



Guía de aplicación

# AUTOEVALUACIÓN ÉTICA DE IA PARA ACTORES DEL ECOSISTEMA EMPRENDEDOR

Copyright © 2021. Banco Interamericano de Desarrollo.  
Esta obra se encuentra sujeta a una licencia Creative Commons  
IGO 3.0 Reconocimiento-NoComercial-SinObrasDerivadas  
(CC-IGO 3.0 BY-NC-ND) (<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier  
uso no-comercial otorgando el reconocimiento respectivo al BID.  
No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que  
no pueda resolverse amistosamente se someterá a arbitraje de  
conformidad con las reglas de la CNUDMI (UNCITRAL). El uso  
del nombre del BID para cualquier fin distinto al reconocimiento  
respectivo y el uso del logotipo del BID, no están autorizados por  
esta licencia CC-IGO y requieren de un acuerdo de licencia adicional.

Note que el enlace URL incluye términos y condiciones adicionales  
de esta licencia.

Las opiniones expresadas en esta obra son de los autores y  
no necesariamente reflejan el punto de vista del BID, de su  
Directorio Ejecutivo ni de los países que representa, así como  
tampoco del Comité de Donantes de BID Lab (FOMIN) ni de  
los países que representa.



fAIR LAC, la iniciativa del Grupo BID que promueve el uso ético y responsable de la inteligencia artificial, a través de BID Lab ha elaborado una herramienta práctica de autoevaluación ética de IA para emprendedores, que permite llevar a cabo un análisis de la solución tecnológica basada en IA y manejo de datos. Este diagnóstico ayuda a los emprendedores a mejorar su desarrollo de producto, mientras que identifica las principales áreas de atención para prevenir errores, sesgos, discriminaciones y exclusiones resultantes del despliegue tecnológico. La autoevaluación ética de IA para emprendedores que tiene en sus manos es el primer producto de fAIR LAC para emprendedores, y es una guía orientativa en español con enfoque multidisciplinario que contempla seis dimensiones principales: 1. Conceptualización y diseño, 2. Gobernanza y seguridad, 3. Involucramiento humano en los sistemas de IA, 4. Ciclo de vida de la IA (datos y algoritmos), 5. Actores relevantes y 6. Comunicaciones. El propósito es que los emprendedores tengan una referencia rápida de cuáles son los aspectos más importantes a considerar en cada una de estas dimensiones, a fin de tener una visión completa sobre las implicaciones éticas de sus productos y así establecer las medidas pertinentes de mejora y mitigación de riesgos.

La principal innovación de este documento radica en dos aspectos principales: el primero que no pone el peso de la autorregulación únicamente en los emprendedores, sino que involucra dos actores clave para el ecosistema: los financiadores de proyectos y las aceleradoras. El segundo aspecto es que las preguntas orientadoras se corresponden con tres niveles de desarrollo de las empresas desde etapas tempranas en fases de ideación hasta emprendimientos maduros o pequeñas y medianas empresas (PyMEs) desarrollando productos innovadores. Asimismo, si bien el documento está pensado principalmente para orientar en el desarrollo e implementación de soluciones basadas en IA, también es útil para aquellas soluciones basadas en el manejo de datos.

Invitamos a los interesados a descargar la publicación y a formar parte del viaje del emprendedor en su camino por desarrollar soluciones tecnológicas de impacto social que contribuyan a un desarrollo regional sostenido que no deje a nadie atrás.

# AUTORES

# AGRADECIMIENTOS

**César Said Rosales Torres**

**César Buenadicha**

**Tetsuro Narita**

Agradecemos por su apoyo a **Irene Arias**, CEO de BID Lab y

**Marcelo Cabrol**, gerente del Sector Social del BID.

Por su tiempo y valiosos aportes, expresamos un agradecimiento especial al equipo de BID Lab: **Carolina Carrasco, Ana Castillo, Smeldy Ramírez, William Ernest, Paula Auerbach, Jessica Leite, Patricio Aznar y Cecilia Franco**.

De igual manera, el equipo de BID Lab agradece la colaboración y retroalimentación de los siguientes expertos en el uso de inteligencia artificial, así como al Consejo Asesor del proyecto. El listado se presenta a continuación en orden alfabético:

**Adrián Soto**, profesor de la Facultad Ingeniería y Ciencias Universidad Adolfo Ibáñez. Chile.

**Cecilia Tham**, cofundadora y CEO de All Women. España.

**Constanza Gómez Mont**, fundadora y CEO de C Minds. México.

**Enrique Cortés Rello**, director del Hub de Inteligencia Artificial, Instituto Tecnológico y de Estudios Superiores de Monterrey. México.

**Francesca Gabetti**, fundadora y CEO de Team EQ. España.

**Irene Velasco**, fundadora y CEO de helKi. México.

**Iván Caballero**, fundador y CEO de Citibeats. España.

**Jesús Cepeda**, cofundador y CEO de OS City. México.

**Juan Carlos Holguin**, cofundador y CEO de LinkIn. Ecuador.

**Juan Eduardo Orlandi**, gerente general de Magical. Chile.

**Juan Roberto Hernández Villalobos**, coordinador de fAIR Jalisco del Instituto Tecnológico y de Estudios Superiores de Monterrey. México.

**Leo Prieto**, fundador y CEO de Odd Industries. Chile.

**Leopoldo Bertossi**, profesor de la Facultad Ingeniería y Ciencias de la Universidad Adolfo Ibáñez. Chile.

**Lorena Barrenechea**, asesora legal del Banco Interamericano de Desarrollo.

**Mara Balestrini**, consultora BID Lab y experta en transformación digital e interacción humano-computadora. Argentina/España.

**María Paz Hermosilla**, directora del GobLab de la Universidad Adolfo Ibáñez. Chile.

**Natalia González**, consultora del Banco Interamericano de Desarrollo.

**Ricardo Baeza-Yates**, director de Investigación, Instituto de IA Experiencial, Northeastern University, EE.UU.

**Roberto Sánchez**, consultor del Banco Interamericano de Desarrollo.

**Romina Garrido**, investigadora asociada del GobLab de la Universidad Adolfo Ibáñez. Chile.

**Sylvia Chebi**, directora ejecutiva de ThalesLab. Uruguay.

# ÍNDICE

<b>Introducción</b>	<b>6</b>
<b>¿A quién se dirige esta matriz de progresión y cómo se utiliza?</b>	<b>12</b>
<b>Conceptualización y diseño</b>	<b>14</b>
1.1 Objetivo de la IA en las operaciones de negocio	15
1.2 Evaluación del ecosistema digital	17
1.2.1 Valoración técnica	17
1.2.1.1 Determinación del nivel de madurez tecnológica (TRL por sus siglas en inglés)	
1.2.2 Valoración legal	19
1.2.3 Valoración sectorial	22
1.3 Impacto social y ambiental	25
<b>Gobernanza y Seguridad: Estructuras de gobernanza interna y control de la IA</b>	<b>28</b>
2.1 Estructura corporativa para la gobernanza de la IA	30
2.2 Gestión de riesgos y controles internos	32
2.2.1 Seguridad de los datos	34
<b>Involucramiento humano en los sistemas de IA</b>	<b>36</b>
3.1 Interacción seres humanos-IA	
Determinación del nivel de supervisión humana	38
<b>Ciclo de vida de la IA: Gestión de operaciones del sistema de IA</b>	<b>40</b>
4.1 Fuente y manejo de Datos	41
4.1.1 Tipo de datos	42
4.1.2 Procesamiento de datos	44
4.1.3 Integridad y confidencialidad de datos	45
4.1.4 Privacidad mediante diseño	46
4.1.5 Interoperabilidad	47
4.1.6 Ensayos y validación	48

# ÍNDICE

---

4.2 Desarrollo de modelos: Algoritmos	49
4.2.1 Trazabilidad	49
4.2.2 Explicabilidad de los procesos y resultados de la IA	50
4.2.3 Replicabilidad	52
4.2.4 Reproducibilidad	53
4.2.5 Auditabilidad	54
4.2.6 Mantenimiento	55
<b>Relacionamiento con actores clave</b>	<b>56</b>
5.1 Transparencia	58
5.2 Consumidores y usuarios del sistema	59
5.3 Acceso a financiamiento	60
<b>Comunicación</b>	<b>63</b>
<b>Anexo 1 - Descripción de las Etapas de Desarrollo de las Empresas</b>	<b>66</b>
<b>Fuentes</b>	<b>69</b>

---

# INTRODUCCIÓN

# INTRODUCCIÓN

**Durante la última década, la inteligencia artificial (IA) ha experimentado un crecimiento significativo, a pesar de haber tenido falsos comienzos en décadas anteriores causados principalmente por expectativas demasiado optimistas sobre el nivel de desarrollo tecnológico de la IA, la carencia de infraestructura de datos históricos, estructurados y de alta calidad en diversos sectores de la economía, así como diversos factores de mercado y uso comercial a gran escala de estas tecnologías (MGI, 2017).**

Actualmente el panorama es muy distinto, gracias a los desarrollos científicos de los últimos años y la aplicación comercial a escala de la IA en sectores relacionados con el desarrollo de vehículos autónomos, procesamiento de lenguaje natural, reconocimiento facial, visión computarizada, entre otros (MGI, 2017; OCDE, 2018). Esta dinámica a su vez ha despertado mayor interés por parte de los inversionistas, así como de los gobiernos por promover ecosistemas que faciliten el crecimiento de esta tecnología y que a su vez regulen

aspectos importantes como la protección de datos, la interacción humana-IA, y el impacto que esta podría tener en la vida social y política de las personas (OCDE, 2018).

No obstante, a pesar de estos avances, el grado de adopción de la IA permanece relativamente bajo en el mundo comparado con otras tecnologías y altamente concentrado en zonas geográficas como Estados Unidos, Canadá, China y Europa, cada uno con modelos de gobernanza e inversión muy distintos (Gigler 2020).



Asimismo, existe una brecha de financiamiento y acceso a recursos entre grandes corporaciones que utilizan la IA y startups. Por su parte, en América Latina y el Caribe el nivel de desarrollo de la IA permanece bajo comparado con el resto de regiones puntero y concentrado en países como Argentina, Brasil, Chile, Colombia, México, Perú y Uruguay, donde principalmente se utiliza para desarrollos tecnológicos basados en modelos de terceros, chatbots, procesamiento de lenguaje natural y conversión de texto, y en menor medida para tareas de clasificación y predicción (Everis & Endeavor 2018). Además de la disparidad en los niveles de desarrollo e inversión con Estados Unidos, Canadá, China y Europa, América Latina y el Caribe se enfrenta a bajos niveles de inversión tanto pública como privada, así como limitaciones para el desarrollo profesional de expertos en manejo de datos y la creación de cadenas de valor de datos sobre las cuales los sistemas de IA basados en aprendizaje automático puedan operar de forma más eficiente (OCDE 2018).

En este contexto, desde finales de 2019 el Banco Interamericano de Desarrollo (BID) y BID Lab lideran una iniciativa denominada fAIr LAC para promover la adopción y el uso ético y responsable de la IA en la región.

El objetivo principal es el de aprovechar el inmenso potencial de la tecnología para crear servicios sociales más eficientes, justos y personalizados para los ciudadanos de América Latina y el Caribe (BID 2020). Cada vez más países hacen uso de este tipo de tecnologías que toca prácticamente todas las áreas del desarrollo en la región, por lo que resulta fundamental respetar la privacidad de los datos de los ciudadanos y prever posibles sesgos en la construcción de algoritmos (Google 2018).

En el marco de esta gran alianza, BID Lab inició en 2020 el proyecto “Entrepreneurial Journey: articulación de un ecosistema ético y responsable de startups de inteligencia artificial”, o también denominado fAIr LAC Entrepreneurial Journey, que tiene por objetivo coadyuvar al fortalecimiento de ecosistemas emprendedores para la adopción de la IA, a través del co-diseño de actividades, productos y servicios con emprendedores, inversionistas, aceleradoras de negocios, expertos en datos, academia, sociedad civil y gobiernos. Esto se logra a través de herramientas y servicios que permitan a los emprendedores hacer uso de esta tecnología de modo dinámico e innovador, en línea con principios éticos que maximicen beneficios y minimicen riesgos.

La autoevaluación ética para actores del ecosistema emprendedor que tiene en sus manos es el primer producto del proyecto fAIr LAC Entrepreneurial

Journey, y es una guía orientativa en español para empresas, inversionistas y aceleradoras sobre aspectos éticos a considerar en el desarrollo en implementación de soluciones tecnológicas basadas en el manejo de datos de IA, de acuerdo con su nivel de madurez. El resultado de este ejercicio resultará en la identificación de áreas de mejora que permitirán mejorar la calidad de los productos, servicios, procesos, modelos y sistemas en desarrollo, así como identificar las principales áreas de riesgo en materia de ética.

Lo innovador de este documento radica en dos aspectos principales: el primero que no pone el peso de la autorregulación únicamente en los emprendedores, sino que involucra dos actores clave para el ecosistema: los financiadores de proyectos y las aceleradoras. El segundo aspecto es que las preguntas orientadoras se corresponden con tres niveles de desarrollo de las empresas desde etapas tempranas en fases de ideación



hasta emprendimientos maduros o pequeñas y medianas empresas (PyMEs) desarrollando productos innovadores. Asimismo, si bien el documento está pensado principalmente para orientar en el desarrollo e implementación de soluciones basadas en IA, también es útil para aquellas soluciones basadas en el manejo de datos.

Por otro lado, la elaboración de este producto descansa en los siguientes principios para la adopción ética y responsable de la IA para el desarrollo de América Latina y el Caribe, a la vez que se consideran los principios éticos para la IA de la OCDE (OCDE 2019), a saber:

### IA para el desarrollo

**El objetivo de este proyecto es promover iniciativas que utilicen la IA para propósitos de impacto social positivo y escalable, a fin de contribuir en la mejora del bienestar económico y social de los ciudadanos de la región.**

### Integridad

**El desarrollo y uso de sistemas de IA deberían ser coherentes con su propósito de impacto social, además de respetar el uso de los datos y metadatos de clientes y terceros con sistemas que permitan garantizar su fiabilidad.**

### Confianza

**La confianza en el uso de la tecnología de IA debería estar basada en un enfoque de prevención de daños y mitigación de riesgos, así como la transparencia de sus operaciones. Todo ello a fin de garantizar principios de equidad y no discriminación en el uso de esta tecnología.**

Finalmente, cabe señalar que la dimensión ética de este producto no solo se observa a través del cumplimiento legal de las obligaciones de las empresas, sino a través de la solidez de las propuestas y los objetivos

de uso tecnológico para el beneficio, empoderamiento y protección de los bienes comunes, resaltando sobre todo la dignidad humana y los principios sociales comúnmente aceptados.

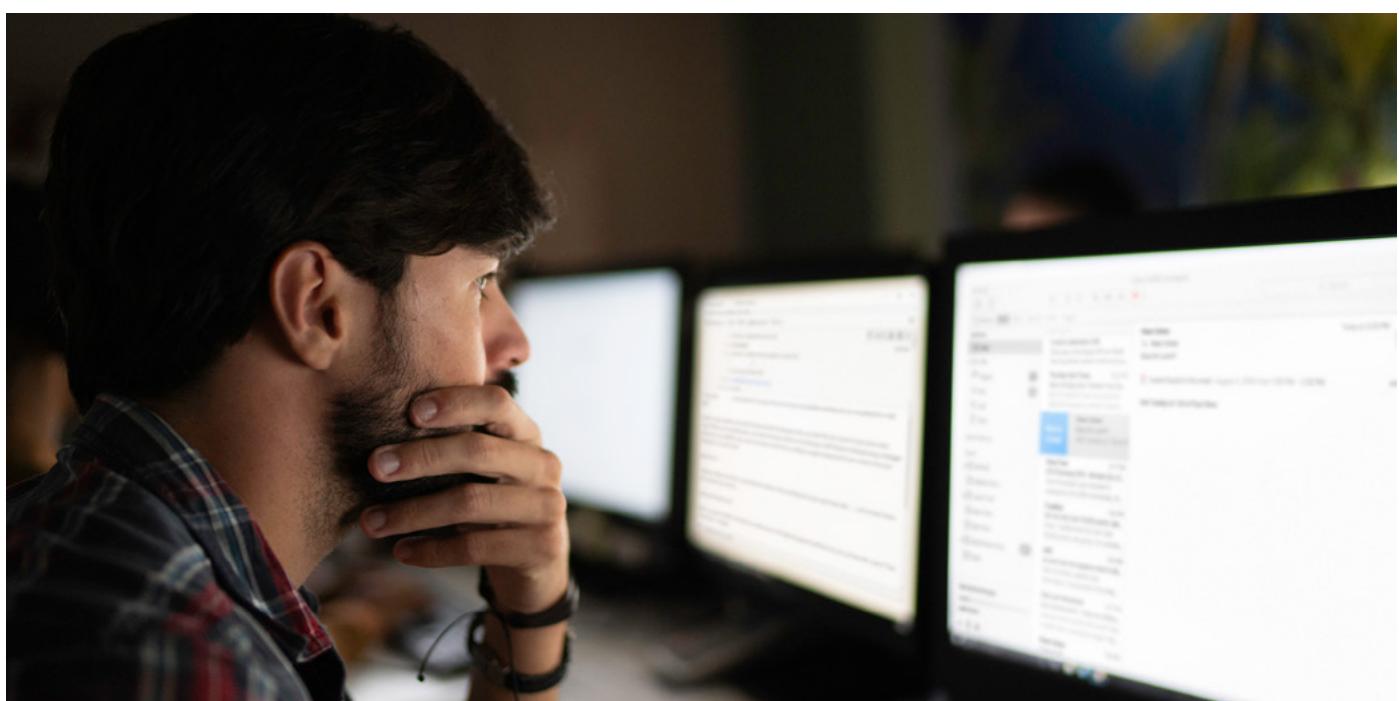


# ¿A QUIÉN SE DIRIGE ESTA MATRIZ DE PROGRESIÓN Y CÓMO SE UTILIZA?

**Esta matriz de autoevaluación ética es una herramienta de adopción voluntaria dirigida a quienes diseñan, desarrollan, despliegan o utilizan tecnologías de IA o conducente a ella, lo que incluye a empresas en diferentes etapas de maduración, inversionistas y aceleradoras. También es una herramienta de utilidad para todos aquellos interesados en los procesos de adopción de la IA de un modo ético y responsable.**

La herramienta es una guía general no exhaustiva de preguntas orientadoras que permitan identificar el propósito de uso de la IA, fortalecer las estructuras de gobernanza y controles internos en torno a los sistemas de IA, determinar el mejor equilibrio de involucramiento humano en la IA, dar seguimiento a los procesos de control de datos y algoritmos, así como sugerencias sobre el relacionamiento con actores clave como usuarios y clientes, inversionistas y entidades supervisoras.

Se trata de una matriz de autoevaluación ética progresiva, ya que está dividida en tres etapas (cuyo detalle se encuentra en el Anexo 1 de este documento) de acuerdo al nivel de desarrollo del emprendimiento, al nivel de desarrollo tecnológico y a los actores que participan en cada una de estas etapas. Dichas etapas no constituyen conceptos rígidos preestablecidos, sino que ofrecen una orientación mínima para usar y entender mejor las preguntas orientadoras en cada



una de las dimensiones. En este sentido, si bien las preguntas pretenden llevar un orden según el nivel de desarrollo de las empresas, algunas de sus preguntas pueden ser aplicables a las diferentes etapas según las particularidades del negocio o sector en el que se opere. Asimismo, es relevante recordar que esta guía no es un listado de cumplimiento, sino un listado versátil y adaptable a las necesidades de cada emprendedor, inversionista, o cualquier interesado en el uso ético y responsable de la IA.

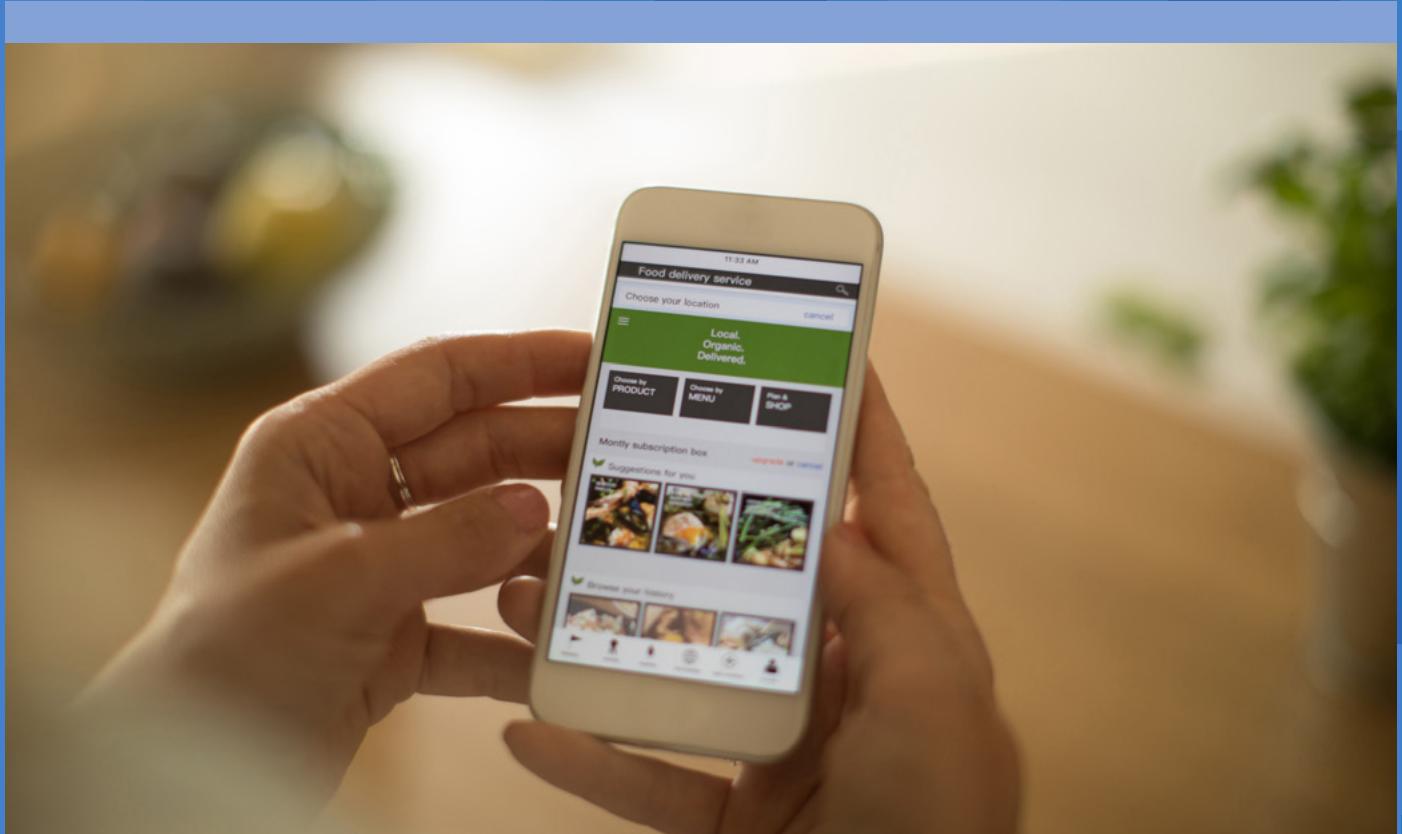
Esta es la primera versión de la matriz, la cual ha contado con la **revisión, comentarios y orientaciones de emprendedores, inversionistas, expertos legales y científicos de datos, a fin de garantizar la facilidad de uso y pertinencia del documento.**

La elaboración de este producto se basa en un **proceso iterativo en el que se tiene planificada una ronda de consultas amplia con más empresas, académicos, sociedad civil, gobiernos, inversionistas, abogados y científicos de datos** para mejorar la matriz y llevar a cabo una fase de pruebas piloto que permita identificar acciones estratégicas para el robustecimiento del ecosistema emprendedor en materia de IA.

Esperamos que esta herramienta sea de utilidad para todos aquellos interesados en el uso **ético y responsable de la IA y que contribuya al fortalecimiento del ecosistema a fin de promover la competitividad de las empresas en la región altiempo que se alcanza un impacto social positivo.**



# 1



## CONCEPTUALIZACIÓN Y DISEÑO

# 1 CONCEPTUALIZACIÓN Y DISEÑO

**Esta primera sección comprende diversos aspectos que se deberían contemplar de forma previa a la adopción de la IA como tecnología base para los emprendimientos. Dentro de estos aspectos se incluye la determinación del propósito operativo de la IA, así como aquellas valoraciones legales, técnicas y sectoriales necesarias para definir el grado de idoneidad de la tecnología en las operaciones de negocio. Asimismo, se incluye un apartado de preguntas guía relacionadas con el impacto social.**

---

## 1.1 OBJETIVO DE LA IA EN LAS OPERACIONES DE NEGOCIO

El primer paso para la adopción y uso ético y responsable de la IA es definir con precisión la utilidad que generará dicho sistema para el modelo de negocio, así como su idoneidad dentro de la estrategia corporativa. Esta etapa requiere de una valoración realista y tangible sobre las capacidades y limitaciones que tendría el sistema de IA (WEF & IMDA 2020). Asimismo, es recomendable llevar a cabo un análisis pormenorizado de casos de uso similares a fin de maximizar las posibilidades de éxito y reducir el margen de incertidumbre de la iniciativa con base en la experiencia previa de aplicación

de la IA. Incluso, a fin de entender la utilidad del sistema de IA se pueden llevar a cabo comparativas de solución con o sin IA.

En este primer momento conviene tener en mente que para la adopción ética y responsable de la IA, es recomendable pensar en soluciones ante problemas reales para los cuales la IA es la tecnología más adecuada, más allá de simplemente incorporar la IA para ofrecer al mercado productos interesantes que usen esta tecnología (MGI 2017; Szramke 2017; EC 2018; IMDA & PDPC 2020).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿La organización o empresa se ha planteado un <b>objetivo con métricas de impacto para sus operaciones de negocio en las que la IA sea la tecnología más adecuada</b>, más allá de solo ofrecer en sus productos o servicios la incorporación de una tecnología innovadora?</p> <p>¿Qué <b>tipo de tareas</b> se pretenden realizar con el sistema de IA, por ejemplo, toma de decisiones automáticas, predicciones, recomendaciones, interacción con los usuarios, etc?</p> <p>¿Qué <b>tipo de categoría tecnológica de IA</b> se va a utilizar, por ejemplo: aprendizaje automático o machine learning (ML), aprendizaje profundo o deep learning, redes neuronales artificiales, visión artificial o computer vision, procesamiento de lenguaje natural o natural language processing (NLP), agentes virtuales, robótica inteligente, o vehículos autónomos?</p>	<p>¿Se han explorado ejemplos de <b>casos de uso en los que se hayan utilizado tecnologías de IA o similares</b> a los propósitos establecidos por la compañía para su caso particular?</p> <p>¿Se han incorporado las lecciones aprendidas de <b>casos de uso similares</b> en cuanto al tipo de tecnología de IA y sector?</p>	<p>¿Se han realizado <b>estudios comparativos entre diferentes tecnologías</b> o métodos para determinar si el uso de la IA es el más sencillo y adecuado? Por ejemplo mediante el uso de técnicas de programación o de simulación que no hagan uso de IA.</p> <p>En caso de que la IA sea el conjunto de tecnologías más adecuado para los propósitos de la empresa, ¿se ha hecho una <b>comparación entre los diferentes modelos de IA</b> para determinar cuál sería el más eficiente? Por ejemplo, <i>entre modelos basados en reglas o modelos basados en el aprendizaje automático a través de big data</i>.</p> <p>Una vez determinado el tipo de modelo de IA, ¿se ha determinado qué <b>tipo de algoritmo</b> resultaría el más adecuado para utilizar? Por ejemplo: <i>algoritmos de regresión, árboles de decisión, clusterizaciones, o redes neuronales</i>.</p>
<b>INVERSIONISTAS</b>	<p>¿El planteamiento del despliegue de la IA, además de tener un claro propósito comercial, <b>está en línea con la tesis de inversión de los inversionistas</b>?</p>	<p>¿La empresa se encuentra incursionando hacia la IA <b>proveniente de otro tipo de tecnología de uso de datos</b>?</p>	<p>¿La selección del tipo de IA es <b>congruente con los propósitos para los que se va utilizar la IA</b>?</p>
<b>ACELERADORAS</b>	<p>¿Los socios fundadores de la empresa cuentan con un <b>entendimiento</b> mínimo sobre las ventajas y riesgos de la adopción de la IA?</p>	<p>¿Los socios y empleados de la empresa cuentan con un <b>entendimiento</b> mínimo sobre las ventajas y riesgos de la adopción de la IA?</p>	Idem.

# 1.2 EVALUACIÓN DEL ECOSISTEMA DIGITAL

La evaluación del ecosistema digital permite a las empresas llevar a cabo un reconocimiento de la infraestructura tecnológica que permita valorar la viabilidad y utilidad del desarrollo de un sistema de IA a través de determinadas características técnicas, legales y sectoriales (Galdon 2020).

## 1.2.1 Valoración técnica

Esta dimensión permite evaluar en términos técnicos la disponibilidad de datos e infraestructura crítica mínima para garantizar la operabilidad del sistema de IA con un enfoque de prevención de riesgos (EC 2018; IMDA & PDPC 2020).

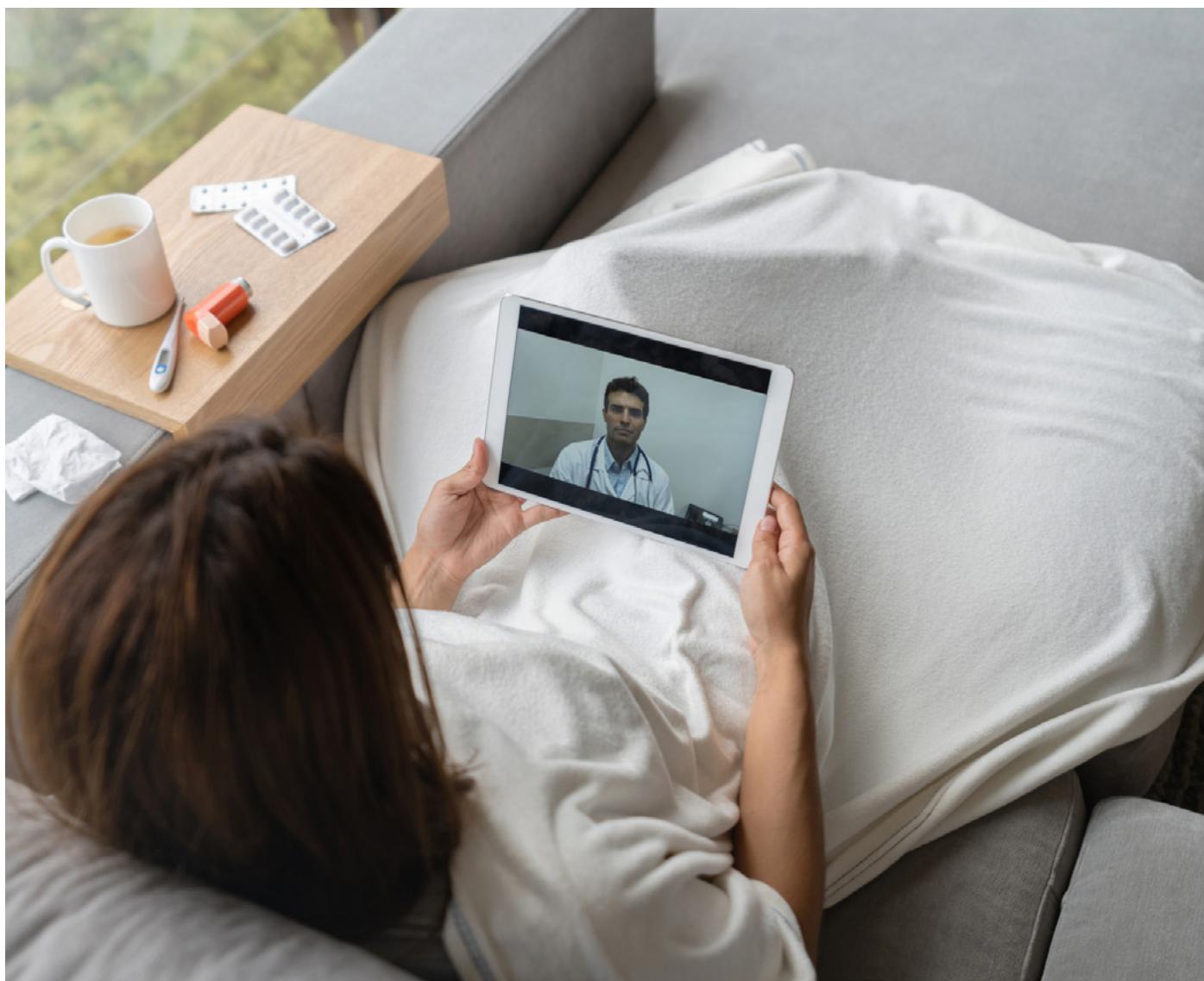


	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Existen diversas <b>fuentes de datos históricos</b> que permitan la creación de un modelo de IA? Por ejemplo <i>data lakes</i>, iniciativas de gobierno como <i>smart cities</i>, etc.</p> <p>¿Es posible <b>acceder a los datos</b> de dichas fuentes?</p> <p>¿Cómo se garantiza la obtención de <b>datos completos, actualizados y de alta calidad</b>?</p> <p>¿Existe una <b>base de tecnología de información y telecomunicación</b> suficiente para garantizar la operatividad constante del modelo de IA?</p> <p>¿La iniciativa ha tomado en consideración las recomendaciones emitidas por <b>organismos internacionales de normalización</b> que faciliten el cumplimiento de las obligaciones legales?</p>	<p>¿Se ha considerado la necesidad de <b>requerir datos adicionales con la finalidad de reducir sesgos</b>?</p> <p>Con respecto a la <b>arquitectura del sistema de IA</b>, ¿se han establecido el conjunto de normas y restricciones sobre las cuales el sistema de IA debería operar y no transgredir?</p> <p>En el caso de utilización de modelos de IA para robótica, ¿se han tomado consideraciones para determinar la <b>suficiencia de todos los elementos técnicos requeridos con base en un modelo SMPA</b> (Sense-Model-Plan-Act) de modo que se puedan reconocer todos los elementos del entorno y que la consideración de planes y actuación del sistema cumplan con las expectativas de comportamiento y resultados? En este sentido es importante definir claramente las expectativas de comportamiento del sistema.</p>	Idem.
<b>INVERSIONISTAS</b>	¿A través de qué <b>tipo de interfaz</b> se va a interactuar con los clientes, sitio web, app, etc?	¿La información del sistema de IA es también accesible a usuarios que utilicen <b>diversos tipos de plataformas tecnológicas</b> ?	¿Cuál es el <b>nivel de madurez tecnológica</b> de la iniciativa?
<b>ACCELERADORAS</b>	¿Se requiere de algún <b>soporte de negocios o técnico</b> en materia de datos para consolidar el proyecto?		

### **1.2.1.1 Determinación del nivel de madurez tecnológica (TRL por sus siglas en inglés)**

Particularmente para algunos casos que pretenden iniciar rondas de financiamiento, es recomendable determinar el nivel de madurez tecnológica del desarrollo que se propone. Usualmente organizaciones internacionales como la Unión Europea utilizan un marco conceptual con la siguiente progresión (European Commission 2014-2015):

<b>Nivel 1</b> <b>Principios básicos observados y reportados;</b>	<b>Nivel 6</b> <b>Demostración de prototipo en entorno relevante;</b>
<b>Nivel 2</b> <b>Concepto y aplicación tecnológica formulada;</b>	<b>Nivel 7</b> <b>Demostración de prototipo de sistema en entorno operativo;</b>
<b>Nivel 3</b> <b>Prueba experimental del concepto;</b>	<b>Nivel 8</b> <b>Sistema completo y certificado a través de pruebas y demostraciones;</b>
<b>Nivel 4</b> <b>Validación de componentes y disposición en entorno de laboratorio;</b>	<b>Nivel 9</b> <b>Sistema probado con éxito en entorno real.</b>
<b>Nivel 5</b> <b>Valoración de componente y disposición en entorno relevante;</b>	



Asimismo, dependiendo del nivel de impacto de la solución de IA en la vida social de las personas, debería cumplir con normativas de protección a comunidades vulnerables tales como niños, discapacitados y poblaciones históricamente desfavorecidas o en riesgo de exclusión (Reglamento General de Protección de Datos de la Unión Europea 2016-2018). En este mismo sentido, debería cumplir con los estándares internacionales en materia

de derechos humanos, sobre todo para aquellas iniciativas encaminadas a la solución de problemas sociales en materia de salud, educación, trabajo y economía. En particular, es necesario verificar que el uso de IA respete el derecho a la no discriminación de las personas, ya que el uso de datos sesgados podría generar la perpetuación de tratos desiguales a grupos poblacionales que tradicionalmente sufren de un trato desfavorable.

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿El sistema cumple con las <b>regulaciones en materia de protección y privacidad de datos</b>, protección al consumidor desde su diseño?</p> <p>¿Existen regulaciones sectoriales específicas que resulten aplicables al sistema? ¿el sistema cumple con estas regulaciones?</p> <p>¿El sistema de IA tendrá algún tipo <b>de interacción con poblaciones vulnerables</b> que requieran de consideraciones especiales?</p>	<p>¿El sistema tendrá algún <b>impacto relevante en la vida social, económica o política de la comunidad</b> que requiera de la revisión de normativas particulares en este respecto?</p>	<p>En caso de que el sistema de IA esté siendo utilizado para resolver problemas sociales, ¿se han revisado los <b>estándares internacionales más relevantes en materia de protección de derechos humanos</b>, tales como el derecho a la no discriminación y otros derechos fundamentales, principalmente en el marco de la Convención Interamericana sobre Derechos Humanos, protocolos y resoluciones adicionales?</p>
<b>INVERSIONISTAS</b>	<p>¿Se ha determinado si la iniciativa cumple con la normativa aplicable y si el eventual <b>incumplimiento de las obligaciones legales</b> de la empresa podrían causar la interrupción del negocio, sanciones o multas que lo hagan inviable o, en casos limitados en la región, transferir la responsabilidad a las entidades financieradoras?</p> <p>Incluso si se cumple con las normativas nacionales, ¿la empresa cumple las normativas del resto de países donde tiene operaciones?</p>	<p>¿La información del sistema de IA es también accesible a usuarios que utilicen <b>diversos tipos de plataformas tecnológicas</b>?</p>	<p>¿Cuál es el <b>nivel de madurez tecnológica</b> de la iniciativa?</p>
<b>ACELERADORAS</b>	<p>¿Se ha llevado a cabo un <b>análisis de factibilidad legal</b> del modelo de IA?</p> <p>¿La empresa ha protegido su propiedad intelectual mediante el registro de una patente o correspondiente según el caso?</p> <p>En el caso de que <b>los modelos de negocio no estén contemplados en la legislación vigente</b>, ¿se han estructurado los mecanismos legales que permitan su operabilidad? Tales como permisos y autorizaciones o medidas de mitigación de riesgos.</p>		

### 1.2.3 Valoración sectorial

La IA es un conjunto de tecnologías que por su versatilidad puede ser utilizada en una gran variedad de sectores e industrias. Esta diversidad requiere de la conducción de análisis específicos sobre las dinámicas que se presentan

en cada industria. Esta sección ofrece algunos parámetros mínimos a tener en consideración para la identificación de oportunidades de uso ético y responsable de la IA (EC 2018; Omidyar 2019; Galdon 2020; MacCarthy 2020; WEF & IMDA 2020).



	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Qué tipo de necesidad cubre el sistema de IA, una identificada previamente o una que no había sido registrada previamente en el sector?</p> <p>En caso de que la empresa haya decidido construir sus propios datos, ¿esta ha implementado una <b>metodología que permita conocer el impacto sistémico</b> del sector en el que trabajará?</p> <p>¿Los <b>datos</b> con los que se cuenta son <b>representativos</b> de la base de usuarios del sector?</p> <p>¿Los datos con los que se cuenta provienen de fuentes <b>heterogéneas</b> dentro del mismo sector y existen suficientes fuentes de datos para el uso de sistemas de IA basados en aprendizaje automático?</p> <p>¿Los ingenieros y desarrolladores conocen el entorno sectorial específico en el cual van a trabajar en términos de cadenas de valor y sistemas de datos?</p> <p>¿Se ha conversado con los tomadores de decisión la importancia de la calidad de los datos para el impacto en la toma de decisiones?</p>	<p>¿La empresa cuenta con un grupo especializado en la gobernanza de IA que realice <b>evaluaciones periódicas sobre preocupaciones públicas o de política pública emergentes</b> sobre el uso del sistema de IA en el sector en el que opera?</p> <p>¿Se han adoptado <b>estándares sectoriales</b> para el mejor uso de datos y tecnologías como la IA?</p> <p>¿Existen métodos continuos de formación y <b>capacitación interna y externa acerca de la ética, pensamiento sistémico de la IA y ciencia de datos</b>?</p>	<p>¿Se cuenta con la <b>colaboración de otros actores clave</b> del sector para las revisiones periódicas del funcionamiento del sistema de IA?</p> <p>¿Se han establecido <b>acuerdos o protocolos de protección de datos con el resto de socios</b> y actores dentro de la cadena de valor sectorial?</p> <p>¿Se han llevado a cabo <b>auditorías internas que evalúen la ejecución de las mejores prácticas establecidas en las etapas tempranas</b>?</p>
<b>INVERSIONISTAS</b>	<p>¿Qué tanto <b>conocen las empresas el sector</b> para el que plantean desarrollar una solución de IA?</p> <p>¿El producto o solución <b>resuelve un problema relevante para el sector</b>?</p> <p>¿Se ha <b>comunicado a los inversionistas</b> que la adquisición y estructuración de datos de calidad requiere de consideración dentro de las partidas de la inversión?</p>	<p>¿Los <b>beneficios</b> que ofrece el sistema podrían ser escalables a otras áreas de la cadena de valor o incluso a otros sectores relacionados?</p>	

## ACELERADORAS

¿Cuentan con asesores, mentores o herramientas para **comunicar la importancia de las consideraciones éticas** en el uso de la IA y ciencia de datos?

¿Existe un mecanismo de **acercamiento entre emprendedores y actores del sector** que permita conocer el impacto de la solución tecnológica desarrollada?



# 1.3 IMPACTO SOCIAL Y AMBIENTAL

De acuerdo con el tipo de funcionalidad que tenga el sistema de IA, hay ocasiones en las que este podría tener repercusiones de gran calibre para el bienestar social y ambiental de las comunidades en las que opera. Por lo anterior, resulta fundamental que la organización que adopta o usa la IA tenga en cuenta medidas que aseguren la diversidad, la no discriminación y la equidad en sectores como la educación, salud, medio ambiente, economía,

y mercados laborales (EC 2018; Omidyar 2019).

Asimismo, el sistema de IA deberá tener en especial consideración su impacto o influencia en poblaciones vulnerables, tales como: menores de edad, personas con discapacidad, grupos históricamente desfavorecidos y en riesgo de exclusión (Reglamento General de Protección de Datos de la Unión Europea 2016-2018).



	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿El sistema respeta principios fundamentales de equidad y no discriminación en general y hacia poblaciones vulnerables?</p> <p>¿El sistema de IA es la mejor herramienta para generar un <b>impacto social</b> positivo o existen otras alternativas?</p> <p>¿Se ha analizado el <b>potencial impacto social negativo del uso de IA, para analizar la gravedad de los errores?</b> Por ejemplo a través de una <i>matriz de confusión</i> que permita visualizar el desempeño de los algoritmos</p> <p>¿Se ha valorado cuál es <b>el alcance del impacto social que tendrá el uso del sistema de IA</b>, con el fin de establecer <b>medidas proporcionales de mitigación de riesgos?</b> No tendría el mismo nivel de impacto el desarrollo de un sistema de recomendaciones a usuarios, que el de un sistema de toma de decisiones para la selección de beneficiarios de un programa de asistencia social o de salud.</p> <p>En caso de que el sistema de IA esté diseñado para ser usado por menores de edad, personas con discapacidad o que se encuentran en condiciones de riesgo de exclusión, ¿qué <b>medidas de protección existe para estos colectivos?</b></p>	<p>¿El sistema de IA <b>alimenta algún proceso de toma de decisiones que tengan impacto social?</b> En caso afirmativo, ¿qué <b>medidas se toman para asegurar principios de equidad y no discriminación?</b></p> <p>¿Existe una estrategia o procedimientos para <b>contrarrestar sesgos provenientes</b> de la realidad, tales como discriminaciones y prejuicios?</p> <p>¿Existen los mecanismos que permitan <b>señalar problemas vinculados con sesgos y discriminación por el uso de la IA?</b></p> <p>Aunado a lo anterior, ¿el sistema cuenta con <b>mecanismos de corrección de datos</b> que permitan conocer y corregir los criterios de una <b>decisión automatizada?</b></p> <p>Derivado del punto anterior, ¿existen <b>mecanismos que permitan a los usuarios presentar, tramitar y dar seguimiento a reclamaciones?</b></p> <p>¿La política de consentimiento sobre el uso del sistema de IA incluye <b>provisiones con respecto a la autorización de uso a menores de edad por sus padres o tutores, así como personas discapacitadas?</b></p>	<p>¿La solución <b>genera algún tipo de asimetría de poder o de información?</b> En caso afirmativo, ¿existen los mecanismos para identificar oportunamente y corregir estos desbalances?</p> <p>¿Se lleva a cabo algún procedimiento de <b>evaluación y corrección</b> en caso de que se presente un impacto social negativo?</p> <p>En su debido caso, ¿se han establecido <b>mecanismos que midan el impacto ambiental del despliegue del sistema de IA?</b></p> <p>¿Se han diseñado <b>medidas para reducir posibles impactos negativos en el uso del sistema de IA</b> durante el ciclo del proyecto?</p>
<b>INVERSIONISTAS</b>	<p>¿Existe un <b>planteamiento y delimitación sobre un problema social específico a resolver</b> para el que la IA sea la tecnología más sencilla y adecuada de usar?</p> <p>¿Se ha involucrado o <b>consultado a la comunidad relevante</b> sobre el desarrollo del sistema de IA?</p>	<p>¿Se cuenta con <b>métricas para valorar el impacto social y económico</b> del despliegue del sistema de IA?</p>	<p>¿La empresa lleva a cabo una <b>valoración de impacto periódica</b> sobre el uso del sistema de IA en torno al problema específico que aborda, utilizando metodologías como por ejemplo la <i>Pirámide de Necesidades de Maslow</i>?</p>

## ACCELERADORAS

¿El producto es comercialmente viable y **ofrece beneficios a los consumidores, individuos y comunidades** donde se utilizará?

¿La empresa **documenta todos los posibles riesgos del uso del sistema de IA** y las acompaña de estrategias de mitigación?



# 2



## GOBERNANZA Y SEGURIDAD: ESTRUCTURAS DE GOBERNANZA INTERNA Y CONTROL DE LA IA

# 2

## GOBERNANZA Y SEGURIDAD: ESTRUCTURAS DE GOBERNANZA INTERNA Y CONTROL DE LA IA

**Las empresas pueden crear o actualizar sus marcos de gobernanza y controles internos a fin de garantizar la transparencia y rendición de cuentas de la IA vinculada con su desarrollo, despliegue y utilización. La selección de estos modelos depende de la estructura de la organización, tamaño y recursos disponibles, así como del sector particular en el que opera (Omidyar 2019; WEF & IMDA 2020).**

**Esta sección ofrece algunas preguntas orientadoras que ayudan a analizar e identificar las mejores estructuras de gobernanza corporativa que ayuden a garantizar la adopción ética y responsable de la IA.**



## 2.1 ESTRUCTURA CORPORATIVA PARA LA GOBERNANZA DE LA IA

La estructura corporativa de las empresas que desean adoptar o que utilizan la IA debería contar con arreglos organizacionales que permitan la identificación, evaluación, documentación y resolución de tensiones fundamentales entre sus principios y los problemas técnicos que resulten del uso de esta tecnología (WEF & IMDA 2020). En este sentido, se debe valorar qué tipo de modelo de gestión sobre la IA o datos resulta más adecuado, por ejemplo, entre sistemas centralizados o descentralizados con cierto grado de intervención humana o sin ella (Google 2018).

De forma transversal se recomienda la creación de un consejo de ética con ejecutivos y directivos directamente

involucrados con las diversas áreas de operación de la IA, tales como la dirección y el consejo de administración, el departamento jurídico, desarrollo de productos, garantía de calidad, recursos humanos, compras y adquisiciones, desarrolladores y operaciones diarias, a fin de dirimir controversias resultantes del uso de los sistemas (WEF & IMDA 2020). Derivado de lo anterior, es recomendable que las empresas cuenten con estructuras de comunicación directa con grupos externos de supervisión pública o sectorial que permitan el intercambio de buenas prácticas, el debate en torno a problemas o controversias, así como la notificación oportuna de problemas emergentes resultantes del uso de dicha tecnología.



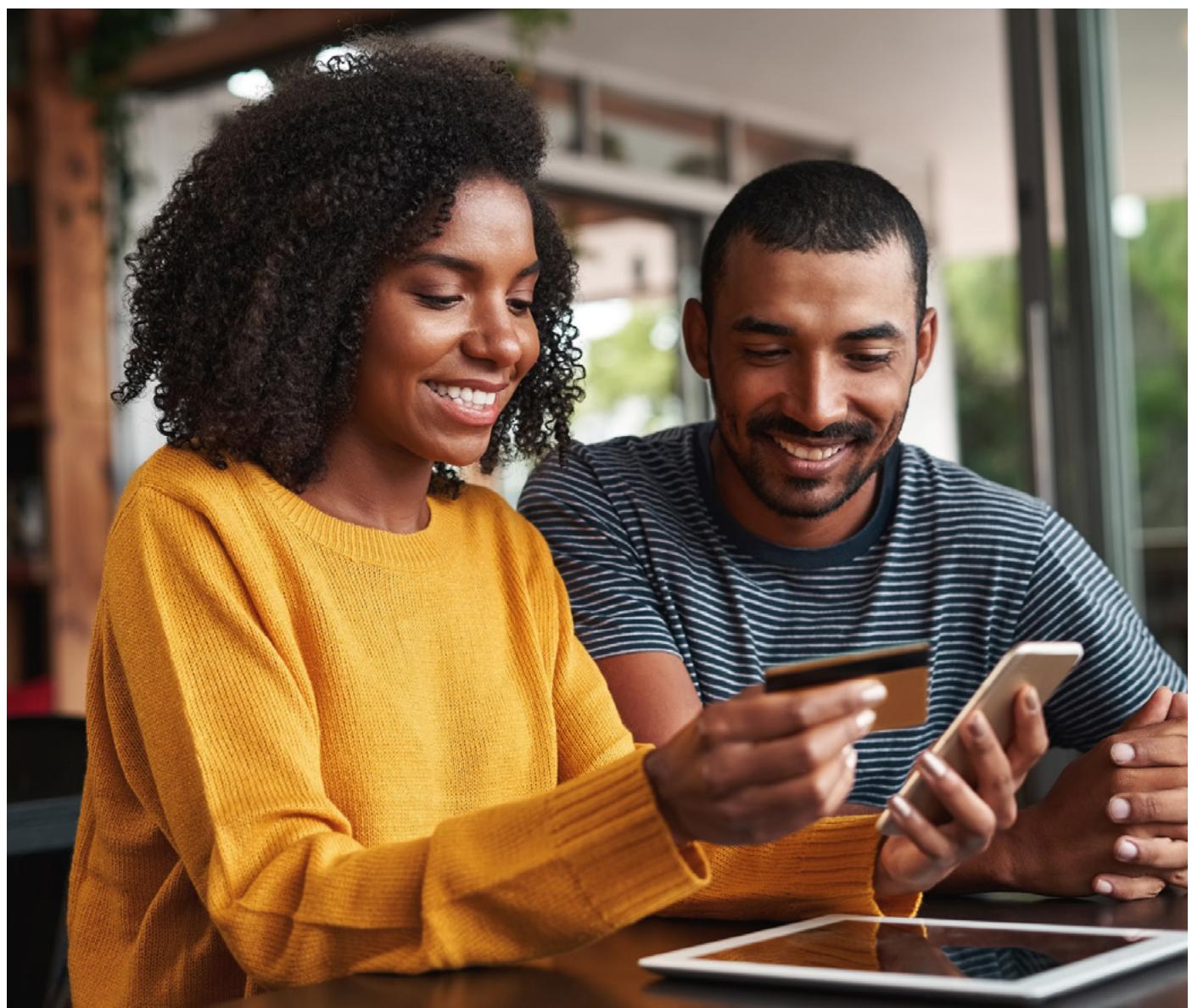
	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se ha designado dentro de la empresa a un responsable para la valoración ética de la adopción de la IA?</p> <p>¿Se ha realizado una <b>valoración sobre si la estructura corporativa</b> de la empresa debería modificarse en algún modo para facilitar la adopción de la IA, especialmente en caso de crecimiento acelerado de la empresa?</p> <p>¿El <b>personal</b> encargado de la supervisión, manejo y control del sistema de IA <b>está adecuadamente capacitado y conoce sus roles y responsabilidades?</b></p> <p>¿Los equipos vinculados al manejo del sistema de IA tiene a su <b>disposición las herramientas y capacidades necesarias</b> para cumplir de forma adecuada con sus responsabilidades?</p>	<p>¿Se ha considerado la creación de una <b>mesa directiva de valoración</b> ética sobre la adopción de la IA?</p> <p>¿Se prevé la utilización de una <b>guía externa proveniente de una fuente confiable para supervisar temas vinculados con la ética</b> y rendición de cuentas, además de las iniciativas internas?</p> <p>¿Se ha considerado la <b>adaptación de las estructuras de gestión de riesgos</b> para la incorporación de procesos de control de la IA?</p> <p><b>¿Existe un Código de Conducta</b> sobre el uso y operaciones del sistema de IA?</p>	<p>¿Existen mecanismos institucionales para la <b>coordinación con respecto al sistema de IA con terceras partes</b>, tales como proveedores, consumidores, distribuidores y otros trabajadores?</p> <p>¿Se ha considerado la creación de un <b>órgano colegiado para el control de la IA</b> compuesto por los representantes de departamentos que trabajan directamente con el sistema de IA?</p> <p>¿Se ha elaborado un <b>Código de Conducta sobre la IA</b> para distribuir a todos los empleados de la empresa para que conozcan los procesos de gobernanza y control de la IA?</p>
<b>INVERSIÓNISTAS</b>	<p>¿Se ha compartido de forma clara la <b>estructura corporativa de la organización y de los roles específicos</b> que tienen los ejecutivos en materia de gobernanza y control de la IA?</p> <p>¿La empresa cuenta con <b>Chief Technology Officer, Chief Product Officer y/o Chief Data Officer o Data Protection Officer</b>? ¿Se le ha asignado la responsabilidad específica de liderar la valoración ética de la adopción de IA?</p> <p>¿La empresa cuenta con un <b>equipo sólido de científicos de datos</b> e ingenieros?</p>	<p>¿Existe una distinción entre las <b>responsabilidades y el involucramiento del personal técnico y el personal de negocios</b> en cuanto a sus atribuciones con respecto al sistema de IA?</p>	
<b>ACELERADORAS</b>	<p>¿Se proveen <b>mecanismos de aprendizaje y educación para el desarrollo de capacidades de los empleados</b> conducentes a la adopción y mantenimiento de la IA?</p>		

## 2.2 GESTIÓN DE RIESGOS Y CONTROLES INTERNOS

Los sistemas de gobernanza de la IA deben estar basados en un enfoque de gestión de riesgos que permita identificar de manera temprana posibles dificultades para el sistema y mitigar los daños no intencionales que pudieran haber resultado de un mal uso de la IA. Por lo anterior, se sugiere que los sistemas de gestión de riesgos estén basados en modelos de protección de la privacidad por defecto (EC 2018; Galdon 2020; Omidyar 2019, Hoepman 2020), a fin

de que los datos personales usados en el sistema de IA estén protegidos en todo momento y que el usuario tenga el control sobre sus datos.

Por su parte, los controles internos deben describir cómo se lleva a cabo la supervisión del procesamiento de datos, incluyendo detalles sobre quién procesa, quién controla, y con quién se comparten los datos.



## EMPRESAS

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
	<p>¿Existen <b>mecanismos de gestión de riesgos</b> que puedan ser expandidos para incluir a los sistemas de IA?</p> <p>En el caso de hacer uso de datos masivos y tecnologías invasivas, entre otras, ¿se ha contemplado la posibilidad de llevar a cabo <b>evaluaciones de impacto en protección de datos</b> (<i>EIPD</i> o <i>PIA</i> en inglés <i>Privacy Impact Assessment</i>)?</p> <p>¿El sistema de IA cuenta con un <b>sistema de gestión de datos</b> en el que el usuario pueda tener un nivel de gestión mínimo sobre su propia información personal?</p> <p>¿Se cuenta con mecanismos de <b>seguridad para los elementos críticos</b> del sistema de IA durante todo el periodo operacional del sistema?</p> <p>¿La <b>junta directiva de la empresa supervisa y aprueba</b> las metodologías de gestión de riesgos?</p>	<p>¿Se han realizado evaluaciones dentro de la organización que permitan <b>identificar riesgos al personal o a las estrategias de negocios</b>?</p> <p>¿Existe una <b>evaluación periódica sobre posibles riesgos y espacios de discusión</b> para tratar preocupaciones relativas a los algoritmos?</p>	<p>¿La junta directiva de la empresa ha adoptado alguna <b>herramienta que permita identificar riesgos clave de acuerdo con el nivel de madurez tecnológica</b> de la empresa, tales como las emitidas por el <i>Instituto Nacional de Estándares y Tecnología (NIST)</i>, por sus siglas en inglés) u otros organismos internacionales de estandarización en torno a la ciberseguridad?</p> <p>¿Se llevan a cabo evaluaciones regulares y <b>actualizaciones de los sistemas de gestión de riesgos</b>?</p> <p>¿Se cuenta con una estrategia de <b>gestión de conocimiento</b> que permita tener un seguimiento y control sobre los procesos del sistema de IA?</p>
	<p>¿La empresa cuenta con mecanismos de <b>gestión de riesgos basados en una política de privacidad por defecto</b>?</p> <p>¿La junta directiva de la empresa participa de forma regular en las <b>valoraciones sobre el sistema de gestión de riesgos y controles internos</b> para asegurar la integridad del sistema de IA?</p>		

## 2.2.1 Seguridad de los datos

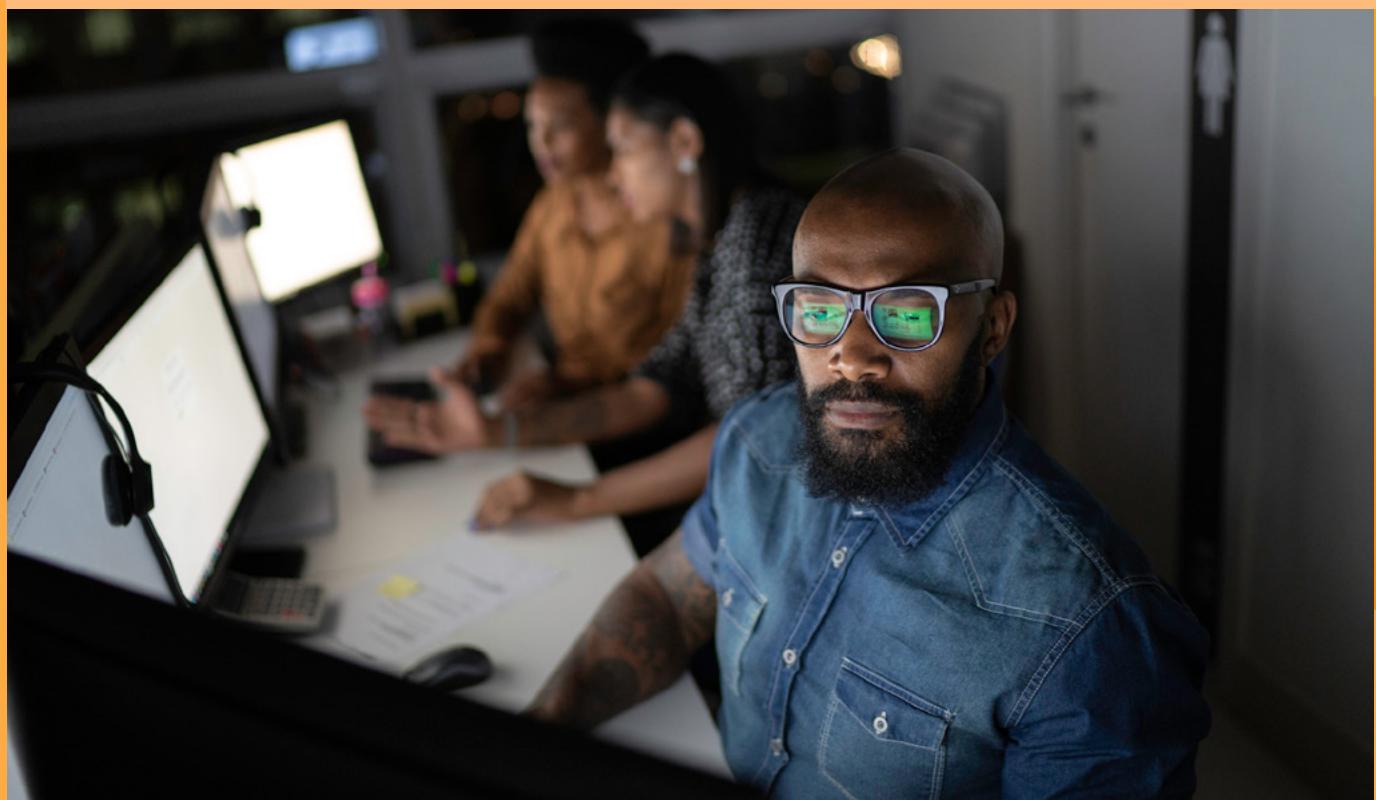
La seguridad de datos es uno de los elementos más importantes dentro de las evaluaciones de un sistema de IA, debido a que se trata de los insumos sobre los cuales opera el sistema además del potencial contenido privado por parte de los usuarios. De este modo, resulta importante valorar las acciones conducentes a la protección técnica y organizacional del sistema de IA ante vulnerabilidades y agentes maliciosos

que puedan alterar los datos, causar fallas en los modelos de la IA, así como su infraestructura crítica en materia de software y hardware. En la gran mayoría de los casos, para que se consideren seguros, los desarrolladores de los sistemas de IA deberían contemplar explícitamente aquellos casos en los que se pudieran presentar potenciales abusos por parte de agentes malintencionados (EC 2018; Google 2018; Omidyar 2019; Méndez 2020; WEF & IMDA 2020).



	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se han realizado <b>evaluaciones sobre las vulnerabilidades</b> del sistema de IA?</p> <p>¿Se ha considerado el <b>nivel de riesgo del sistema de IA para cada uso específico</b>?</p> <p>¿Se ha considerado <b>hasta qué punto el sistema puede ser utilizado para propósitos diferentes a los establecidos por la organización</b>? Existen numerosos ejemplos de sistemas de IA utilizados en campañas electorales que han tenido repercusiones importantes en los sistemas políticos y democráticos.</p>	<p>¿Se han establecido umbrales de aceptabilidad y mecanismos de gobernanza de la IA que sean accionables ante una posible eventualidad?</p> <p>Para el caso de <b>modelos de rastreo y proximidad de los usuarios</b>, ¿se ha considerado el uso de protocolos descentralizados para la protección de datos tales como DP-3T o equivalentes?</p>	<p>¿Se ha valorado los posibles <b>riesgos del uso de la IA a terceros, incluyendo el daño potencial, caracterización de la audiencia afectada y la severidad del daño</b>?</p> <p>¿La organización cuenta con protocolos de <b>emergencia ante potenciales ataques ciberneticos y pérdida de datos, entre otros</b>, tales como cambios técnicos inmediatos o intervención de seres humanos para reducir riesgos?</p> <p>¿Se han llevado a cabo <b>simulaciones de ciberataques</b> en los que hayan puesto a prueba los <i>mecanismos de intervención humana o desconexión total o parcial del sistema</i>?</p>
<b>INVERSIONISTAS</b>	<p>¿Qué tipo de <b>modelos de seguridad de datos</b> se utilizan, tales como protocolos centralizados o descentralizados?</p>	<p>¿Se han valorado los <b>riesgos ante posibles multas o sanciones por incumplimiento legal en materia de seguridad de datos</b> que pudieran comprometer seriamente las operaciones y rentabilidad del negocio?</p>	<p>¿Se considera alguna <b>cláusula de cobertura ante algún posible daño causado</b> por el sistema de la IA?</p> <p>En caso de ser necesario ¿la empresa ha considerado la <b>contratación de seguros de cobertura contra posibles daños causados por el sistema de la IA</b>?</p> <p>¿El proyecto toma en consideración las <b>responsabilidades atribuibles en materia de regulación de protección al consumidor</b>?</p>

# 3



## INVOLUCRAMIENTO HUMANO EN LOS SISTEMAS DE IA

# 3

## INVOLUCRAMIENTO HUMANO EN LOS SISTEMAS DE IA

**Uno de los elementos más relevantes para el uso ético y responsable de la IA se relaciona con la identificación de las tareas que requieren supervisión, intervención o interacción humana, ya sea para corregir imprecisiones del sistema o generar una interacción adecuada entre usuarios y dicho sistema (Google 2018; IMDA & PDPC 2020).**

**Del mismo modo, y desde la posición de un enfoque de prevención de riesgos, el grado de involucramiento humano adecuado en los sistemas de IA debería potenciar la autonomía humana al tiempo que se minimizan posibles condicionamientos, engaños o subordinación injustificada a quienes utilizan esta tecnología.**

A continuación, se comparte una tabla de clasificación de la IA por tipo de uso e interacción con seres humanos elaborada por la División de Sector Social del BID,

que permite valorar la correspondencia del nivel de involucramiento humano según el tipo de proceso de toma de decisión:

	HUMANO INVOLUCRADO	SIN HUMANOS INVOLUCRADOS
OBJETIVO ESPECÍFICO	Inteligencia asistida	Inteligencia automatizada
SISTEMAS ADAPTATIVOS	Inteligencia aumentada	Inteligencia autónoma

## 3.1 INTERACCIÓN SERES HUMANOS-IA: DETERMINACIÓN DEL NIVEL DE SUPERVISIÓN HUMANA

La determinación del nivel de involucramiento humano en los procesos, resultados y toma de decisiones del sistema de IA dependerá en gran medida del impacto que tengan dichos procesos en los usuarios. El ejemplo más común utilizado para tener una idea de la gradualidad del impacto de los sistemas de IA en su entorno es la comparación entre un sistema de IA que se limita a emitir sugerencias de

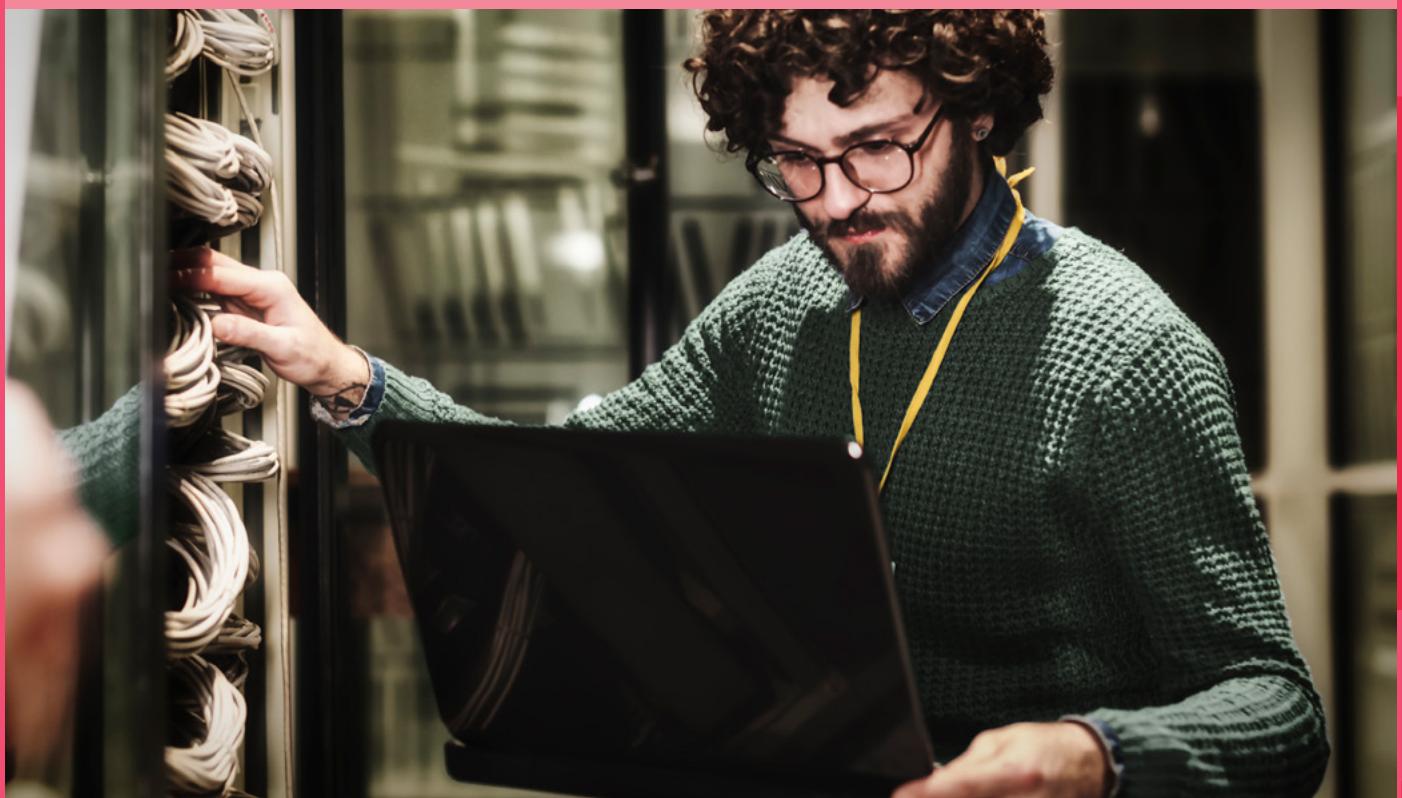
compra a los usuarios con base en su comportamiento y preferencias, versus un sistema de IA que tiene implicaciones para la seguridad vial de los usuarios (MGI 2017; EC 2018). Es evidente que ambos sistemas requieren de diferentes grados de supervisión e intervención humana, a fin de prevenir o corregir posibles resultados adversos esperados o inesperados del sistema (Google 2018).



## EMPRESAS

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
	<p>¿Se ha considerado el uso de herramientas o modelos para <b>determinar el nivel de involucramiento humano con base en el impacto del sistema en la vida de los usuarios?</b> Por ejemplo, a través de metodologías como <i>human in/out/over the loop</i>.</p> <p>¿Se ha considerado el <b>costo operativo de los diferentes niveles de involucramiento humano</b> en el sistema de IA, al destinar personal especializado dedicado a estas labores?</p> <p>Dependiendo del grado de impacto del sistema en la vida de los usuarios, ¿el sistema cuenta con una opción de <b>detención inmediata para la intervención humana?</b></p>	<p>¿Se ha seleccionado el <b>nivel apropiado de control humano</b> de acuerdo con el sistema de IA y casos de uso pertinentes?</p> <p>¿Las actividades o procesos en los que la IA va a operar <b>requieren del reconocimiento de matices emocionales</b> o simplemente de la <b>ejecución trata de actividades repetitivas?</b></p> <p>¿Se ha considerado la elaboración de <b>indicadores clave de desempeño (KPIs por sus siglas en inglés)</b> del sistema de IA que ayuden a determinar la pertinencia del involucramiento humano <b>en caso de que el sistema se deteriore?</b></p>	<p>En caso de la utilización de mecanismos autónomos de IA y de aprendizaje automático, ¿existen <b>mecanismos de control más específicos</b> sobre problemas ya identificados?</p> <p>¿Se ha considerado la elaboración de <b>respuestas de emergencia basadas en el análisis de escenarios</b> ante la posibilidad de que las medidas de mitigación de riesgo fallen?</p> <p>En caso de que el sistema de IA tenga la capacidad de entender o sentir a sus interlocutores, ¿qué <b>medidas se han tomado para evitar problemas de empatía o apego de los usuarios?</b></p>

# 4



## CICLO DE VIDA DE LA IA: GESTIÓN DE OPERACIONES DEL SISTEMA DE IA

# 4

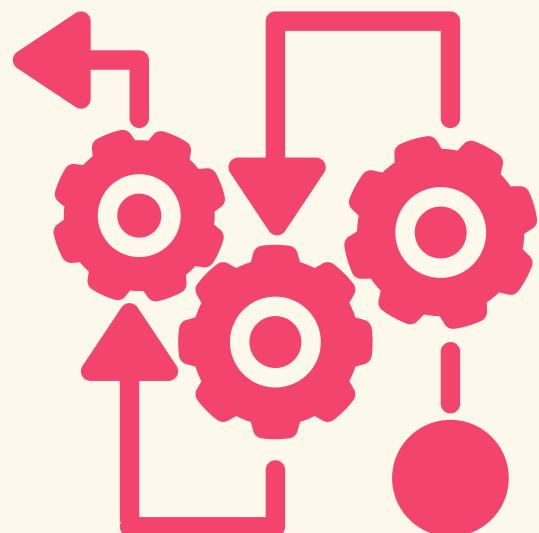
# CICLO DE VIDA DE LA IA: GESTIÓN DE OPERACIONES DEL SISTEMA DE IA

**Esta sección está enfocada al análisis de las operaciones del sistema de IA desde la recolección y estructuración de datos desde una perspectiva de privacidad por diseño o por defecto, hasta el procesamiento de estos datos a través de los algoritmos de IA.**

## 4.1 FUENTE Y MANEJO DE DATOS

El insumo base de los sistemas de IA son los datos. La mayoría de sistemas basados en el aprendizaje automático (a diferencia de los sistemas basados en reglas) requieren de una gran cantidad de datos históricos y de alta calidad a fin de perfeccionar su comportamiento y los resultados tanto en materia de predicción, como de toma de decisiones automáticas. En este sentido, la calidad de los datos es de gran importancia para el uso responsable de la IA, pues es una forma de garantizar la correcta operatividad de dicho sistema de modo que no tenga sesgos o problemas que afecten a los usuarios (Posey 2019; WEF & IMDA 220). Por otro lado, muchos de los datos que utilizan los sistemas de IA suelen contener información personal o metadatos relacionados con el comportamiento de sus usuarios, de

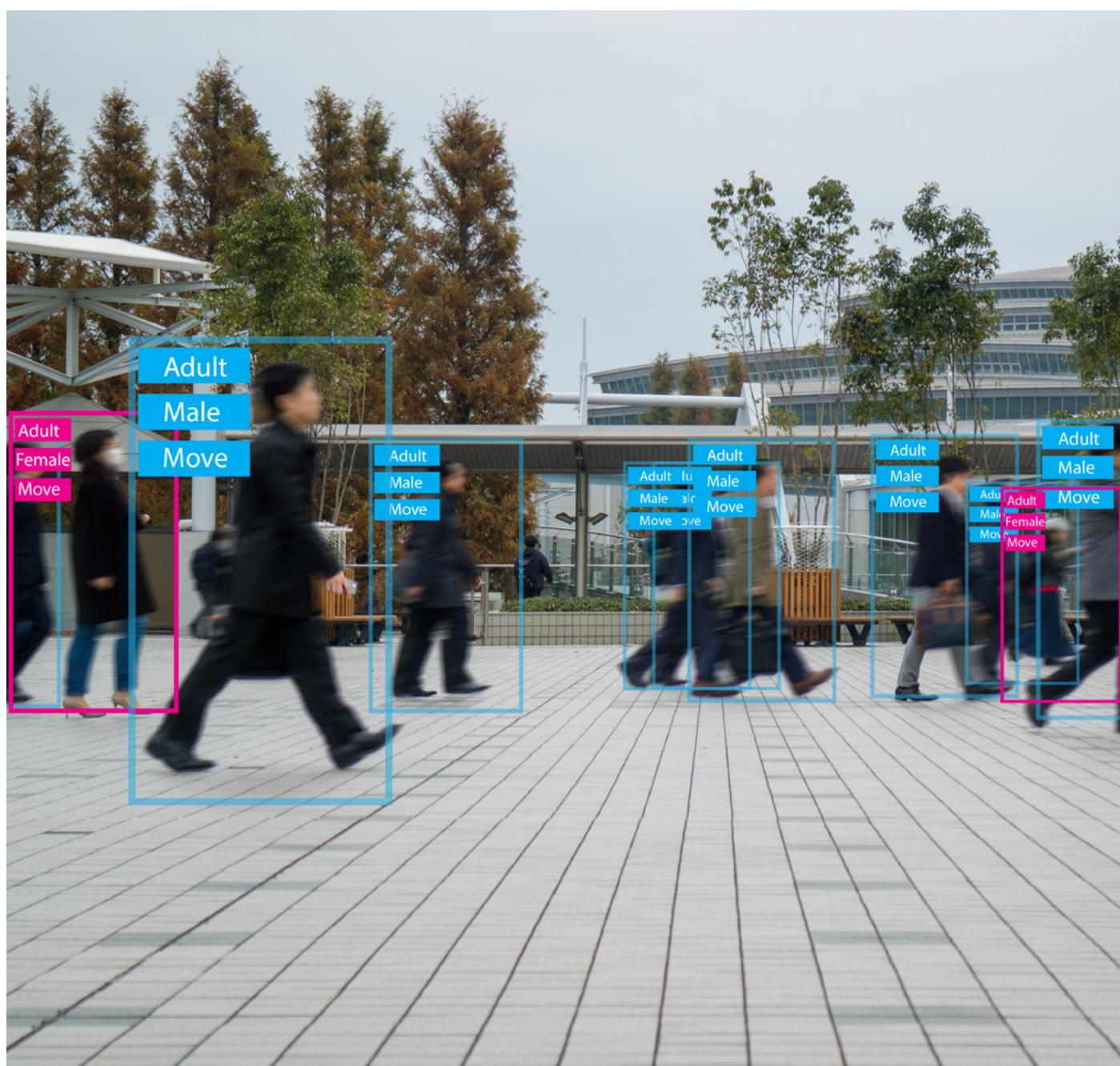
modo que la protección de estos datos es otro elemento crítico para el uso ético y responsable de la IA.



#### 4.1.1 Tipo de datos

Antes del despliegue o la adopción de un sistema de IA conviene realizar una valoración sobre el tipo de datos que se requieren para el correcto funcionamiento del sistema, así como el tipo de datos con los que se dispone para su alimentación. Actualmente los datos son un elemento que ha adquirido un valor comercial cada vez mayor, de modo que su intercambio requiere de un cuidadoso tratamiento a fin de no violar diversos aspectos de la privacidad de

los usuarios (MGI 2017). Por otro lado, en los últimos años se ha generado una dinámica de competitividad por datos que los usuarios suelen proteger cada vez más, tales como su geolocalización y opiniones o sentimiento en torno a diferentes productos, servicios o temas, de modo que hay una conciencia cada vez mayor sobre el cuidado de estos elementos que deberían ser considerados por aquellas empresas interesadas en el uso de este tipo de datos.



	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se ha valorado con precisión qué <b>tipo de datos va a requerir el sistema de IA</b>, incluyendo su <b>disponibilidad, acceso a fuentes y formatos</b> necesarios?</p> <p>¿Existen medidas que permitan <b>rastrear el origen y destino</b> de los datos utilizados por el sistema de IA?</p> <p>¿Los datos son completos, están <b>actualizados y provienen de fuentes confiables</b>?</p> <p>¿Se ha considerado el uso de <b>métodos para la extracción de datos</b> tales como ETL (<i>extraer, transformar, cargar</i>) para garantizar la compatibilidad de datos y formatos?</p>	<p>¿Se han realizado evaluaciones sobre el <b>control de calidad</b> sobre el uso de <b>datos provenientes de fuentes externas</b>?</p> <p>¿Se cuenta con <b>inventarios o repositorios de datos, taxonomías, y documentación</b> de mecanismos de control sobre los datos que estén basados en los principios de <i>gobernanza de datos de agencias de normalización como ISO o IEEE</i>?</p> <p>¿Se cuenta con mecanismos de monitoreo que permitan <b>valorar si los cambios en la fuente de origen generan cambios sustanciales</b> en los resultados del sistema de IA?</p>	<p>¿El sistema de IA utiliza herramientas de software especializado para el <b>almacenamiento, recuperación o transformación de datos</b> que permitan garantizar su consistencia? Algunos ejemplos son: <i>NoSQL</i> o el modelo <i>MapReduce</i> usado por <i>Apache Hadoop</i>?</p> <p>¿Se cuenta con un equipo destinado a la <b>gestión de la política de datos</b>?</p> <p>¿La empresa tiene la capacidad de <b>verificar la validez y representatividad de los datos</b> obtenidos a través de terceros (incluyendo estrategias de verificación a través de metadatos)?</p>
<b>INVERSIONISTAS</b>	<p>¿Cuál es la <b>fuente de los datos</b> que utiliza el sistema de IA y cuál es la estrategia de datos?</p> <p>¿La empresa cuenta con su <b>propia fuente generadora de datos (base de datos única)</b> o depende de terceros para su obtención?</p>	<p>¿Existen <b>convenios de asociación</b> con otras empresas a fin de garantizar el mantenimiento de <b>bases de datos únicas no accesibles a competidores</b>?</p>	<p>¿La empresa utiliza <b>datos históricos, datos estáticos o datos en tiempo real</b>?</p> <p>¿La empresa tiene la <b>capacidad de procesar datos de entrenamiento</b> para hacer más eficientes sus algoritmos?</p>
<b>ACELERADORAS</b>	<p>¿Hay disponibilidad de expertos técnicos que puedan ofrecer orientación sobre <b>arquitectura de datos, recolección, almacenamiento y transformación de datos</b>?</p> <p>¿Se ha valorado cuáles son los <b>instrumentos legales que permitan procesar los datos</b> para el sistema de IA?</p>		

#### 4.1.2 Procesamiento de datos

El procesamiento de datos se refiere a la recolección, almacenamiento, organización, estructuración, alteración, consulta, uso, transmisión, diseminación y cualquier operación ejecutada en estos datos (Omidyar 2019; WEF & IMDA 2020). Conviene recordar que diversos instrumentos internacionales prohíben en principio el procesamiento de datos que revelen el origen étnico o racial de las personas, opiniones políticas, creencias religiosas o filosóficas, afiliación o pertenencia a sindicatos laborales, datos genéticos y biométricos, preferencia u orientación sexual, así como estado de salud, excepto si se cumplen condiciones en ciertos casos,

como, entre otras, el consentimiento explícito del usuario (EC 2018; Reglamento General de Protección de Datos de la Unión Europea 2016-2018). Por lo anterior, se recomienda que se consulten los instrumentos y regulaciones apropiadas según normativas internacionales y nacionales. Las preguntas orientadoras que se presentan a continuación están basadas en su mayoría en las disposiciones del Reglamento General de Protección de Datos de la Unión Europea, como uno de los instrumentos más avanzados en dicha materia y que podrían constituir insumos para la elaboración de regulaciones en diversos países.

EMPRESAS	ETAPA I	ETAPA II	ETAPA III
	<p>¿El proceso de <b>recolección de datos forma parte de los protocolos de minimización de datos</b> que recolecten únicamente la información personal de los usuarios solo cuando sea necesario y durante un tiempo limitado bajo condiciones de almacenamiento seguro?</p> <p>¿Se han establecido <b>protocolos para el procesamiento de datos</b> con respecto al almacenamiento, divulgación, y transformación de procesamiento, así como para su tratamiento por terceros?</p> <p>¿Se ha establecido una <b>política de consentimiento informado y divulgado</b> sobre el uso de datos de las personas?</p>	<p>En caso de que ocurra una <b>edición manual de datos</b>, ¿existen los protocolos suficientes para garantizar la calidad, auditabilidad y trazabilidad de los datos modificados?</p> <p>¿Existen provisiones que permitan <b>salvaguardar la privacidad</b> de la información personal que utiliza el sistema de IA, tales como agregación y anonimato de datos?</p> <p>¿Se han contemplado las <b>regulaciones en torno a las políticas de seguridad para la limitación de almacenamiento de datos personales</b>, incluyendo la protección contra el uso no autorizado, pérdida accidental, destrucción o daño de datos?</p>	<p>¿Qué medidas toma la organización para <b>mitigar el sesgo originado a causa del reforzamiento de patrones</b>?</p> <p>¿Se ha considerado la elaboración de un <b>modelo de IA adversario al sistema original</b> para ponerlo a prueba y valorar la confiabilidad de los resultados?</p> <p>¿La organización <b>fracciona bases de datos voluminosas con el fin de mitigar riesgos</b> y validar el modelo de IA desde varias perspectivas?</p>

#### 4.1.3 Integridad y confidencialidad de datos

La valoración sobre la integridad de los datos está estrechamente vinculada a la calidad de los mismos en toda la cadena de uso dentro del sistema de IA, pues la introducción de datos incompletos o mal estructurados podrían tener repercusiones importantes en el desempeño del sistema. Los diversos conjuntos de datos de los que se alimenta el sistema deberían ponerse

a prueba de forma regular y estar bien documentados tanto si los datos fueron desarrollados de forma interna o si se adquirieron por medio de un tercero (Reglamento General de Protección de Datos de la Unión Europea 2016-2018; EC 2018). Además de utilizar las preguntas guía a modo de orientación sobre este tema, se recomienda revisar los principios de privacidad de la OCDE, así como consultar las normas respectivas de ISO e IEEE.

EMPRESAS	ETAPA I	ETAPA II	ETAPA III
	<p>¿Se han establecido procedimientos para <b>garantizar la calidad e integridad de datos</b>?</p> <p>¿Se reconocen las posibles <b>limitaciones originadas por la composición de los datos</b> usados?</p> <p>¿Se consideran amplios <b>márgenes de diversidad y representatividad de datos</b> con respecto a los usuarios?</p> <p>¿Existe la <b>posibilidad de que el usuario opte por eliminar su historial y registro</b> del sistema?</p>	<p>¿Se ha considerado la implementación de <b>tests que valoren la integridad de los datos</b>? Por ejemplo: <i>Database Integrity Testing por servicio, área, tecnología o disciplina</i>.</p> <p>¿Cómo se <b>verifica que las bases de datos mantengan un estado consistente o que no sean accedidas por terceros de forma indebida</b>?</p> <p>¿Se han realizado <b>pruebas a la diversidad y representatividad de datos con respecto a poblaciones particulares</b> que hayan reportado casos de uso problemáticos?</p>	<p>¿Se utilizan métodos de investigación y herramientas <b>técnicas para mejorar el entendimiento de los datos</b>, su dinamismo, el modelo de la IA y su rendimiento?</p> <p>¿Se ha considerado la <b>obtención de certificación por parte de agencias independientes en torno a la seguridad y privacidad de datos</b>?</p>

#### 4.1.4 Privacidad mediante diseño

El manejo a datos personales de los usuarios para el uso de sistemas de IA debería estar regulado mediante protocolos de acceso a estos datos que determinen con precisión quién y bajo

qué circunstancias se puede acceder a dicha información. El personal de la empresa que acceda a ello debe estar debidamente acreditado y calificado para su tratamiento (KPMG 2018; Galdon 2020; Hoepman 2020).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>En caso de contar con bases de datos que incluyan información personal de los usuarios, ¿qué <b>mecanismos de protección de esta información</b> existen?</p> <p>¿Se ha considerado el uso de <b>modelos de IA que minimicen la dependencia del uso de datos personales</b>?</p> <p>¿Se han establecido <b>mecanismos de consentimiento para el uso de datos personales</b> y la posibilidad de retirar los datos si se solicita?</p> <p>¿Se ha considerado en el diseño de procesamiento y archivo de datos que los datos personales <b>no se retengan más del tiempo necesario para su uso</b>?</p> <p>¿Se llevan a cabo <b>evaluaciones de impacto en la protección de datos personales</b> (<i>PIA: Privacy Impact Assessment</i>) sobre los productos y servicios vinculados con el sistema de IA, incluyendo al personal?</p>	<p>¿Se han tomado medidas de protección de datos como <b>encriptación, anonimización y agregación</b>?</p> <p>¿Se siguen <b>protocolos de protección de datos</b> provenientes de <b>estándares internacionales como normas ISO o IEEE</b>?</p> <p>¿Se tiene un <b>registro del acceso a los datos personales</b>? Por ejemplo: quién accede, cuándo, desde qué ubicación.</p>	<p>En caso de la existencia de personal destinado a la supervisión de datos personales, ¿se ha <b>involucrado a esta persona en el diseño del sistema desde el inicio</b>?</p> <p>¿Existen mecanismos de <b>divulgación de la política de protección de datos</b> que aseguren su conocimiento por parte de los empleados de la empresa y terceros?</p>
<b>INVERSIONISTAS</b>	<p>Las entidades receptoras de financiamiento o inversión deberían <b>conocer y seguir las políticas de privacidad</b> de datos de sus contrapartes.</p>		

#### 4.1.5 Interoperabilidad

La interoperabilidad se refiere a la capacidad de un sistema de compatibilizar procesos y fuentes de datos con otros proveedores o clientes (Galdon 2020), a través del uso de formatos similares que permitan el correcto funcionamiento de los sistemas de IA, así como de sintaxis similares entre diversas bases de datos.

Sin embargo, en algunas ocasiones se suele omitir esta dimensión y asumir que los datos provenientes de otras instituciones se encuentran en formatos compatibles y con contenidos estructurados del mismo modo. Sin embargo, esto suele ser un gran reto para el funcionamiento de muchos sistemas, sobre todo la compatibilidad sintáctica (Potgieter 2018).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se ha realizado un <b>mapeo de las fuentes con las que se requiere integrar el sistema de IA</b>, incluyendo tipo de formatos utilizados y estructura legible de los datos?</p> <p>¿Quiénes son las contrapartes (<b>sector público o privado</b>) con las que se va a realizar la integración operativa?</p> <p>¿Se requiere de la <b>celebración de contratos</b> para la integración operativa?</p> <p>¿Se ha considerado la <b>obtención de datos por medio de proveedores que cumplan con estándares de protección de datos</b>?</p>	<p>¿Se conocen detalles como la ubicación física de servidores y bases de datos?</p> <p>¿Qué tipo de <b>rediseño de procesos se requiere para integrar bases de datos externas</b> dentro del flujo de trabajo?</p> <p>¿Existe la posibilidad de que el <b>usuario obtenga su registro de datos de modo que lo pueda portar y usar con otro sistema de IA</b>?</p>	<p>¿Se han adoptado medidas específicas para compatibilizar la <b>sintaxis y semántica de las fuentes de datos</b> con los requerimientos específicos de cada sistema de IA? Por ejemplo, en el sector de la salud a través de recursos similares al <i>Fast Healthcare Interoperability Resources (FHIR)</i> en Estados Unidos para el sector de la salud.</p>
<b>INVERSIONISTAS</b>	<p>¿Existe un <b>historial previo de integración operativa</b> de la empresa con las contrapartes propuestas que incluya contactos, contratos o algún otro vínculo?</p>	<p>¿Es <b>sostenible la relación de interoperabilidad</b> entre la empresa y sus contrapartes?</p>	
<b>ACELERADORAS</b>	<p>¿Qué <b>iniciativas análogas de interoperabilidad</b> pueden ser desarrolladas y utilizadas? Por ejemplo, iniciativas como smart cities o data lakes.</p>		

#### 4.1.6 Ensayos y validación

Los procesos de ensayo y validación permiten identificar y resolver posibles errores del sistema de IA dentro de un entorno estable. Estos ensayos requieren

de bases de datos precisas y lo más cercanas a la realidad, a fin de garantizar el correcto funcionamiento del sistema de IA (Galdon 2020; Omidyar 2019).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>Durante el diseño de los ensayos, ¿se determinan los <b>parámetros básicos para la realización de validaciones que permitan conocer el grado de fiabilidad</b> del sistema de IA?</p> <p>¿Se realizan <b>ensayos y validaciones para cada uno de los subgrupos de la población</b> objetivo del sistema de IA?</p> <p>¿Los ensayos realizados <b>se corresponden con las expectativas de operatividad</b> en el ambiente en el que se utilizará el sistema de IA?</p>	<p>¿Se cuenta con <b>diferentes bases de datos</b> para llevar a cabo actividades diferenciadas de entrenamiento al algoritmo, prueba y validación?</p> <p>¿Se llevan a cabo <b>test de reproducibilidad en ambientes de producción a escala</b>?</p> <p>¿Se ha considerado la <b>creación de un modelo de ensayo</b> que pueda contar con la <b>retroalimentación de los usuarios</b>?</p>	<p>¿Se llevan a cabo <b>simulaciones sobre la recolección de datos de diferentes fuentes</b>, a fin de asegurar la calidad de datos y verificar la validez del modelo con el mundo real?</p> <p>¿Se llevan a cabo pruebas con métodos como A/B sobre <b>distintas versiones del modelo</b> de IA para comprobar su fiabilidad?</p>
<b>INVERSIONISTAS</b>	<p>¿Los <b>resultados de los ensayos se comparten</b> con los entes financieros?</p>		

## 4.2 DESARROLLO DE MODELOS: ALGORITMOS

Los algoritmos son la estructura de los sistemas de IA, pues son los encargados de procesar los datos para realizar diversas funciones, entre las que se incluyen predicciones, clusterizaciones y categorizaciones, así como toma de decisiones automatizadas (MGI 2017; Google 2018; IMDA & IDPC 2020). Existen diversos tipos de algoritmos basados en reglas o en el aprendizaje a través de datos masivos. Su clasificación también puede incluir el nivel de supervisión o instrucción humana que requieran. Por su complejidad conceptual y alto nivel de abstracción matemática, los algoritmos suelen ser explicados por el tipo de resultado que generan, más que por la forma en que actúan sus procesos internos (Google 2018; IMDA & IDPC 2020).

Con base en lo anterior, la construcción de confianza en torno a la IA pasa necesariamente por el escrutinio de los

diferentes modelos de algoritmo utilizados, así como por sus resultados. Esta sección es una guía temática sobre los temas más relevantes en torno al uso ético y responsable de los algoritmos.

### 4.2.1 Trazabilidad

La trazabilidad de los algoritmos utilizados en la IA se refiere a la documentación de los insumos y procesos del sistema de IA que conducen a determinados resultados. Este registro es importante, ya que es uno de los elementos que contribuyen en la transparencia y generación de confianza en torno a la adopción de la IA. Además, permite identificar de manera precisa cualquier problema que se pueda presentar a lo largo de las operaciones complejas desde la recolección de datos, su estructuración, procesamiento, interpretación y resultado (EC 2018; Google 2018; WEF & IMDA 2020).



EMPRESAS	ETAPA I	ETAPA II	ETAPA III
	<p>¿Existen formas de <b>documentar</b> los siguientes elementos?: <i>Objetivos del sistema, datos, aproximación de la investigación, descripción del algoritmo y sus parámetros de evaluación de desempeño, errores y notas del equipo técnico.</i></p> <p>¿Se cuenta con registros sobre el proceso de <b>obtención de datos provenientes de fuentes abiertas?</b></p> <p>¿Se han implementado métodos de <b>documentación</b> y seguimiento a los procesos de IA basados en <b>reglas preestablecidas o autoaprendizaje?</b></p>	<p>Para sistemas de IA basados en reglas preestablecidas, ¿se ha <b>documentado el proceso de construcción del modelo?</b></p> <p>Para sistemas de IA basados en el autoaprendizaje, ¿se tiene un sistema de <b>rastreo de las fuentes de datos que explique el proceso de selección y cómo ha ocurrido?</b></p>	<p>En cuanto a los métodos de prueba y validación de algoritmos:</p> <p>Para los sistemas basados en reglas preestablecidas, ¿se tiene información <b>sobre los datos usados para las pruebas y validaciones?</b></p> <p>Para los sistemas basados en aprendizaje automático, ¿se <b>cuenta con el rastreo de los datos de diversas fuentes y los resultados obtenidos en diversas simulaciones?</b></p>

#### 4.2.2 Explicabilidad de los procesos y resultados de la IA

Se trata de la capacidad de explicar tanto los procesos técnicos de un sistema de IA, como de las decisiones que resultan de la utilización de dicho sistema. Una explicación efectiva es aquella que es comprensible para las diferentes audiencias a quienes se expone. A medida que los sistemas de IA tienen mayor impacto en la vida de las

personas, mayor será su obligación de explicar los procesos y resultados del uso de la IA. Asimismo, se trata de una obligación reglamentaria que ha surgido de mecanismos como el Reglamento General de Protección de Datos de la Unión Europea y que ha sido incorporada en diversas legislaciones en América Latina y el Caribe (Google 2018; Dassen et al. 2019; Omidyar 2019; Molnar 2020; WEF & IMDA 2020).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Hasta qué punto se puede <b>entender de forma sencilla los procesos y resultados</b> del sistema de IA?</p> <p>¿Se puede explicar <b>por qué el sistema realizó determinada predicción o tomó cierta decisión</b> de modo que todos los usuarios lo puedan entender?</p> <p>¿Se ha diseñado el modelo de IA tomando en cuenta distintos <b>factores de interpretación</b> incluyendo variedades del lenguaje?</p>	<p>En el caso de que los resultados de la operación del sistema de IA tenga un impacto en la toma de decisiones de la empresa, ¿<b>se comunica de forma oportuna</b> al resto de empleados y socios afectados por las decisiones?</p> <p>¿Se cuenta con <b>diferentes modelos de explicación correspondientes con las diversas características o acciones</b> del modelo de IA?</p> <p>¿Se ha considerado la <b>construcción de narrativas de explicación simples</b> a través de parámetros de medición del tipo: alto/medio/bajo, porcentajes u otras herramientas que ayuden a la comprensión del sistema?</p>	<p>¿Se cuenta con un <b>documento que explique al público cómo funciona el sistema de IA, incluyendo parámetros de medición de desempeño y mantenimiento</b>?</p> <p>¿Se cuenta con un <b>modelo más sencillo del sistema de IA creado a efectos de proveer explicaciones</b> didácticas?</p> <p>¿Se ha considerado el uso de <b>técnicas especializadas para la explicación de los resultados de modelos</b> de IA tales como <i>LIME</i>, <i>SHAP</i>, <i>LOCO</i>, <i>PDP</i>, <i>ICE</i>?</p> <p>¿La empresa cuenta con <b>infraestructura suplementaria para ofrecer explicaciones de forma automática y proactiva a los usuarios</b> de acuerdo con los diferentes resultados de predicción o toma de decisiones generados por el sistema de IA?</p>
<b>INVERSIÓNISTAS</b>	<p>¿Cuál es el modelo de negocio de la empresa y <b>cómo genera valor a través del uso de la IA</b>?</p>	<p>¿Se ha valorado hasta qué punto se puede <b>examinar la interpretabilidad del modelo de IA</b> y acceder a los procesos de trabajo del modelo para comprenderlos mejor?</p>	
<b>ACCELERADORAS</b>	<p>¿Existe la obligación de explicar los procesos y resultados del sistema de IA? Y <b>¿se cuenta con los mecanismos que permitan esta explicación?</b></p>		

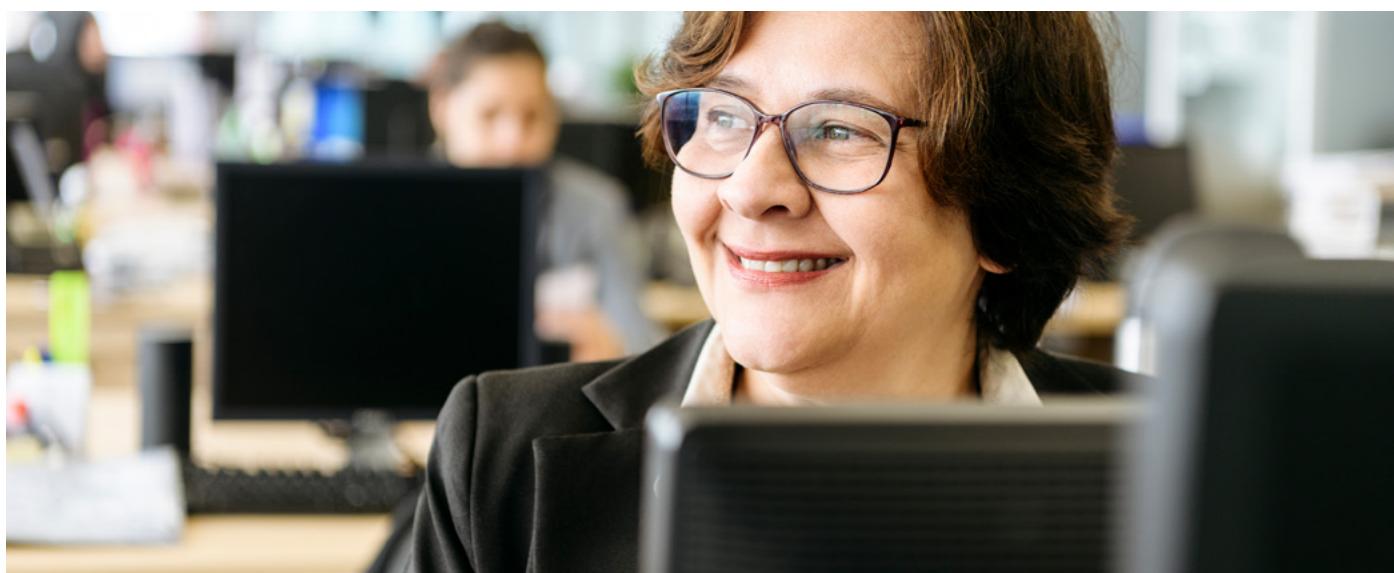
#### 4.2.3 Replicabilidad

La replicabilidad consiste en la ejecución repetida de una operación que genere un mismo resultado de forma consistente. Es por esta razón que algunos modelos de IA sustituyen ciertas medidas de explicabilidad por la demostración de la consistencia en los resultados de un determinado modelo, a través de repeticiones controladas. Esto ocurre cuando los procesos del modelo de IA suelen ser muy complejos o abstractos de un modo tal que el entendimiento sobre el sistema resulta más sencillo y útil

para comunicar el grado de fiabilidad del sistema (EC 2018; Google 2018; Omidyar 2019; WEF & IMDA 2020).

Además, en algunas ocasiones es conveniente llevar a cabo pruebas de replicabilidad para valorar de modo *contrafactual* qué tan “justo” es un sistema con respecto a diversos grupos poblacionales, a fin de evitar la perpetuación de sesgos, sobre todo en sectores como seguros, préstamos, contrataciones y procuración de justicia de forma predictiva (Kusner et al. 2017).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Existe la necesidad de llevar a cabo <b>test de replicabilidad y pruebas contrafactuales</b> para asegurar que determinados grupos poblacionales no serán afectados de forma negativa por las decisiones automatizadas del sistema de IA?</p>	<p>En caso de usar medidas de replicabilidad, <b>¿se han identificado resultados excepcionales</b> y se ha realizado una valoración de cómo abordarlos?</p> <p>En caso de realizar ensayos de replicabilidad, <b>¿hay datos actualizados para cada prueba?</b></p>	<p>¿Se llevan a cabo <b>ensayos de replicabilidad comparativos entre diferentes versiones del mismo modelo de IA?</b></p>



#### 4.2.4 Reproducibilidad

El concepto de reproducibilidad se refiere a la capacidad que tiene el modelo de IA de ser reproducido en otro ambiente externo al de la organización que lo creó, a diferencia del concepto de replicabilidad, el cual hace referencia

a la posibilidad de obtener los mismos resultados dentro de un mismo ambiente. Por definición, este concepto requiere de dos entornos en los cuales se lleven a cabo pruebas para verificar los resultados del modelo (Google 2018; Omidyar 2019; WEF & IMDA 2020).

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se ha determinado qué <b>condiciones son necesarias para garantizar la reproducibilidad del sistema?</b></p> <p>¿Se ha entrado en contacto con alguna <b>organización externa con la que se podrían llevar a cabo actividades de reproducibilidad</b> del sistema de IA?</p>	<p>¿Se ha destinado a un <b>equipo especializado para revisar si se pueden producir resultados similares</b> utilizando los mismos modelos de IA en diferentes entornos?</p> <p>¿Se ha <b>documentado y operacionalizado los procesos de prueba y verificación</b> de la confiabilidad de los sistemas de IA?</p>	<p>¿Se ha considerado la <b>elaboración de archivos de reproducibilidad</b> para fases de prueba de comportamiento del sistema de IA?</p> <p>¿Se han establecido mecanismos de <b>comunicación para hacer saber a los usuarios el grado de fiabilidad</b> que tiene el sistema de IA, con base en los resultados de las pruebas de reproducibilidad?</p>
<b>INVERSIONISTAS</b>	<p>¿La compañía es la que ha desarrollado su sistema de IA o <b>depende de un desarrollador para la reproducción y prueba de resultados?</b></p>		

#### 4.2.5 Auditabilidad

La auditabilidad de los sistemas de IA se suele reservar para casos muy específicos en los que un tercero valora los procesos y resultados de los algoritmos. En este sentido, los modelos de IA que podrían tener una mayor propensión a ser auditados son aquellos que podrían tener un impacto en los derechos humanos o en la seguridad de las personas. Si bien

se trata de un concepto cuya aplicación es limitada, es conveniente tomar en consideración algunos aspectos preparativos que faciliten el proceso de recolección y estructuración de información en caso de requerir una auditoría (Reglamento General de Protección de Datos de la Unión Europea 2016-2018).

	ETAPA I	ETAPA II	ETAPA III
EMPRESAS	<p>¿Se han establecido <b>mecanismos que faciliten la auditabilidad</b> del sistema? Por ejemplo a través de la <i>documentación y trazabilidad de procesos y resultados</i>.</p> <p>¿Se ha planteado la forma en que <b>se debería compartir la información comercialmente sensible</b> en caso de auditoría?</p>	<p>En casos de que el sistema de IA tenga implicaciones para la seguridad de los usuarios, <b>¿se podría auditar de forma independiente?</b></p>	<p>¿La empresa cuenta con <b>certificaciones de agencias normalizadoras</b>, que faciliten la entrega de documentación a los organismos auditores?</p>
INVERSIONISTAS	<p>¿La empresa se hace responsable del tratamiento de datos, <b>facilitando la trazabilidad de los flujos de datos</b>?</p> <p>¿Se lleva a cabo un <b>reporte periódico sobre la aplicación de códigos de gobernanza</b> a los inversionistas?</p>	<p>¿Las prácticas comerciales de la empresa como transacciones, adquisición de licencias u otros acuerdos están <b>alineados con los principios éticos de tratamiento de datos</b>?</p>	

#### 4.2.6 Mantenimiento

Los sistemas de IA requieren de un mantenimiento periódico a fin de contar con datos actualizados que permitan

un funcionamiento eficiente (EC 2018; Galdon 2020; WEF & IMDA 2020)

EMPRESAS	ETAPA I	ETAPA II	ETAPA III
	<p>¿Se llevan a cabo revisiones periódicas y <b>actualizaciones de las bases de datos</b> para asegurar su precisión, validez y relevancia?</p> <p>¿Se ha dedicado personal especializado para la <b>revisión y actualización periódica de los datos y algoritmos</b>?</p>	<p>¿Se ha contemplado el uso de <b>sistemas automáticos de notificación sobre la disponibilidad de nuevos datos</b> relevantes para el sistema de IA?</p> <p>¿Se ha <b>verificado si los sistemas se comportan de forma inesperada</b> en ciertos contextos?</p>	<p>¿Se utilizan diversos <b>modelos alternativos del sistema de IA para validar su operatividad</b> con respecto a los diferentes elementos o tipos de datos que se procesan, esto con el fin de mitigar el sesgo del sistema?</p>



# 5



## RELACIONAMIENTO CON ACTORES CLAVE

# 5

## RELACIONAMIENTO CON ACTORES CLAVE

**El relacionamiento con los actores clave se refiere de forma amplia a la implicación de los actores relevantes en todo el ciclo de vida de la IA, lo que incluye comunicaciones, colaboración, consultas, y rendición de cuentas con los titulares de la información, usuarios, clientes, inversionistas, socios, proveedores, entes reguladores y la sociedad civil en general (EC 2018; Omidyar 2019; WEF & IMDA 2020).**

---



# 5.1

## TRANSPARENCIA

	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿La organización ha <b>contemplado y diseñado piezas dirigidas a los diferentes actores que estarán involucrados</b> con el sistema de IA o que serían afectados de algún modo por este?</p> <p>¿La organización ha valorado y diseñado piezas de comunicación con base en <b>escenarios en los que sería necesaria una explicación sobre el funcionamiento</b> o los resultados del sistema de IA?</p>	<p>¿Se ha considerado la creación de una <b>estrategia de explicación según se trate de datos, modelos, elementos humanos, interferencias, impacto, etc.</b>?</p> <p>¿Se ha tomado en consideración la progresividad de uso del sistema de IA por parte de las diferentes audiencias a fin de <b>evitar una fatiga de comunicaciones</b>?</p> <p>¿Los <b>canales de comunicación y retroalimentación</b> son supervisados por personal calificado?</p>	<p>¿Se ha considerado el uso de diversas <b>herramientas y medios audiovisuales para mejorar la estrategia de comunicación</b> sobre el sistema de IA?</p> <p>En aquellos momentos en los que la explicación de las operaciones del sistema de IA no resulte útil, ¿se ha considerado la utilización de <b>técnicas de explicación contrafactual basadas en los test de replicabilidad</b>?</p>
<b>INVERSIONISTAS</b>	<p>¿La compañía tiene una <b>política de rendición de cuentas y transparencia o similar</b> con sus clientes y socios?</p>		<p>¿La <b>empresa comunica de forma eficiente los resultados de las evaluaciones de impacto en la protección de datos personales</b> de sus clientes, en caso de ser necesario? Sobre todo si el sistema de IA lleva a cabo tareas de toma de decisiones automatizadas que tienen un impacto en los clientes.</p>

## 5.2 CONSUMIDORES Y USUARIOS DEL SISTEMA

EMPRESAS	ETAPA I	ETAPA II	ETAPA III
	<p>¿Se <b>notifica a los usuarios que se está interactuando con un sistema de IA y qué implicaciones tiene</b>, antes de iniciar la interacción, por ejemplo con chatbots?</p> <p>¿Se <b>comunica a los usuarios la razón por la que se hace uso de esta tecnología y qué funcionalidades</b> concretas están vinculadas con la IA?</p> <p>¿Se comunica al usuario la <b>expectativa y el alcance del comportamiento</b> del sistema de IA?</p> <p>¿Existe la posibilidad de <b>que el usuario opte por salir del sistema</b> y eliminar todos sus datos personales?</p>	<p>Además de las herramientas de comunicación basada en texto, <b>¿se cuenta con infográficos y otras herramientas multimedia</b> que permitan al público entender el funcionamiento del sistema de IA?</p> <p>Para ciertos casos: Una vez que se genera algún resultado o una decisión derivada del sistema de IA, <b>¿se explica al usuario los factores que condujeron a tal resultado?</b></p> <p>Especialmente en casos de otorgamiento de créditos, ayudas de gobierno o coberturas de seguros.</p> <p>¿Existe un medio de comunicación a través del cual <b>el usuario pueda solicitar una revisión del resultado</b> del sistema de IA que le haya afectado de algún modo?</p>	<p>¿Se cuenta con una <b>política de explicación</b> que contemple los diversos casos en los que se requiere ofrecer información adicional a los usuarios?</p> <p>¿Se cuenta con <b>herramientas pedagógicas y talleres</b> que permitan a los clientes tener un mejor nivel de entendimiento e interacción con la empresa sobre el comportamiento del sistema de IA?</p> <p>En caso de utilización de chatbots para el relacionamiento con clientes, ¿este cuenta con un <b>método de grabación de la conversación, para efectos de mejora y aprendizaje únicamente</b>?</p>

## 5.3 ACCESO A FINANCIAMIENTO

El acceso a financiamiento para los proyectos relacionados con el uso de la IA es el resultado de un punto de encuentro entre el fondo de capital emprendedor y la startup. Se trata de un proceso de construcción de confianza y selección mutua basado en el intercambio de información para reducir la brecha de conocimiento entre ambas partes sobre las aplicaciones de la IA y el rendimiento de la inversión. En este punto es crucial tener una comunicación precisa sobre los alcances y limitaciones del

sistema de IA, e incluso de herramientas que no son propiamente IA, pero que llevan a cabo procesamiento de grandes cantidades de datos.

De igual manera, es fundamental que los equipos que se presenten a rondas de financiamiento incluyan un balance de profesionales de negocios, expertos sectoriales y científicos de datos que puedan ofrecer una vista completa y sólida de la solución que se presenta (EC 2018; Gonfalonieri 2019).



## EMPRESAS

<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<p>¿Se está promoviendo el uso de la <b>IA como mejor herramienta</b> para resolver el problema planteado?</p> <p>¿Existe una <b>narrativa simple y clara que vincule la identificación de problemas y las soluciones</b> que ofrecerá el sistema de IA con la <b>estrategia corporativa</b> en materia 42 de identificación de nicho de mercado, competidores y FODA?</p> <p>¿Se comunica con <b>precisión y transparencia el tipo de sistema de IA o tecnología propia que se utiliza</b>, incluso si solo se trata de un sistema de procesamiento de grandes cantidades de datos? En ocasiones se <i>suele confundir la realización de tareas automatizadas y de análisis de datos con la aspiración de utilizar estas herramientas para llevar a cabo predicciones o toma de decisiones autónomas e inteligentes como lo hacen los sistemas de IA.</i></p> <p>¿El pitch de la empresa se centra en <b>mostrar el valor añadido y las facilidades de implementación de la solución</b> más allá de enfocarse en el uso de un modelo específico de IA?</p> <p>¿La <b>empresa cuenta con un equipo diverso y balanceado de profesionales</b> en áreas de negocios, ciencia de datos, mercadotecnia, distribución, ventas y especialistas sectoriales?</p> <p>¿Qué tipo de <b>cláusulas de garantía</b> ofrece el fondo para proteger su inversión y es factible/aceptable cumplir con ellas?</p>	<p>¿Se han llevado a cabo <b>pruebas piloto del sistema de IA con indicadores</b> incipientes que den cuenta de su capacidad de despliegue en el mercado?</p> <p>¿Con qué <b>indicadores clave de desempeño (KPI)</b> se cuentan para comunicar la solidez del modelo de IA?</p> <p>¿Cuál es la <b>estrategia de relacionamiento con socios</b> estratégicos dentro de la cadena de valor?</p> <p>¿Se comunican elementos que muestran la posibilidad de <b>escalabilidad del modelo de negocio</b>, incluyendo estrategias de “go to market”?</p>	<p>¿La empresa cuenta con <b>indicadores que permitan transmitir el interés y la validación de clientes</b> que justifiquen el <i>retorno de inversión del software como servicio (SaaS) versus otras soluciones digitales?</i></p>

## INVERSIONISTAS

<p>En caso de no contar con un sistema propiamente de IA, pero cuyos elementos permitirían una posible adopción en el corto o mediano plazo <b>¿Con qué grado de intervención humana cuenta el sistema de IA</b> y en qué fase de transición se encuentra la empresa?</p> <p>¿Qué <b>historial tiene la empresa</b> en el uso de IA para soluciones en el mundo real y cuáles han sido sus retornos de inversión?</p> <p>¿Qué <b>nivel de dependencia tiene en la obtención de derechos de propiedad intelectual</b> sobre datos y modelos de IA <b>en caso de no contar con tecnología propia</b>?</p>	<p>¿La empresa cuenta con la <b>capacidad de convertirse en líder de segmento</b> y dominar su mercado?</p> <p>Para productos dirigidos a consumidores usuarios ¿el sistema de IA utiliza los datos que le brinda el usuario para <b>crear mayor personalización y vínculo con el servicio</b>?</p>	<p>¿Cuál es el <b>nivel de impacto social</b> que ha tenido la empresa mediante la utilización de su solución de IA y <b>qué tan sostenible es su retorno de inversión</b>?</p>
---	---	---

# 6

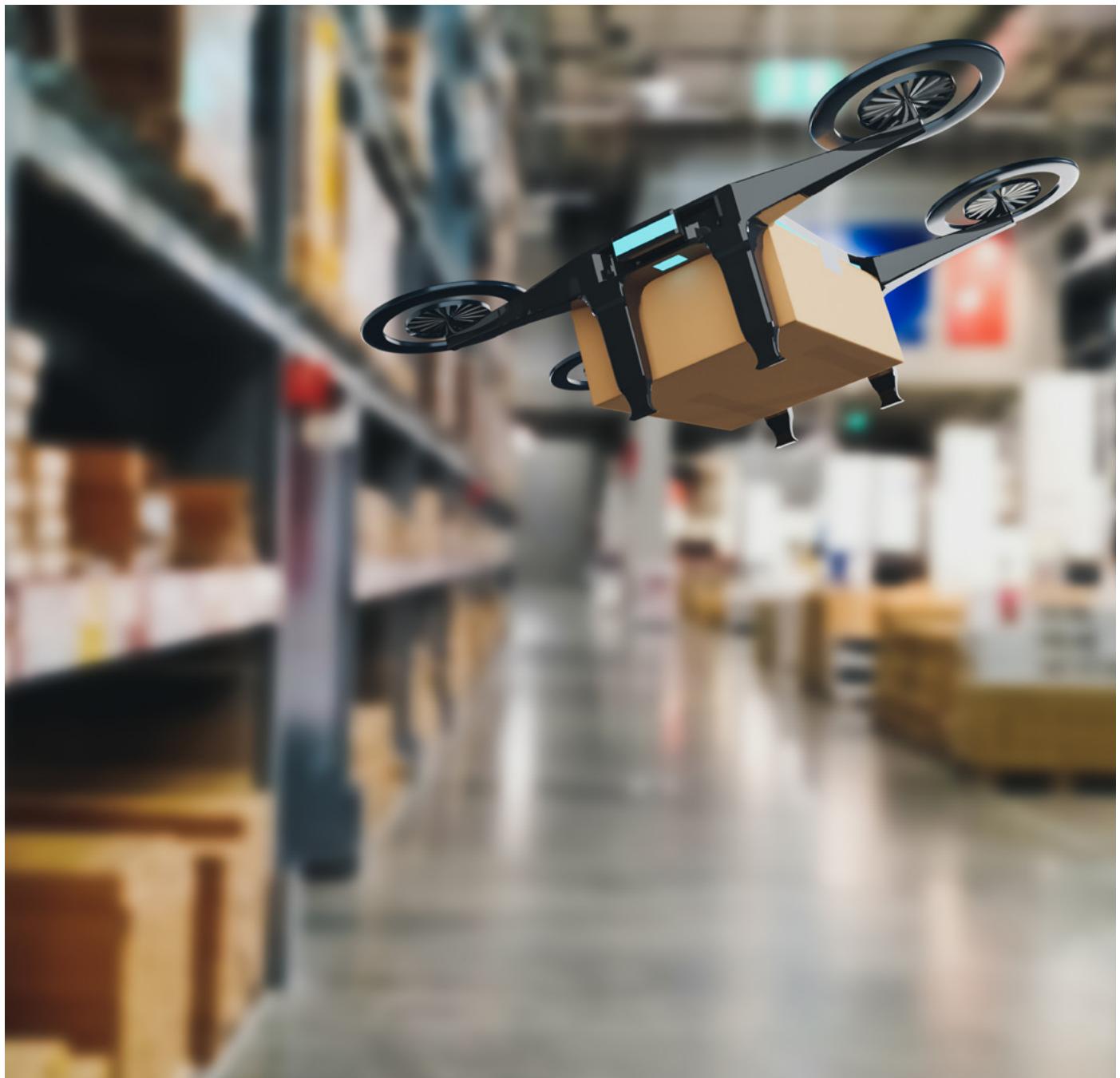


## COMUNICACIÓN

# 6 COMUNICACIÓN

**Las empresas que hacen uso de sistemas de IA deberían llevar a cabo comunicaciones proactivas, claras y sistemáticas a las partes interesadas sobre las capacidades y limitaciones de los sistemas de IA. En este sentido, resulta importante manejar las expectativas sobre el alcance tecnológico (Omidyar 2019; Engler 2020; WEF & IMDA 2020).**

---



	<b>ETAPA I</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>EMPRESAS</b>	<p>¿Se comunica de forma clara los <b>términos de servicio</b> ofrecido por el sistema de IA?</p> <p>¿Se han establecido <b>canales de comunicación con los usuarios</b> a fin de que estén enterados de que interactúan con un sistema de IA?</p> <p>¿Se han ofrecido <b>alternativas de interacción con seres humanos?</b></p> <p>¿Se han establecido mecanismos de <b>recepción de retroalimentación</b> por parte de los usuarios finales del sistema?</p> <p>¿Se han compartido los <b>canales de comunicación, procedimientos y puntos de contacto para resolver problemas y preguntas</b> sobre el sistema de IA?</p>	<p>¿Se ha comunicado a las partes interesadas, <b>incluyendo accionistas y usuarios sobre sesgos o errores previamente identificados</b> o percibidos?</p> <p>¿Se ha considerado algún mecanismo para <b>involucrar a las partes interesadas en el desarrollo o retroalimentación</b> del sistema de IA?</p> <p>¿<b>Se ha comunicado al interior de la empresa la adopción de la IA</b>, incluyendo a trabajadores posiblemente afectados y representantes?</p>	<p>¿Se han tenido en cuenta <b>diversos escenarios sobre los cuales se debería comunicar</b> de forma clara ciertas operaciones críticas del sistema de IA?</p> <p>¿Se ha tomado en consideración la posibilidad de que <b>las explicaciones corren el riesgo de generar confusión, sesgo de confirmación o fatiga cognitiva de las audiencias?</b></p> <p>¿Se ha elaborado una <b>taxonomía o glosario que clarifique conceptos como “justicia”</b> con respecto al uso del sistema de IA?</p> <p>¿Existe una métrica o <b>parámetros para la determinación de conceptos como “justicia”</b> en el uso del sistema de la IA?</p>
<b>INVERSIONISTAS</b>	<p>¿Es posible o <b>necesario determinar el grado de comprensión de las contrapartes sobre los términos de servicio del sistema de IA?</b></p>		

# **ANEXO 1**

# **ANEXO 1 - DESCRIPCIÓN DE LAS ETAPAS DE DESARROLLO DE LAS EMPRESAS**

**Con la finalidad de abarcar el mayor espectro posible de emprendimiento y sectores que hagan uso o pretendan incorporar la IA a su negocio, esta matriz de autoevaluación ética desarrolla tres etapas para la autoevaluación del uso ético y responsable de la IA. Estas etapas otorgan a la matriz su condición de progresión, ya que las evaluaciones cambian según el nivel de desarrollo del emprendimiento (Churchill y Lewis 1983; Dibner 2018; Areitio 2019), el nivel de alistamiento tecnológico (Comisión Europea 2014-2015) y la participación de diferentes actores en las diversas etapas.**

El componente principal que determina las etapas es el nivel de desarrollo de los emprendimientos desde una óptica de rondas de financiamiento, incluyendo las etapas semilla y early stage en la etapa I, startups en serie A y B en la etapa II y startups en serie C y corporaciones en etapa III. Esto a su vez, se suele corresponder con el nivel de desarrollo tecnológico que han alcanzado los emprendimientos, de acuerdo con los nueve niveles de madurez tecnológica (o

technology readiness levels - TRL) de la Comisión Europea categorizados en tres niveles: investigación, desarrollo y despliegue. Finalmente, cabe destacar que cada etapa conlleva una interacción distinta entre actores, de este modo las incubadoras suelen operar en la etapa I, mientras que las aceleradoras podrían tomar un rol más activo en las etapas I y II, y los inversionistas roles más determinantes en la etapa III.

<b>DETERMINANTES DE LAS ETAPAS</b>	<b>ETAPA II</b>	<b>ETAPA II</b>	<b>ETAPA III</b>
<b>Nivel de desarrollo de la startup</b> <i>*Basado en la conceptualización de Harvard Business Review, Medium, Mattermark, Crunchbase, The Venture City.</i>	<b>Semilla/early stage</b> Ideación y desarrollo de producto.  Llegada a los primeros usuarios y descubrimiento del mercado.  Definición del modelo de negocio.  Primeras contrataciones.  Alistamiento Serie A	<b>Later stage Path to growth</b>  Startups en Serie A y B + bridge rounds.  <b>Serie A</b> Camino hacia la consolidación. Inicio de distribución de producto o servicio a mayor escala.  <b>Serie B</b> Tracción comercial con clientes y modelo de negocio bien definido.  Escalar modelo de negocio, base de usuarios y adquisiciones.	<b>Mature</b>  Startups en serie C y corporaciones.  <b>Serie C</b> Adquisición de mayor capital y crecimiento acelerado.  Identificación de rentabilidad, pero con necesidad de mayor capital.  Internacionalización.  Más adquisiciones y contrataciones.  <b>Corporaciones</b> Mayor consolidación en el mercado y rentabilidad.
<b>Nivel de madurez tecnológica (TRL)</b> <i>*Basado en los niveles de alistamiento tecnológico desarrollados por la Comisión Europea 2014-2015.</i>	<b>Investigación</b>  <b>Nivel 1</b> - Principios básicos observados y reportados.  <b>Nivel 2</b> - Concepto y aplicación tecnológica formulada.  <b>Nivel 3</b> - Prueba experimental del concepto.	<b>Desarrollo</b>  <b>Nivel 4</b> - Validación de componentes y disposición en entorno de laboratorio.  <b>Nivel 5</b> - Valoración de componente y disposición en entorno relevante.  <b>Nivel 6</b> - Demostración de prototipo en entorno relevante.	<b>Despliegue</b>  <b>Nivel 7</b> - Demostración de prototipo de sistema en entorno operativo.  <b>Nivel 8</b> - Sistema completo y certificado a través de pruebas y demostraciones.  <b>Nivel 9</b> - Sistema probado con éxito en entorno real.
<b>Actores involucrados</b>	Emprendedor  Inversionista (ángel o company builder)  Incubadora	Emprendedor  Inversionista (capital emprendedor/private equity)  Incubadora/Acelerador	Emprendedor  Inversionista

# FUENTES

**Areitio, Andy (2019).** *De idea a scale up: Fases de una start up.* Medium. The Venture City.  
<https://medium.com/theventurecity/fases-de-la-startup-de-idea-a-scale-up-73c8e81cd37>

**Banco Interamericano de Desarrollo (BID) (2020).** *Adopción ética y responsable de la Inteligencia Artificial en América Latina y el Caribe.*

**Bartram, Robert (2018).** *The New Frontier for Artificial Intelligence.* ISO Committee on Artificial Intelligence. ISO/IEC JTC 1/SC 42.  
<https://www.iso.org/news/ref2336.html>

**Berkman Klein Center for Internet & Society at Harvard University (2020).** *Principled Artificial Intelligence. A Map of Ethical and Rights-Based Approaches to Principles for AI.*

**Cabrol, Marcelo; Baeza-Yates, Ricardo; González Alarcón, Natalia; Pombo, Cristina (2020).**  
¿Es la privacidad de los datos el precio que debemos pagar para sobrevivir a una pandemia?. Banco Interamericano de Desarrollo.  
<http://dx.doi.org/10.18235/0002292>.  
<https://publications.iadb.org/es/es-la-privacidad-de-los-datos-el-precio-que-debemos-pagar-par-a-sobrevivir-una-pandemia>

**Churchill, Neil C. & Lewis, Virginia L. (1983).**  
*The Five Stages of Small Business Growth.* Harvard Business Review. May 1983 Issue.  
<https://hbr.org/1983/05/the-five-stages-of-small-business-growth>

**Comisión Europea (2018).** *Directrices Éticas para una IA Fiable.* Grupo Independiente de Expertos de Alto Nivel sobre Inteligencia Artificial.

**Dassen, Thommy; Hou, Naiwen; Kronseder, Veronika (2019).** Chapter 2. *Introduction to Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE).*  
[https://compstat-lmu.github.io/iml\\_methods\\_limitations/pdp.html](https://compstat-lmu.github.io/iml_methods_limitations/pdp.html)

**Dibner, Gil (2018).** *There are only three startup stages.* Medium. Angular Ventures.  
<https://medium.com/angularventures/there-are-only-three-stages-for-startups-b8783d6b0f1>

**Engler, Alex (2020).** *The Case for AI Transparency Requirements.* Brookings Institution.  
<https://www.brookings.edu/research/the-case-for-ai-transparency-requirements/>

**European Commission (EC) (2014-2015).** *Technology Readiness Levels. Horizon 2020 Work Programme 2014-2015. Annex.*  
[https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014\\_2015/annexes/h2020-wp1415-annex-g-trl\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf)  
\_\_\_\_\_. (2018). *Trustworthy AI Assessment List. Independent High Level Expert Group on Artificial Intelligence.*

**European Parliament Research Service (2019).** *A Governance Framework for Algorithmic Accountability and Transparency.*  
[https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS\\_STU\(2019\)624262\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2019/624262/EPRS_STU(2019)624262_EN.pdf)

**European Union (2019).** *Annual Report 2019.* European Data Protection Supervisor.

**Everis & Endeavor (2018).** *El impacto de la Inteligencia Artificial en el emprendimiento.*  
<http://www.endeavor.cl/wp-content/uploads/El-impacto-de-la-IA-en-el-emprendimiento-en-América-Latina-everis-y-Endeavor.pdf>

**Galdon, Gemma (2020).** *Protocolo de Evaluaciones Éticas.* Éticas Consulting.

**Gigler, Bjorn-Soren (2020).** *Closing The Investment Gap For Artificial Intelligence And Blockchain In Europe.* Panodissey.  
<https://panodissey.com/en/article/entrepreneurship/forging-new-frontiers-in-finance-for-digital-innovations-m4qcv3ktj9ft>

**Gonfalonieri, Alexandre (2019).** *How to get Funding for AI Startups.* KDnuggets.  
<https://www.kdnuggets.com/2019/06/funding-ai-startups.html>

**González Alarcón, Natalia; Pombo, Cristina (2020).** *¿Cómo puede la inteligencia artificial*

ayudar en una pandemia?. Banco Interamericano de Desarrollo.

<http://dx.doi.org/10.18235/0002300>.

<https://publications.iadb.org/es/como-puede-la-inteligencia-artificial-ayudar-en-una-pandemia>

**Google (2018).** Whitepaper: Perspectives on Issues in AI Governance.

**Hoepman, Jaap-Henk (2020).** Privacy Design Strategies (The Little Blue Book). Computing Science Department. Radboud University.

<https://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf>

**IEEE P7000TM Standards.**

**Infocomm Media Development Authority and Personal Data Protection Commission, Singapore (IMDA & IDPC) (2020).** Model Artificial Intelligence Governance Framework. Second Edition.

**ISO/IEC JTC 1/SC 42 Artificial intelligence.**

**KPMG (2018).** La importancia de los Privacy Impact Assessment (PIA) en la protección de datos.

<https://www.tendencias.kpmg.es/2018/01/la-importancia-de-los-privacy-impact-assessment-pia-en-la-proteccion-de-datos/>

**Kusner, Matt; Loftus, Joshua; Russell, Chris & Silva, Ricardo (2017).** Counterfactual Fairness.

Neural Information Processing Systems.

<https://papers.nips.cc/paper/6995-counterfactual-fairness.pdf>

**Lewis, Barnaby (2019).** Standards Cooperation is Key to Making AI and Smart Cities a Reality. ISO Committee on Artificial Intelligence. ISO/IEC JTC 1/SC 42.

<https://www.iso.org/news/ref2381.html>

**Lockett, Kath (2019).** Embracing the Power of Technology. ISO Committee on Artificial Intelligence.

ISO/IEC JTC 1/SC 42. <https://www.iso.org/news/ref2451.html>

**MacCarthy, Mark (2020).** AI needs more regulation not less. Brookings Institution.

<https://www.brookings.edu/research/ai-needs-more-regulation-not-less/>

**McKinsey Global Institute (MGI) (2017).** Artificial Intelligence: The Next Digital Frontier?. Discussion Paper.

**Méndez, Manuel Ángel (2020).** El Gobierno se une a una iniciativa digital de rastreo cuestionada: "Es un caballo de Troya". El Confidencial. Apps. [https://www.elconfidencial.com/tecnologia/2020-04-17/gobierno-apps-rastreo-pepp-pt-dp3t-coronavirus\\_2553776/](https://www.elconfidencial.com/tecnologia/2020-04-17/gobierno-apps-rastreo-pepp-pt-dp3t-coronavirus_2553776/)

**Molnar, Christoph (2020).** Interpretable Machine Learning: A Guide for Making Blackbox Models Explainable. <https://christophm.github.io/interpretable-ml-book/>

**Official Journal of the European Union (2016-2018).** General Data Protection Regulation 2016/679.

**Omidyar Network (2019).** Due -Diligence Checklist for Privacy-Aware Investing. Race to the Top Initiative.

\_\_\_\_\_ (2019). Progression Matrix. Race to the Top Initiative.

\_\_\_\_\_ (2019). The Emerging Market in Digital Trust: New Obligations and New Opportunities for Venture Capital.

**Organisation for Economic Co-operation and Development (2018).** Private Equity Investment in Artificial Intelligence.

\_\_\_\_\_ (2019). OECD Principles on Artificial Intelligence.

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

**Parlamento Europeo (2016).** Reglamento General de Protección de Datos de la Unión Europea.

<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32016R0679&from=es>

**Posey, Luke (2019).** Intro to AI Ethics. Towards Data Science.

<https://towardsdatascience.com/intro-to-ai-ethics-2a31fbbe6e87>

**Potgieter, Luke (2018).** Semantic Interoperability: Are you training your AI by mixing data sources that look the same but aren't?. KDnuggets.

<https://www.kdnuggets.com/2018/10/semantic-interoperability-training-ai-mixing-different-data-sources.html>

**Szramke, Tomasz (2017).** Types of AI Algorithms. Growly. <https://www.growly.io/types-of-ai-algorithms/>

**World Economic Forum in collaboration with  
the Info-communications Media Development  
Authority of Singapore (WEF & IMDA) (2020).**

*Companion to the Model AI governance  
Framework - Implementation and Self-Assessment  
Guide for Organizations.*

**Wright, Ray (2019).** *Interpreting Black  
Box Machine Learning Models Using Partial  
Dependence (PDP) and Individual Conditional  
Expectation (ICE) Plots.* SAS Institute.  
<https://www.sas.com/content/dam/SAS/support/en/sas-global-forum-proceedings/2018/1950-2018.pdf>