/extender, o descartar.

AGENTE RESPONSABLE

Comité de evaluación, con insumos de las partes afectadas.

UMBRAL / DISPARADOR

Evidencia suficiente sobre las hipótesis originales.

CONSECUENCIA

Sin evidencia suficiente, no procede recomendar una ruta de escalamiento; la ausencia de incidentes no es evidencia suficiente por sí sola.

Fase 7 CIERRE

OBJETIVO

Formalizar el fin del ejercicio y su destino regulatorio.

ACTIVIDADES CLAVE

- Emitir la resolución de cierre: terminación de la intervención temporal y, si aplica, transición a un instrumento permanente o a un nuevo ciclo.
- Publicar un resumen de resultados, salvo restricciones legítimas de confidencialidad, como ejercicio de transparencia regulatoria.
- Incorporar los aprendizajes al inventario de evidencia del regulador, para este caso y para otros instrumentos del catálogo ágil.

ENTREGABLE

Resolución de cierre y resumen público de resultados.

AGENTE RESPONSABLE

Comité de evaluación.

UMBRAL / DISPARADOR

Emisión formal de la resolución de cierre.

CONSECUENCIA

Sin resolución de cierre formal, el ejercicio no puede considerarse concluido para efectos de responsabilidad regulatoria.

DIAGRAMA 3.2

Actividades y puntos de control por fase

ACTIVIDAD · DECISIÓN · PASA A LA FASE INDICADA · FIN DEL PROCESO

Fase 1 — JUSTIFICACIÓN

1. Verificar incertidumbre regulatoria genuina o limbo normativo
2. Confirmar que el marco vigente no responde a cambios del entorno o la tecnología (o que el diseño institucional no está listo)
3. Determinar el potencial de escalamiento (criterio V-R-I)
4. Comparar instrumentos y elegir la combinación más conveniente (matriz impacto-certidumbre)

¿Existe incertidumbre regulatoria genuina o potencial de escalamiento?

- SÍ → a Fase 2
- NO → termina aquí

Fase 2 — PLANEACIÓN

1. Identificar actores, sistemas y población participante
2. Definir elegibilidad y proceso de selección
3. Establecer límites de la intervención regulatoria
4. Fijar duración máxima y salvaguardas

¿Términos de referencia incluyen salvaguardas explícitas?

- SÍ → a Fase 3
- NO → a definir salvaguardas

Fase 3 — DISEÑO

1. Definir hipótesis o resultados esperados
2. Establecer indicadores y umbrales de alerta
3. Diseñar mecanismo de reporte
4. Definir comité de evaluación con criterios de inclusión y diversidad

¿Todos los indicadores tienen umbral cuantificable?

- SÍ → a Fase 4
- NO → a rediseño

Fase 4 — DESPLIEGUE / EJECUCIÓN

1. Activar el instrumento con actores seleccionados
2. Mantener canal de comunicación para incidentes
3. Registrar evidencia de forma continua

¿Algún incidente cruza el umbral de riesgo?

- SÍ → pausa, a Fase 7
- NO → a Fase 5

Fase 5 — MONITOREO

1. Revisar indicadores contra umbrales fijados
2. Documentar desviaciones de la hipótesis
3. Mantener trazabilidad de decisiones

¿Indicadores fuera de rango?

- SÍ → a Fase 3 o a Fase 7
- NO → a Fase 6

Fase 6 — RETROALIMENTACIÓN

1. Sistematizar resultados frente a hipótesis
2. Identificar ruta: escalar, ajustar o descartar
3. Consultar a las partes afectadas

¿Evidencia suficiente sobre las hipótesis originales?

- SÍ → a Fase 7
- NO → a Fase 4 (extiende)

Fase 7 — CIERRE

1. Emitir resolución de cierre
2. Publicar resumen de resultados
3. Incorporar aprendizajes al inventario de evidencia

¿Se emitió la resolución de cierre?

- SÍ → fin del proceso
- NO → no se considera concluido

Fuente: Elaboración propia. Una pregunta de sí/no cierra cada fase y determina si el proceso avanza, regresa a un paso anterior, salta a otra fase o concluye.

3.7 SECCIÓN 7 DE 7

Metodología para diseñar y ejecutar un sandbox regulatorio

Hemos escogido el caso particular de un sandbox para ilustrar de manera específica los pasos del método propuesto. Sin embargo, como se menciona desde el inicio de este manual, para poder desarrollar con detalle un sistema de regulación ágil es necesario consultar y aplicar la metodología descrita pormenorizadamente en los tres primeros anexos de este documento.

Pasando a la metodología de esta sección, ésta consiste de siete fases, en cada una de las cuales se describe su objetivo, sus actividades clave, su entregable y, siguiendo el formato de acción establecido en el manual, el agente responsable, el umbral o disparador que activa una decisión, y las consecuencias asociadas. En línea con el espíritu ágil que anima este documento, ninguna fase depende de plazos fijos como único criterio de avance: los plazos funcionan como límites máximos, no como estándares, y la activación real proviene de umbrales cuantificables.

La metodología se describe tomando como ejemplo el sandbox, pero la secuencia de actividades, preguntas y análisis es aplicable para los otros 11 modelos regulatorios incluidos en la Tabla 2.1.

3.7 SECCIÓN 7 DE 7

Fase 1 JUSTIFICACIÓN

OBJETIVO	ACTIVIDADES CLAVE
Determinar si el sandbox es el instrumento adecuado para el caso, y no una elección por defecto.	• Verificar que existe incertidumbre regulatoria genuina, un limbo normativo o un marco de pensamiento lineal inoperante. • Confirmar que el caso tiene potencial de escalamiento. Aplicando el método V-R-I (vulnerabilidad-riesgo-impacto), un sandbox se justifica típicamente cuando la vulnerabilidad o el impacto son altos pero el riesgo aún no está caracterizado con evidencia suficiente como para normar o regular en firme. • Evaluar antes si el caso se puede resolver con alternativas más ligeras como elaborar una guía interpretativa, autoimposición de mejores prácticas u organización de una mesa de diálogo.
ENTREGABLE	**AGENTE RESPONSABLE**
Nota de justificación (1-2 páginas) con el criterio V-R-I aplicado y la alternativa descartada.	Unidad regulatoria, administrativa o de innovación que propone el sandbox.
UMBRAL / DISPARADOR	**CONSECUENCIA**
Si no existe incertidumbre regulatoria genuina, un limbo normativo o un marco de pensamiento lineal inoperante, el proceso termina aquí.	No se abre un sandbox para lo que ya tiene respuesta regulatoria clara.

FICHAS DE FASE · CONTINUACIÓN

Fase 2 PLANEACIÓN

OBJETIVO	ACTIVIDADES CLAVE
Definir el alcance, los límites y las reglas del juego antes de aceptar participantes.	• Identificar qué actores deben participar, qué sistemas de IA, qué sectores y la muestra de la población afectada que entra al sandbox. • Definir criterios de elegibilidad para participantes (empresas, instituciones públicas, universidades) y un proceso de selección transparente. • Establecer los límites de la excepción regulatoria: qué normas se suspenden temporalmente y cuáles permanecen vigentes sin excepción (protección de datos y derechos fundamentales nunca se relajan). • Fijar la duración máxima del sandbox y las condiciones de salvaguarda para los usuarios o afectados durante la prueba (consentimiento informado, posibilidad de reversión, compensación ante daño).
ENTREGABLE	**AGENTE RESPONSABLE**
Términos de referencia del sandbox (participantes, elegibilidad, alcances, salvaguardas, duración).	Regulador sectorial, unidad administrativa de la empresa, institución educativa u otra, todas con validación de la unidad jurídica.
UMBRAL / DISPARADOR	**CONSECUENCIA**
Publicación de convocatoria incluyendo explícitamente en los términos de referencia las salvaguardas no negociables.	Sin salvaguardas explícitas, la convocatoria no se publica.

FICHAS DE FASE · CONTINUACIÓN

Fase 3 DISEÑO

OBJETIVO

Construir el protocolo técnico de prueba y las métricas que se van a observar.

ACTIVIDADES CLAVE

- Definir las hipótesis a probar (por ejemplo: "el sistema reduce el sesgo de género en x puntos porcentuales; no se incrementa el riesgo de incumplimiento").
- Establecer los indicadores de monitoreo y sus umbrales de alerta; cada indicador debe tener un valor que detone una acción específica.
- Diseñar el mecanismo de reporte (frecuencia, formato, responsable) entre las y los participantes y el agente regulador.
- Definir el comité o mesa de evaluación que tomará decisiones durante el sandbox, y que estará a cargo de la gobernanza del ejercicio.

ENTREGABLE

Protocolo de prueba con hipótesis, indicadores, umbrales y mecanismo de reporte.

AGENTE RESPONSABLE

Equipo técnico conjunto (agente regulador + participantes).

UMBRAL / DISPARADOR

El protocolo no se aprueba si algún indicador carece de umbral de alerta cuantificable.

CONSECUENCIA

Indicador sin umbral cuantificable → protocolo regresa a rediseño.

Fase 4 DESPLIEGUE / EJECUCIÓN

OBJETIVO

Poner en marcha la prueba en condiciones controladas y reales.

ACTIVIDADES CLAVE

- Activar el sandbox con el grupo de participantes seleccionado, bajo el perímetro regulatorio acordado en planeación.
- Mantener un canal de comunicación abierto entre las y los participantes y el agente regulador para incidentes previstos o imprevistos en el protocolo.
- Registrar evidencia de forma continua para sostener el aprendizaje, incluso si el sandbox se interrumpe antes de tiempo.

ENTREGABLE

Bitácora de ejecución y registro de incidentes.

AGENTE RESPONSABLE

Participantes (generan evidencia) y agente regulador (supervisa).

UMBRAL / DISPARADOR

Cualquier incidente que cruce el umbral de riesgo definido en el protocolo.

CONSECUENCIA

Cruce de umbral → pausa inmediata del sandbox, sin esperar al reporte periódico.

FICHAS DE FASE · CONTINUACIÓN

Fase 5 MONITOREO

OBJETIVO

Verificar en tiempo real que la prueba se mantenga dentro de los parámetros definidos.

ACTIVIDADES CLAVE

- Revisar los indicadores contra los umbrales fijados en la fase de diseño, con la periodicidad acordada. La periodicidad es un límite máximo, no un estándar: si un umbral se cruza antes, se actúa antes.
- Documentar desviaciones entre lo observado y la hipótesis original.
- Mantener trazabilidad de las decisiones tomadas durante la prueba: quién decidió qué, con base en qué evidencia.

ENTREGABLE

Reportes de monitoreo periódicos y registro de desviaciones.

AGENTE RESPONSABLE

Comité de evaluación definido en la fase de diseño.

UMBRAL / DISPARADOR

Cruce de umbral de riesgo, o cruce de umbral positivo de desempeño.

CONSECUENCIA

Umbral de riesgo → ajuste del protocolo o terminación anticipada. Umbral positivo → posible ampliación del alcance del sandbox.

Fase 6 RETROALIMENTACIÓN

OBJETIVO

Convertir la evidencia generada en aprendizaje aplicable, dentro y fuera del sandbox.

ACTIVIDADES CLAVE

- Sistematizar los resultados frente a las hipótesis planteadas: ¿se confirmaron, se refutaron, quedaron sin resolver?
- Identificar si la evidencia sugiere que el caso puede pasar a regulación permanente, si no requiere más que criterios de interpretación de la regulación existente, requiere un sandbox adicional con condiciones ajustadas, o no debe escalar.
- Consultar a las partes afectadas, no sólo a las y los participantes, sobre su experiencia durante la prueba, cuando el caso lo permita.

ENTREGABLE

Informe de resultados con recomendación de ruta: regular, emitir criterios, extender el sandbox, o descartar.

AGENTE RESPONSABLE

Comité de evaluación, con insumos de las partes afectadas.

UMBRAL / DISPARADOR

Evidencia suficiente sobre las hipótesis originales.

CONSECUENCIA

Sin evidencia suficiente de afectaciones, no procede recomendar regulación permanente; la ausencia de incidentes no es evidencia suficiente por sí sola.

FICHAS DE FASE · CONTINUACIÓN

Fase 7 CIERRE

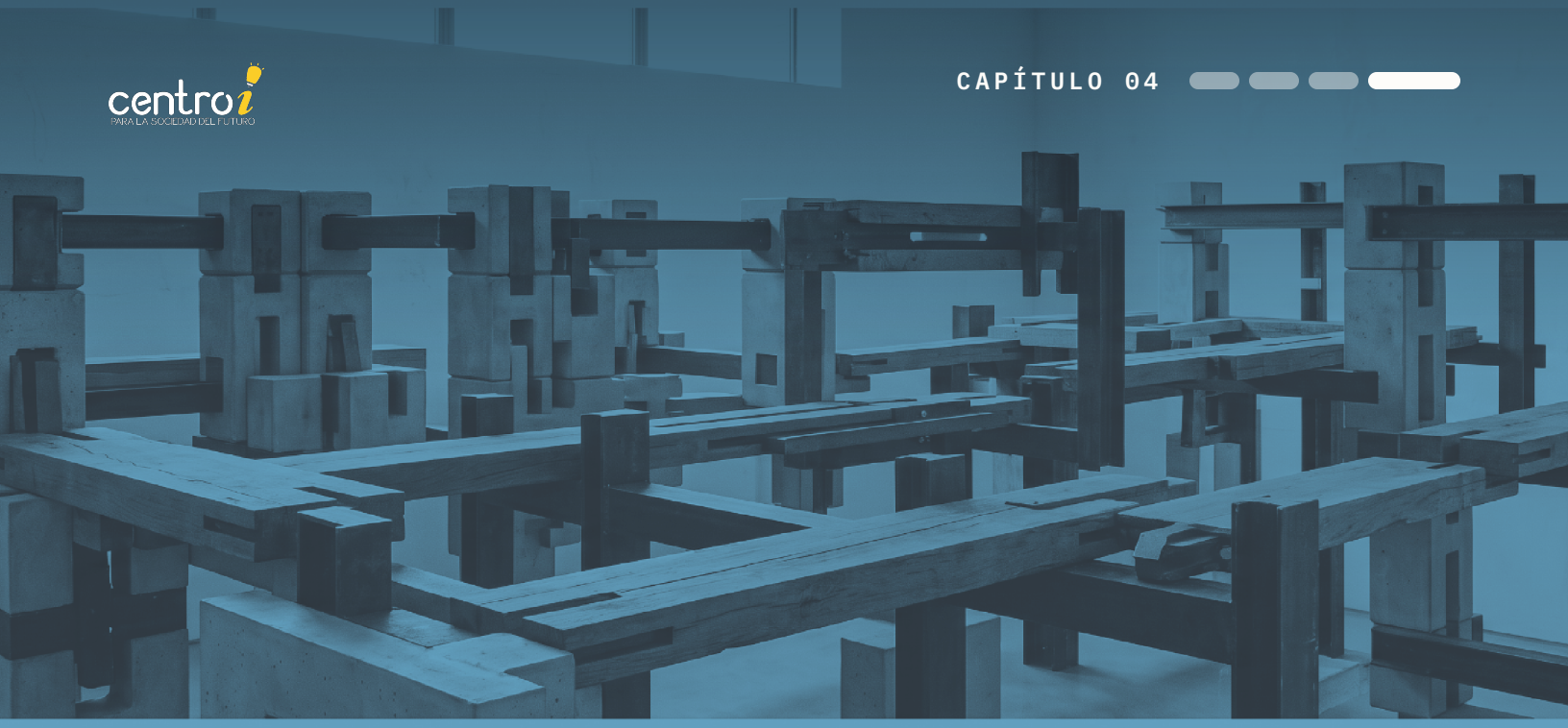
OBJETIVO	ACTIVIDADES CLAVE
Formalizar el fin del sandbox y su destino regulatorio.	· Emitir la resolución de cierre: terminación del espacio experimental regulatorio y, si es aplicable, transición a marco regulatorio permanente o a un nuevo ciclo de prueba. · Publicar un resumen de resultados, salvo restricciones legítimas de confidencialidad comercial, como ejercicio de transparencia regulatoria. · Incorporar los aprendizajes del sandbox al inventario de evidencia del regulador, para casos futuros similares.
ENTREGABLE	**AGENTE RESPONSABLE**
Resolución de cierre y resumen público de resultados.	Comité de evaluación.
UMBRAL / DISPARADOR	**CONSECUENCIA**
Emisión formal de la resolución de cierre.	Sin resolución de cierre formal, el sandbox no puede considerarse concluido para efectos de responsabilidad regulatoria.

El siguiente diagrama despliega la secuencia completa de actividades dentro de cada una de las siete fases, junto con el punto de decisión que cierra cada fase: una pregunta de sí/no que determina si el proceso avanza a la siguiente fase, regresa a un paso anterior, salta a otra fase, o concluye.

DIAGRAMA 3.3

Actividades y puntos de control por fase — sandbox regulatorio

Fuente: Elaboración propia.



04

Conclusiones y recomendaciones

CONCLUSIONES

Este documento ha sido elaborado con el propósito de contribuir al alivio de las tensiones generadas por la inédita aceleración con la que la inteligencia artificial está evolucionando y permeando todos los ámbitos de la vida. En particular, nos hemos enfocado en las tensiones que surgen por el hecho de que los marcos vigentes de gobernanza y regulación tecnológica que seguimos aplicando fueron concebidos para contextos y situaciones que simplemente ya no existen. Esa brecha estructural no es un problema coyuntural que se resuelva actualizando leyes o reglamentos, se trata más bien de un fenómeno nuevo provocado por la aceleración que ahora define los cambios en el desarrollo de la IA. Así, no basta con pretender actualizar leyes, regulaciones, guías o lineamientos con mayor regularidad, sino de modificar el enfoque para hacer normas dentro de un marco de incertidumbre inescapable, que nos ayude a aprender y mejorar de forma continua.

El documento analiza un catálogo de modelos de regulación ágil que puede ser de utilidad, pero su valor agregado es de orden metodológico: se parte de una concepción ecosistémica para desarrollar una herramienta práctica, reproducible y conceptualmente sólida para entender cómo funcionan los distintos modelos de regulación ágil que se han desarrollado, sus características, requerimientos, ventajas y desventajas. Una vez entendidos, entonces seleccionar, diseñar y aplicar estos instrumentos en función de las características específicas del problema de gobernanza que se enfrenta. Como se ha explicado en el texto, ese enfoque permite que una entidad pública o privada, independientemente de su escala y sector, pueda aplicar un proceso de análisis que va desde la construcción del mapa sistémico del caso hasta la selección de la combinación de instrumentos más adecuada, el diseño de su implementación con criterios de agilidad reales (umbrales automáticos verificables, ciclos de monitoreo continuo, criterios de cierre) y mecanismos de retroalimentación que convierten la experiencia acumulada en conocimiento regulatorio.

El enfoque sistémico permite identificar los puntos de incidencia o palanca en donde las medidas de política generan la propagación sistémica deseada, es decir, aquellos nodos en donde se originan las cadenas causales sobre las que se quiere incidir y que la regulación lineal no alcanza a identificar ni conectar. La matriz impacto-certidumbre, la clasificación de riesgos en cuadrantes y el diseño de estrategias de intervención por cadena causal son herramientas derivadas de esa visión sistémica, y su aplicación produce recomendaciones regulatorias cualitativamente diferentes de las que produce el análisis costo-beneficio convencional.

CONCLUSIONES · CONTINUACIÓN

La concepción ecosistémica de la gobernanza que se desarrolla en el manual tiene además una implicación práctica relevante para el contexto mexicano y latinoamericano: permite que el rol regulador no esté restringido a las instancias gubernamentales. Una empresa puede gobernar sus propios sistemas de IA con la misma arquitectura metodológica que usa un regulador sectorial; una universidad puede diseñar protocolos de evaluación para sus proyectos de IA con la misma lógica con la que una agencia pública diseña un sandbox; una organización de la sociedad civil puede construir mecanismos de rendición de cuentas usando las mismas herramientas detonantes y de monitoreo que un supervisor financiero. Esta flexibilidad de escala y de ámbito amplía significativamente el alcance potencial de la metodología propuesta.

Este documento se acompaña de un conjunto de anexos con materiales para facilitar la aplicación de la metodología propuesta. Uno de ellos desarrolla dos ejemplos de aplicación: el sistema judicial de medidas cautelares con sesgo socioeconómico y el motor de evaluación crediticia con sesgo de género. Estas son guías metodológicas completas que recorren todo el proceso: desde la construcción del mapa sistémico y la identificación de los nodos y cadenas causales relevantes, hasta el análisis de ámbitos funcionales, la estimación numérica de riesgos e impactos mediante el método V-R-I (vulnerabilidad-riesgo-impacto), y el diseño de estrategias sistémicas de intervención con acciones congruentes con los principios de la regulación ágil como ciclos de monitoreo automatizado, umbrales detonantes, criterios para detener el despliegue, sandboxes con métricas de evaluación predefinidas, nudges[30] conductuales, y mecanismos de remedio accesibles. Estos casos fueron desarrollados precisamente porque pertenecen a dos de los sectores con mayor exposición a riesgos de alto impacto en México, y porque ilustran la aplicabilidad del enfoque tanto en el ámbito público como en el privado.

Las fichas técnicas que acompañan a este documento y que resumen las características, cualidades y aplicabilidad de cada herramienta de regulación ágil, serán útiles en todas estas tareas.

30 Un nudge o empujón conductual es la creación de condiciones que lleven a las personas o agentes económicos en general, a tomar decisiones en cierto sentido deseado. Este concepto, entre otras aportaciones, le valió a Richard Thaler el Premio Nóbel en Economía en 2017.

RECOMENDACIONES Y AGENDA

La agenda de trabajos pendientes para avanzar en la regulación ágil de la IA está sobrecargada. Es urgente iniciar con la construcción de capacidades institucionales específicas para la regulación ágil de la IA en México. Ningún instrumento del catálogo, ni el sandbox, ni el piloto supervisado, ni los mecanismos de corregulación, podrá funcionar si no contamos con equipos técnicos con formación multidisciplinaria; con sistemas de información que permitan el monitoreo continuo que los umbrales detonantes requieren; y con culturas organizacionales capaces de aprender de la experiencia y de ajustar los instrumentos cuando la evidencia lo demanda.

La globalización de las empresas y mercados, al tiempo que se resienten los costos de la fragmentación regulatoria, genera la presión de armonizar los marcos regulatorios y de gobernanza para cubrir aquellos más exigentes, como es el caso de las normas europeas. Así, se está formando una demanda real de metodologías de implementación en el sector privado: empresas con operaciones transfronterizas, proveedores de tecnología que venden a clientes europeos, e instituciones financieras expuestas a la regulación de la Unión Europea necesitan hoy, no en cinco años, marcos de gobernanza documentados y verificables. Esa demanda privada puede financiar la construcción de capacidades que también beneficien al sector público.

En el ámbito de la educación los pendientes también se acumulan aceleradamente. Entre otras muchas vías, el sistema educativo puede desarrollar protocolos institucionales de gobernanza de IA para proyectos de investigación y para el uso de sistemas de IA en procesos académicos: evaluación de estudiantes, formación de personal académico y administrativo, asignación de recursos aplicando la misma lógica de mapas sistémicos y umbrales que se propone para otros sectores, establecer laboratorios de política educativa que funcionen como espacios de simulación de instrumentos ágiles en colaboración con autoridades educativas y empresas privadas, generando evidencia sobre qué funciona en el contexto institucional local antes de que esa evidencia tenga que generarse directamente en el sistema real, así como producir documentación de casos de aplicación de la metodología en sectores y contextos mexicanos y latinoamericanos específicos, que sirvan como material de enseñanza y como referencia para otros actores de la región.

RECOMENDACIONES Y AGENDA · CONTINUACIÓN

El manual también se pensó para que las empresas privadas encuentren en la guía de selección de instrumentos una herramienta útil para el diseño o actualización de sus marcos de gobernanza de IA, ya sea por ejemplo, para anticipar los requisitos de certificación y gobernanza documentada que implica la armonización otros marcos, como las leyes europeas, utilizando la ISO/IEC 42001 (u otras futuras) como estándar de referencia para los proveedores de sistemas de IA y diseñando cláusulas contractuales con detonantes automáticos de recalibración y auditoría; para la formación de equipos internos con competencias en análisis sistémico y diseño de instrumentos ágiles de regulación; para diseñar mecanismos de rendición de cuentas internos (tableros de equidad, reportes de auditoría con actualización continua, protocolos de operación restringida) que sean operativos antes de que el regulador los exija, aprovechando la ventana de tiempo que la corregulación voluntaria todavía ofrece como diferenciación competitiva; o para participar activamente en esquemas de corregulación sectorial donde los estándares los diseñe el sector pero con respaldo legal del regulador, que es el instrumento más eficiente para contextos donde la diversidad técnica de los sistemas hace difícil que el regulador fije estándares únicos sin perder pertinencia.

Podemos tomar como un último ejemplo el de los fondos de inversión, en donde resulta de gran utilidad la incorporación de un sistema de gobernanza de la IA documentado, verificable y congruente con los principios de regulación ágil de este manual como criterio de diligencia ante las evaluaciones de riesgo regulatorio de portafolios con exposición a sistemas de IA de alto impacto. Asimismo, se vislumbran ventajas interesantes de utilizar los indicadores de justicia algorítmica y los detonantes automáticos como señales tempranas de riesgo reputacional y regulatorio, distinguiendo claramente entre empresas que gobiernan sus sistemas de IA con criterios de proporcionalidad dinámica y las que operan con supuestos estáticos que la evidencia puede invalidar rápidamente, o de promover activamente entre las empresas del portafolio la adopción anticipada de estándares de gobernanza verificables antes de que la presión regulatoria los convierta en requisito obligatorio, capturando el valor de la diferenciación que esa anticipación genera en mercados donde la regulación de IA está endureciéndose cada vez más.

La agenda del Centro-i

En el Centro-i enfrentamos también una agenda sobrecargada. Entendemos la importancia de desarrollar capacidades técnicas internas para concebir los mapas sistémicos y las métricas de equidad que las instituciones públicas y privadas necesitan en el marco de sus esquemas de gobernanza de IA, así como desarrollar y aplicar metodologías de propagación sistémica que permitan estimar con mayor precisión los efectos económicos y sociales de la adaptación generalizada de la IA.

También tendremos que participar activamente en los procesos de diseño de códigos de conducta sectoriales y esquemas de corregulación, garantizando por ejemplo, que las métricas de equidad y los umbrales que se fijen reflejen las condiciones reales de las poblaciones afectadas y no sólo los parámetros técnicos que resultan cómodos para los operadores del sistema; e intentaremos documentar y sistematizar los casos en que los sistemas de IA han producido daños documentables en comunidades específicas, construyendo un corpus de evidencia que alimente los mapas sistémicos y las evaluaciones Vulnerabilidad-Riesgo-Impacto de los reguladores, y que haga visible la dimensión humana de los riesgos que la regulación ágil debe gestionar.

ANEXOS TÉCNICOS

Acompañan a este documento varios anexos técnicos disponibles en la página del Centro-i y un paquete de fichas técnicas imprimibles pensadas para facilitar la aplicación de la metodología desarrollada en sesiones colaborativas, que pueden consultarse aquí:

https://centroi.org/project/regulacion-agil-para-la-gobernanza-de-la-inteligencia-artificial/

Anexo 1. Concepción ecosistémica de la gobernanza y la regulación ágil de la IA	https://centroi.org/regulacion-agil-ia/anexo-1-concepcion-ecosistemica/
Anexo 2. Evaluación del impacto de los sistemas de IA	https://centroi.org/regulacion-agil-ia/anexo-2-evaluacion-de-impacto/
Anexo 3. Ejemplos de aplicación de la metodología desarrollada en este manual	https://centroi.org/regulacion-agil-ia/anexo-3-ejemplos-de-aplicacion/
Anexo 4. Panorama de los sandboxes en Latinoamérica	https://centroi.org/regulacion-agil-ia/anexo-4-sandboxes-latinoamerica/
Anexo 5. Casos de sandboxes para la innovación de la IA en el mundo	https://centroi.org/regulacion-agil-ia/anexo-5-casos-de-sandboxes/
Anexo 6. Fichas de referencia para imprimir	https://centroi.org/regulacion-agil-ia/anexo-6-fichas-para-imprimir/

centroi
PARA LA SOCIEDAD DEL FUTURO