

DOCUMENTO DEL BANCO INTERAMERICANO DE DESARROLLO
FONDO MULTILATERAL DE INVERSIONES

REGIONAL

**INTELIGENCIA ARTIFICIAL RESPONSABLE PARA EL CRECIMIENTO EN AMERICA LATINA Y EL
CARIBE**

RAI 4 GROWTH

(RG-T4463)

MEMORANDO A LOS DONANTES

Este documento fue preparado por el equipo de proyecto integrado por Carolina Carrasco (DIS/CCH), Tetsuro Narita (LAB/INV), Ana Castillo (DIS/CUR), Tatiana Virviescas (DIS/CME), Paula Auerbach (DIS/CEC), Mara Balestrini (LAB/LAB), Arturo Munte (IFD/ICS), Juan Pedeflous (GCL/FML), Delfina Müller (DSP/DVF).

El presente documento contiene información confidencial comprendida en una o más de las diez excepciones de la Política de Acceso a Información e inicialmente se considerará confidencial y estará disponible únicamente para empleados del banco. Se divulgará y pondrá a disposición del público una vez aprobado.

ÍNDICE

RESUMEN DEL PROYECTO

I.	EL PROBLEMA	2
	A. Descripción del problema	2
II.	LA PROPUESTA DE INNOVACIÓN.....	7
	A. Descripción del proyecto	7
	B. Resultados, medición, seguimiento y evaluación del proyecto.....	16
III.	ALINEACIÓN CON EL GRUPO BID, AMPLIACIÓN DE ESCALA Y RIESGOS	17
	A. Alineación con el Grupo BID.....	17
	B. Adicionalidad y Escalabilidad.....	18
	C. Riesgos del proyecto e institucionales	20
IV.	INSTRUMENTO Y PROPUESTA DE PRESUPUESTO.....	21
V.	ORGANISMO EJECUTOR Y ESTRUCTURA DE IMPLEMENTACIÓN	21
	A. Descripción del organismo ejecutor	21
	B. Estructura y mecanismo de implementación.....	22
VI.	CUMPLIMIENTO DE HITOS Y ACUERDOS FIDUCIARIOS ESPECIALES.....	23
VII.	ACCESO A INFORMACIÓN Y PROPIEDAD INTELECTUAL	23

RESUMEN DEL PROYECTO

INTELIGENCIA ARTIFICIAL RESPONSABLE PARA EL CRECIMIENTO EN AMÉRICA LATINA Y EL CARIBE

RAI 4 GROWTH

RG-T4463

En el ecosistema de innovación de América Latina y el Caribe (ALC), la adopción de la Inteligencia Artificial (IA), tecnología que se está convirtiendo en un elemento fundamental para el crecimiento económico, está por debajo de la media global. Aunque el sector privado regional está comenzando a integrar la IA en sus modelos de negocios, productos y servicios, la falta de un acompañamiento adecuado conlleva el riesgo de minar la confianza de la sociedad en estos sistemas. Esto supone un freno para su adopción y, en consecuencia, para la mejora de la competitividad de las empresas que los desarrollan y utilizan. Además, si la IA no se desarrolla de manera responsable, se corre el riesgo de reproducir y amplificar desigualdades a gran escala, como ya ha ocurrido en otras regiones más avanzadas del mundo. La falta de equidad, transparencia y rendición de cuentas puede perpetuar prejuicios y errores humanos.

Asegurar que el desarrollo y uso de la IA se alinea, de manera temprana si es posible, con estándares internacionales en sectores como la inclusión financiera, empleo, educación, salud y justicia resulta fundamental para su adopción y sostenibilidad financiera en el largo plazo. Lo es también para asegurar que los beneficios de esta tecnología alcanzan a todos los grupos de la sociedad, en particular aquellos viviendo en condiciones de pobreza y vulnerabilidad. De este modo, la incorporación de un enfoque de “Responsabilidad por Diseño” en la creación y uso de sistemas de IA permitirá a las empresas materializar la oferta de valor de esta tecnología beneficiando a todos los estratos de la sociedad, lo que a su vez mejorará la condición competitiva de las empresas en el mercado, al desarrollar productos y servicios robustos y fiables. Además, el uso de la IA responsable por parte del tejido empresarial más innovador facilita un posicionamiento de empresas clave para el desarrollo de la economía de la región, facilitando la internacionalización de startups y empresas en crecimiento, gracias al cumplimiento (temprano en sus países) de regulación relacionada con la mitigación de riesgos de la IA: esto les permite ser empresas más competitivas, generando empleos de mayor calidad. Adicionalmente, en estas etapas, las empresas fomentan una cultura que valora la diversidad como motor de creatividad e innovación. Por lo tanto, son actores clave para colaborar en la mitigación de riesgos en áreas como la empleabilidad y la formación.

RAI 4 Growth (Responsible AI for Growth), la solución propuesta, es un programa integrado que fomenta el uso generalizado de estándares y herramientas de IA responsables entre las empresas de IA en la región de ALC, así como la adopción de herramientas de IA por parte de startups, MiPyMEs e inversores de la región. Este programa desarrolla además un conjunto de planes de medición¹ (measurement plans) para que tanto las empresas desarrolladoras como las usuarias de IA puedan comprender cómo estos sistemas responden a la mitigación de riesgos relacionados con la vulnerabilidad económica y de exclusión, la perspectiva de género o la sostenibilidad. De

¹ Un plan de medición de un sistema de software es un conjunto estructurado de métricas y criterios diseñados para evaluar su desempeño, calidad e impacto a lo largo de su ciclo de vida. En el caso de sistemas de IA responsable, un plan de medición se utiliza para asegurar que el sistema cumple con principios éticos y estándares de responsabilidad, tales como la equidad, la transparencia y la no discriminación.

esta manera, tanto empresas usuarias como desarrolladoras de IA podrán beneficiarse de esta poderosa tecnología para una mayor eficiencia y confianza, al tiempo que reducen sus riesgos y aumentan su competitividad. El programa se estructura en base a tres pilares:

- **Herramientas, certificaciones y medición** para la implementación adecuada de IA responsable (RAI por sus siglas en inglés).
- **Capacidades y recursos formativos** para el desarrollo responsable de la IA en el ecosistema de innovación.
- **Fomento de confianza y concientización**, centrado en informar y concienciar a la población sobre el uso crítico de productos y servicios de IA que impacten positivamente en su día a día.

El proyecto cuenta con la experiencia de la Asociación de la Economía Digital (Adigital) de España, a la que se sumarán colaboradores (personas y organizaciones) expertos en tecnología, regulación, ética y concienciación pública, dando impulso a la iniciativa fAIr LAC del Grupo BID, la cual promueve el uso ético y responsable de la inteligencia artificial para contribuir al desarrollo socioeconómico inclusivo de la región.

SIGLAS Y ABREVIATURAS

IA	Inteligencia Artificial
ALC	América Latina y el Caribe
MiPyMEs	Micro, pequeñas y medianas empresas
OCDE	Organización para la Cooperación y el Desarrollo Económico
OE	Organismo Ejecutor
RAI	IA responsable (por sus siglas en inglés)

RESUMEN EJECUTIVO

INTELIGENCIA ARTIFICIAL RESPONSABLE PARA EL CRECIMIENTO EN AMÉRICA LATINA Y EL CARIBE

RAI 4 GROWTH

RG-T4463

País y ubicación geográfica:	Regional con acciones iniciales en Chile, México, Colombia, Uruguay y Ecuador		
Organismo Ejecutor:	Adigital, Asociación Española de la Economía Digital, será Organismo Ejecutor de financiamiento BID Lab por hasta US\$ 1.400.000 y aportará recursos de contrapartida. BID Lab (Fondo Multilateral de Inversiones) será el ejecutor de sus propios recursos por hasta US\$300.000.		
Verticales y Transversales de BID Lab:	Educación, talento y empleo. Género y diversidad.		
Coordinación con otros donantes/ operaciones del Banco:	Coordinación con la iniciativa fAIr LAC y operaciones vinculadas con el desarrollo de conocimiento, herramientas, marcos analíticos, formación, análisis de la IA responsable para el ecosistema de innovación en la región.		
Beneficiarios del proyecto:	Inversores, MiPyMEs y Startups, con énfasis en aquellas cuyos productos o servicios están orientados a poblaciones vulnerables, mujeres o beneficios medioambientales.		
Financiamiento:	Cooperación Técnica:	US\$1.700.000	53%
	Financiamiento total de BID Lab:	US\$1.700.000	
	Contrapartida:	US\$1.500.000	47%
	Presupuesto total del proyecto:	US\$3.200.000	100%
Período de ejecución y de desembolso:	36 meses de ejecución y 42 meses de desembolso.		
Condiciones para el Primer Desembolso	Para el primer desembolso de la cooperación técnica, el OE presentará a satisfacción del Banco: (i) el plan de contrataciones del proyecto, (ii) la designación del gerente de programa, (iii) la designación del coordinador del proyecto.		
Revisión del impacto ambiental y social:	El 20 de mayo de 2024 se analizó y clasificó esta operación de acuerdo con el Marco de Política Ambiental y Social del BID (documento GN-2965-21). Dado que los impactos y riesgos son limitados, se propone clasificar el proyecto en la categoría C.		
Unidad responsable de los desembolsos:	CCH – Para los recursos cuyo Organismo Ejecutor es Adigital LAB/DIS – Para los recursos cuyo Organismo Ejecutor es BID Lab		

I. EL PROBLEMA

A. Descripción del problema

- 1.1 La integración de la Inteligencia Artificial (IA) en el sector empresarial y el ecosistema de innovación de América Latina y el Caribe (ALC) presenta un gran potencial para ayudar a abordar los desafíos que impone la Agenda 2030² para el desarrollo sostenible. Este conjunto de tecnologías puede ayudar a mejorar el impacto del desarrollo económico en sectores críticos como la salud, la agricultura, la industria (debido a su amplia cadena de valor), la educación y la provisión de servicios públicos, estimándose en más de 5% su potencial de contribución al PIB regional a 2030³. También puede contribuir a que el mercado regional (oferta y demanda de soluciones basadas en IA) crezca un 18,30% entre 2023 y 2028⁴.
- 1.2 Una encuesta reciente realizada por NTT Data y MIT Technology Review mostró que el interés en el uso de IA entre las empresas latinoamericanas aumentó del 58% en 2020 al 71% en 2023 y que el 79% de ellas ya ha iniciado algún grado de utilización de herramientas de IA⁵. Sin embargo, el mismo informe constata que la región de ALC todavía está rezagada con respecto al promedio mundial: mientras que la tasa general de implementación efectiva de IA en empresas en la región es del 37%, a nivel mundial es del 42%. Las principales motivaciones del sector empresarial para adoptar estas herramientas guardan relación con la búsqueda de mejoras en competitividad, eficiencia operativa, calidad de las operaciones y disminución de errores.
- 1.3 A pesar de que existe una clara intención de adoptar la IA en el sector privado regional, aún son pocas las empresas que se encuentran preparadas para aprovechar su máximo potencial, así como gestionar de manera adecuada los riesgos que ésta conlleva. Por ejemplo, el reporte 2023 sobre el estado de la IA de Mckinsey señala que solo el 21% de las empresas que declaran haber adoptado herramientas de IA han establecido políticas explícitas sobre su uso o medidas de mitigación de riesgos tales como la falta de transparencia, falta de exactitud, ciberseguridad⁶ u otros de mayor gravedad como el riesgo de generar sesgos dañinos o discriminación algorítmica⁷. Lo anterior es especialmente sensible dado que el uso de algoritmos está afectando y transformando de manera creciente la vida de la población a nivel global, la forma como trabaja, aprende, se recrea o accede a servicios esenciales como la salud, las finanzas, la educación, el empleo, la asistencia social, infraestructuras críticas, agricultura y seguridad alimentaria o incluso las telecomunicaciones⁸. A pesar del bajo nivel de salvaguardas establecidas para la reducción de riesgos en el uso de esta tecnología, un reporte de Statista de julio de 2023 indica que el 70% de la ciudadanía en la región percibe que la IA tiene un potencial que supera cualquiera de sus inconvenientes⁹. La combinación de la acelerada adopción empresarial de la IA con las percepciones favorables de la ciudadanía sobre el uso esta tecnología

² <https://www.un.org/sustainabledevelopment/es/2015/09/la-asamblea-general-adopta-la-agenda-2030-para-el-desarrollo-sostenible/>

³ <https://forbescentroamerica.com/2023/05/17/5-de-las-principales-startups-de-inteligencia-artificial-de-america-latina-y-el-caribe>

⁴ Seizing the opportunity: the future of AI in Latin America, Economist Impact, 2022.

⁵ La inteligencia artificial en America Latina, 2023, NTT Data – MIT Technology Review.

⁶ <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year>

⁷ <https://datos.gob.es/es/blog/invisibilizacion-y-discriminacion-algoritmica>

⁸ The Age of AI has begun, GatesNotes, <https://www.gatesnotes.com/The-Age-of-AI-Has-Begun>

⁹ <https://es.statista.com/grafico/30343/aprobacion-de-los-beneficios-de-la-inteligencia-artificial-en-america-latina/>

confirma una tendencia en la que su veloz introducción alcanzará prácticamente todos los ámbitos de la vida cotidiana¹⁰, lo que a su vez incrementa significativamente el nivel de riesgo e impacto de sus posibles efectos negativos.

- 1.4 Por otro lado, si bien a nivel global las MiPyMEs presentan brechas de productividad respecto a las grandes empresas, éstas son especialmente significativas en América Latina y el Caribe. El valor de la producción de las MiPyMEs regionales representa sólo el 25% del valor total de la producción de la región, generando apenas la mitad de la riqueza que sus pares europeos¹¹. Es en este ámbito donde la automatización y la IA representan una gran oportunidad, no solo para impulsar la competitividad de las PyMEs de la región sino también al ecosistema de empresas emergentes (startups) que ha evolucionado de manera muy relevante en los últimos años¹².
- 1.5 La aceleración y el potencial de autonomía en la toma de decisiones de los sistemas de inteligencia artificial presentan riesgos no solo técnicos, sino también sociales. Estos riesgos podrían amplificar las brechas de desigualdad y perpetuar dinámicas de discriminación y exclusión socioeconómica. Esto se debe a su rápida adopción, a la falta de herramientas y concientización en el sector privado para gestionar estos riesgos y al impacto a gran escala de esta tecnología. A continuación, se describen algunos de los riesgos más relevantes:
- El uso de la IA afectará a la población en condición de pobreza y vulnerabilidad. Solo en ALC (la región del mundo con el nivel más alto de desigualdad socioeconómica del planeta¹³) puede llegar a afectar a más de 180 millones de personas que no cuentan con ingresos para cubrir sus necesidades básicas y 70 millones que ni siquiera tiene el nivel de ingresos suficiente para adquirir una canasta básica de alimentos¹⁴. Tanto la desigualdad socioeconómica como la pobreza son dos dimensiones que pueden verse severamente afectadas por el rápido desarrollo y despliegue de IA si no se incluyen las salvaguardas y estrategias de mitigación de riesgos adecuadas que mitiguen posibles riesgos de discriminación, concentración de la renta y captura asimétrica y unilateral del valor económico¹⁵, así como posibles afectaciones en los mercados laborales por la automatización y efectos adversos en los ingresos de las personas con más bajos ingresos¹⁶. Estos no son problemas a largo plazo, sino desafíos sobre los que hay tomar prontas acciones¹⁷.
 - Otro factor relevante a considerar en el desarrollo de la IA es su impacto en la sostenibilidad ambiental: entrenar un solo modelo avanzado de IA puede producir hasta la misma cantidad de dióxido de carbono que 300 vuelos de ida y vuelta entre Nueva York y San Francisco¹⁸, y se estima que el entrenamiento de modelos fundacionales requerirá una cantidad de energía cercana a un Gigawatt, prácticamente la cantidad de

¹⁰ <https://elpais.com/economia/formacion/2023-08-10/que-queda-al-margen-de-la-inteligencia-artificial.html>

¹¹ Pequeñas empresas, grandes impactos: Apoyando a las PYMES productivas como motor de la recuperación en ALC, PNUD, 2021.

¹² Empresas emergentes (start-ups) en América Latina y el Caribe, Cepal, 2023.

¹³ <https://www.iadb.org/es/noticias/las-complejidades-de-la-desigualdad-en-america-latina-y-el-caribe>

¹⁴ <https://www.cepal.org/es/publicaciones/68702-panorama-social-america-latina-caribe-2023-la-inclusion-laboral-como-eje-central>

¹⁵ <https://www.economist.com/business/2024/03/17/just-how-rich-are-businesses-getting-in-the-ai-gold-rush>

¹⁶ <https://www.imf.org/en/Blogs/Articles/2020/12/02/blog-how-artificial-intelligence-could-widen-the-gap-between-rich-and-poor-nations>

¹⁷ Harnessing the power of AI and emerging technologies. Background paper for the CDEP Ministerial meeting, 15 de noviembre de 2022.

¹⁸ <https://earth.org/the-green-dilemma-can-ai-fulfil-its-potential-without-harming-the-environment/>

energía que genera un reactor nuclear¹⁹. Si bien pocas empresas cuentan con los datos y la supercomputación necesaria para el entrenamiento de modelos y la mayoría utiliza modelos pre entrenados o entrenan sus modelos en hiper escaladores en la nube, es importante que las empresas usuarias incorporen enfoques conscientes respecto a la huella ambiental que deja el uso de sus modelos tecnológicos y que puedan dimensionar el impacto ambiental de estos sistemas. Por parte de las empresas que sí entrenan sus modelos, se observan avances en la definición de políticas de compensación (tanto monetarias, en especie, como en cuanto a la contribución positiva de algunos de esos modelos en la sostenibilidad del medio ambiente²⁰). Si bien al día de hoy no existen estándares al respecto de la medición y mitigación de los efectos ambientales del uso de la IA, las empresas pueden mejorar su transparencia y comenzar a dar los primeros pasos²¹.

- Uno de los principales riesgos del uso y desarrollo de algoritmos de IA que puede afectar a la población vulnerable es la discriminación algorítmica, es decir, que los datos utilizados para entrenar sistemas automatizados de decisión o modelos de IA repliquen, perpetúen y/o amplifiquen sesgos y desigualdades por factores como el género, la raza, la edad y sus interseccionalidades. Para las mujeres en general, la IA puede presentar riesgos extraordinarios para su seguridad económica y personal, así como graves riesgos de discriminación y exclusión si esta tecnología no se desarrolla y utiliza de forma adecuada²². Dicho fenómeno hace necesario que las empresas, independiente de su tamaño, e inversores tomen medidas de mitigación²³, además de explicar o rendir cuentas acerca de cómo lo están haciendo.

1.6 Estos riesgos, aunque generalizables en cualquier sector productivo, se acrecientan en aquellos sectores en los cuales existe evidencia respecto a la prevalencia de sesgos y vulnerabilidades. Se muestran a continuación algunos ejemplos de desafíos específicos en sectores relacionados con el emprendimiento y la innovación:

- En el ámbito de **FinTech**, los desafíos de la IA están estrechamente ligados a la necesidad de evaluar y mejorar los sistemas de inteligencia artificial para evitar problemas significativos relacionados con la equidad y la transparencia. La mitigación/eliminación de sesgo de género y racial en herramientas y algoritmos de calificación crediticia pueden ayudar a resolver el problema estructural de mujeres o minorías étnicas con historial bancario menor o nulo. En todos estos casos la transparencia algorítmica es crucial; la falta de claridad sobre cómo se calculan las puntuaciones crediticias puede erosionar la confianza en estos sistemas. Por último, la privacidad es una preocupación importante, ya que los datos financieros personales pueden ser expuestos o mal utilizados, lo que conlleva a violaciones de confianza.
- En el ámbito de la **salud**, la transparencia en herramientas de diagnóstico puede ayudar a encontrar fallas concretas en el diagnóstico de condiciones en mujeres y en la precisión de diagnósticos para ciertos grupos raciales. En todos estos casos, la transparencia algorítmica es crucial; la falta de comprensión de los procesos de diagnóstico impulsados por IA puede erosionar la confianza de los pacientes en estos

¹⁹ <https://youtu.be/i-o5YbNfmh0?si=cyVGhYCEnOa3cY7U>

²⁰ Fang, M., Li, R., Zhao, X. Improving new energy subsidy efficiency considering learning effect: A case study on wind power. *Journal of Environmental Management*, Volume 357, 2024, 120647. ISSN 0301-4797. <https://doi.org/10.1016/j.jenvman.2024.120647>

²¹ <https://oecd.ai/en/wonk/businesses-regulations-environmental-sustainability>

²² <https://www.brookings.edu/articles/ai-poses-disproportionate-risks-to-women/>

²³ <https://elpais.com/america/lideresas-de-latinoamerica/2023-09-14/al-regular-la-inteligencia-artificial-debemos-colocar-raza-y-genero-en-el-centro-del-debate.html>

sistemas. Por otra parte, los datos de salud sensibles podrían ser filtrados o mal utilizados, poniendo en riesgo la confidencialidad del paciente.

- En el ámbito de la **empleabilidad**, la mitigación/eliminación de sesgo de género y racial en los algoritmos de reclutamiento y emparejamiento de oferta y demanda de empleo puede ayudar a resolver el problema estructural de favorecer inadvertidamente a hombres y ciertos grupos raciales basándose en datos de entrenamiento sesgados. Desde el punto de vista de la privacidad, las verificaciones de antecedentes podrían exponer información personal innecesariamente.
- En el sector **educativo** es esencial abordar sesgos en las herramientas y los algoritmos de recomendaciones de contenido educativo, que pueden reforzar estereotipos y limitar la exposición de estudiantes a diversos campos. Los algoritmos de aprendizaje personalizado también pueden verse impactados por las limitaciones de los sistemas de IA, impidiendo el acceso a la información más relevante para los intereses del estudiante.

- 1.7 Si bien a nivel global se observa un crecimiento del mercado de servicios de Inteligencia Artificial Responsable (RAI)²⁴, existe una brecha significativa en la capacidad especializada para startups, empresas en expansión y MiPyMEs en ALC, lo que impide apreciar los beneficios de la IA. El acceso limitado a recursos, capacitación y educación en IA²⁵ profundiza estas brechas, siendo necesario identificar fórmulas costo efectivas que permitan el desarrollo de herramientas, su adopción masiva y formación de talento en IA responsable²⁶, que sea accesible especialmente para las MiPyMEs que tradicionalmente enfrentan desafíos para la adopción de nuevas tecnologías ya sea por falta de conocimiento, disponibilidad de capacidades o desconocimiento de los procesos de apoyo disponibles²⁷. Por ejemplo, aunque existe gran variedad y cantidad de información sobre la IA responsable de última generación, ésta se encuentra diseminada en múltiples repositorios y sin un acompañamiento para las empresas que quieran implementarla que les indique qué líneas de trabajo son las más adecuadas técnica, regulatoria y éticamente para sus casos de uso concretos²⁸.
- 1.8 La iniciativa fAIr LAC del Grupo BID, basada en los Principios de la OCDE de IA responsable²⁹, fue diseñada, precisamente, para interrumpir esta tendencia y apoyar a la región para adoptar un enfoque de desarrollo y adopción responsable de la IA que sea benéfica para todos. Mediante el impulso de proyectos piloto ya ejecutados –en México, Chile, Colombia y Costa Rica que, desde 2019, ha fortalecido actores del sector público, privado, emprendedores y startups de la región– y el desarrollo de herramientas de autoevaluación orientadas a startups e inversionistas se ha buscado desarrollar capacidades en los actores del ecosistema que les permitan desarrollar, adoptar o financiar soluciones y procesos basados en IA incorporando criterios de ética y responsabilidad. Las principales lecciones aprendidas de estas experiencias son: (i) Al ser la IA un tema pionero para el cual no existen regulaciones, normativas o experiencia previa, su abordaje debe ser

²⁴ <https://www.grandviewresearch.com/industry-analysis/ai-governance-market-report>

²⁵ <https://www.oecd-ilibrary.org/sites/e7a0fd6-en/index.html?itemId=/content/component/e7a0fd6-en>

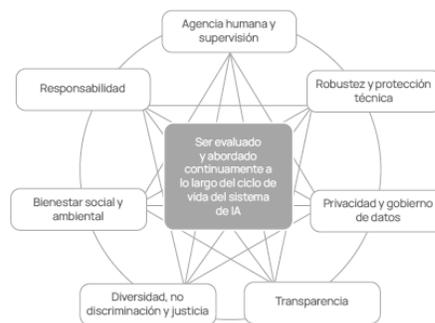
²⁶ <https://impact.economist.com/perspectives/technology-innovation/investment-ai-booming-latin-america-what-will-it-mean-regions-economy>

²⁷ <https://www.eesc.europa.eu/es/our-work/publications-other-work/publications/boosting-use-artificial-intelligence-europes-micro-small-and-medium-sized-enterprises>

²⁸ Por ejemplo, la biblioteca de Holistic.ai es un buen ejemplo de acompañamiento técnico, pero una empresa usuaria de IA o nueva desarrolladora puede tener dificultades para entender cuáles de los casos descritos son más aplicables, o cuál es el impacto potencial que tiene sobre las poblaciones a las que afecta. <https://www.holisticai.com/blog/measuring-and-mitigating-bias-using-holistic-ai-library>

²⁹ <https://oecd.ai/en/ai-principles>

flexible y gradual, de manera que los actores implementen estos principios bajo un esquema de mejora continua y de regulación temprana; (ii) en este mismo sentido, es fundamental continuar generando capacidades en la materia, tanto en los desarrolladores, como los usuarios, reguladores y actores del ecosistema en general. Lo anterior debido a que existen diferentes interpretaciones respecto a qué constituye la IA responsable, siendo importante integrar no solo los elementos tecnológicos, sino también la visión humanista que considere el efecto de estos sistemas en las organizaciones y la vida de las personas, de manera que se pueda apreciar el verdadero potencial que ofrece la IA para mejorar la calidad de vida; (iii) es importante generar procesos participativos, diversos y multiactor en los cuales se pueda reflexionar y dialogar, con lenguaje sencillo, el por qué y para qué analizar y mejorar los sistemas de IA bajo una perspectiva de ética y responsabilidad; (iv) al trabajar programas de innovación es muy importante obtener compromiso del liderazgo como de los cuadros técnicos de las entidades, lo cual mejora la perspectiva de ejecución y monitoreo de las experiencias piloto; (v) para un adecuado desarrollo y adopción de IA ética es fundamental concientizar en la importancia de incorporar el monitoreo de aspectos éticos en las herramientas de gestión de proyectos; (vi) en el desarrollo de herramientas y productos que ayudan a operacionalizar la IA responsable, es deseable que se consideren procesos participativos que faciliten su posterior adopción.



Principios OCDE para la gestión responsable de la IA

- 1.9 Estos desafíos han de ser planteados desde un enfoque de oportunidad. La humanidad ya se ha enfrentado a nuevas tecnologías con bondades y riesgos que se han resuelto en favor del avance de la sociedad. Para ello, es importante entender la oportunidad de impulsar el desarrollo y adopción de la IA, de manera responsable, en MiPyMEs y startups. Este es un tema incipiente y aún no incorporado en regulaciones vigentes de los países en la región³⁰. Por ello, las empresas innovadoras pueden abrir camino y sentar las bases de un mercado tecnológico que permita una mayor competitividad e internacionalización³¹. Pueden además aprovechar la creación de nuevas categorías de mercado (como las propias herramientas de gestión y monitorización de la IA), premiando adicionalmente a las soluciones inclusivas y seguras, y aprovechando las oportunidades de la tecnología, al tiempo que gestionan de manera adecuada sus riesgos. El trabajo con inversores (banca multilateral, fondos de capital de riesgo, inversores ángeles) resulta fundamental para incentivar a las empresas desarrolladoras y usuarias de IA a adoptar enfoques de IA responsable cuando buscan financiamiento, lo cual es especialmente relevante si se

³⁰ <https://www.infobae.com/tecnologia/2023/06/02/que-se-esta-haciendo-en-latinoamerica-para-regular-la-inteligencia-artificial/>

³¹ La 4ª edición del informe de Economía Digital en España, de Adigital (pg. 34), muestra cómo las empresas con mayor cumplimiento en políticas de sostenibilidad y ESG tienen un coste de capital menor, lo cual, unido al mayor nivel de cumplimiento de estas empresas, puede afectar positivamente a su competitividad. https://www.adigital.org/doc/202404_informe-economia-digital.pdf

considera que buena parte del capital de riesgo en la región se enfoca en la financiación de soluciones tecnológicas en sectores de alto impacto social como los servicios financieros³² a través de modalidades como el scoring crediticio alternativo³³. Estos esfuerzos requieren ser apalancados con herramientas y marcos de acción robustos, junto con un aumento de la capacidad de medir los resultados buscados, que permitan alcanzar una audiencia más amplia de startups, empresas en expansión, MiPyMEs, inversores y entidades del ecosistema, allanando el camino para una adopción responsable y equitativa de la IA que impulse el ecosistema de innovación tecnológica regional de la IA, posicionándolo en el escenario internacional, al tiempo que se mitigan los riesgos de la IA y se protege a los más vulnerables. En definitiva, el impulso de la IA responsable por parte de empresas pioneras sirve de faro y espejo al resto del tejido empresarial y social para la puesta en práctica de sistemas de IA con adecuados niveles de control de la privacidad, equidad o transparencia.

II. LA PROPUESTA DE INNOVACIÓN

A. Descripción del proyecto

- 2.1 La utilización de IA ha de ser responsable por diseño³⁴ para que la competitividad del tejido empresarial se produzca y mantenga³¹³⁵³⁶. Esta adopción ha de ser realizada por todo el ecosistema, pero en una primera etapa se requieren empresas pioneras que actúen como modelo a seguir para el resto. Además, acompañar de forma temprana a empresas con impacto directo sobre población vulnerable sirve de ejemplo de acomodación de IA responsable en entornos críticos. Por último, promover la IA responsable en entornos de alta innovación, como la comunidad emprendedora y sus redes, como la de agentes de inversión, crea una enorme oportunidad para potenciar la cultura ambiental, social y de gobernanza empresarial y la vocación de impacto de estas comunidades.
- 2.2 El objetivo del proyecto es contribuir a la mejora de la competitividad de las empresas de ALC, especialmente startups y MiPyMEs, mediante la adopción de estándares de IA responsable, para su desarrollo, su uso y el fomento de una economía digital inclusiva. Es por ello que la iniciativa se denomina **RAI 4 Growth**³⁷, ya que busca potenciar el ecosistema regional como referente global en el desarrollo de IA ética, responsable y competitiva, empoderando y fortaleciendo la propuesta de valor de las empresas (MiPyMEs y startups) y entidades de apoyo al emprendimiento (fondos de inversión, incubadoras, aceleradoras, gremios empresariales, etc.) con herramientas, recursos y capacidades. El proyecto buscará que la tecnología de IA desarrollada y utilizada en LAC no solo esté a la vanguardia de la innovación y la transformación productiva a través de la digitalización, sino que también fomente un despliegue equitativo, accesible y alineado con el bien social. Esto se

³² <https://es.statista.com/grafico/30340/inversion-de-capital-riesgo-en-inteligencia-artificial-en-america-latina/>

³³ <https://www.forbesargentina.com/columnistas/el-impacto-inteligencia-artificial-acceso-credito-inclusion-financiera-n50655>

³⁴ Responsabilidad por diseño: Enfoque que integra principios éticos y de responsabilidad desde el inicio del diseño y desarrollo de sistemas de IA. Este principio asegura que la equidad, la transparencia y la rendición de cuentas, entre otras, sean características predeterminadas, minimizando sesgos, promoviendo explicabilidad y garantizando un uso ético y responsable de la IA.

³⁵ <https://www.vttresearch.com/en/news-and-ideas/ethical-artificial-intelligence-competitive-advantage>

³⁶ Why Addressing Ethical Questions in AI will Benefit Organizations. Capgemini Research Institute. <https://www.capgemini.com/insights/research-library/why-addressing-ethical-questions-in-ai-will-benefit-organizations/> Encuesta a 1.580 ejecutivos de 510 organizaciones y a 4.400 consumidores, donde se muestra relación entre confianza de los consumidores e interacciones éticas con sistemas de IA.

³⁷ Inteligencia Artificial responsable para el crecimiento (por sus siglas en inglés: *Responsible Artificial Intelligence for Growth*).

- logrará mediante la creación de un marco de IA responsable, que incluirá formación, herramientas, certificaciones y planes de medición especializados según la tipología de la solución, beneficiando especialmente a la población más vulnerable.
- 2.3 El proyecto buscará identificar aquellas MiPyMEs, startups o entidades inversoras con interés en el uso y desarrollo de IA responsable, que sirvan como referente y ejemplo al resto del ecosistema. Asimismo, RAI 4 Growth realizará actividades específicas para robustecer con elementos de IA responsable soluciones basadas en esta tecnología que están impactando a poblaciones vulnerables, poblaciones excluidas (por ejemplo por motivos de género) o impactos climáticos.
- 2.4 RAI 4 Growth tendrá tres componentes principales (diálogo, formación y herramientas) y será desplegado inicialmente en Chile, Colombia, Ecuador, México y Uruguay. Estos países -salvo Ecuador- cuentan con una Estrategia Nacional de IA, han implementado proyectos piloto de fAIr LAC y generado redes de actores comprometidos con la temática de la IA responsable, a la cual se busca dar expansión y continuidad con el presente proyecto. Ecuador, por su parte, cuenta con un grupo emergente de startups basadas en IA³⁸, por lo que se considera terreno fértil para implementar y probar las herramientas desarrolladas en un contexto de desarrollo incipiente.
- 2.5 Para maximizar el impacto positivo del proyecto, se buscará que las herramientas, recursos y el desarrollo de capacidades que formarán parte de las actividades del proyecto, y que estarán disponibles para el ecosistema en general, sean adoptadas por empresas y entidades de apoyo comprometidas con la mejora de su competitividad a través de la IA responsable por diseño, formando esto parte de sus objetivos ambientales, sociales y de gobernanza, y por aquellas que trabajen en ámbitos de alto impacto sobre poblaciones vulnerables³⁹, mujeres y minorías, de manera que se puedan verificar los beneficios de la mitigación de riesgos de la IA, especialmente en lo referido a sesgos, exclusión y discriminación.
- 2.6 La iniciativa buscará consolidar un ecosistema propicio para el desarrollo y adopción responsable de la IA que mejore la competitividad de los negocios sin poner en riesgo la inclusión, equidad y derechos de las personas entregando herramientas accionables, formación, conocimiento y sensibilización sobre la temática. Las bases fundamentales de este desarrollo son las herramientas ya probadas por fAIr LAC y Adigital, las que serán consolidadas y mejoradas tanto en su integración como en la experiencia de uso por los usuarios.



RAI 4 Growth - Componentes

³⁸ <https://www.f6s.com/companies/artificial-intelligence/ecuador/co>

³⁹ Tales como: (i) inclusión Financiera; (ii) plataformas educativas; (iii) plataformas de salud; (iv) plataformas de empleo; (v) plataformas de sostenibilidad climática; (vi) plataformas que incorporan y acreditan técnicas para la mitigación de sesgos.

2.7 **Innovación.** El proyecto destaca por su carácter innovador al ser un programa integrado e integrador que se establece como herramienta para fomentar la competitividad de las empresas regionales, que, a su vez, impactará en que las empresas y organizaciones participantes estén concienciadas en evitar sesgos, incluidos aquellos que hoy en día puedan estar generando un impacto negativo en los grupos en condición de exclusión, pobreza y vulnerabilidad que interactúan con tecnologías de IA. El diseño del programa enfatiza un enfoque holístico que no solo adopta tecnologías de IA, sino que también promueve su uso ético y sostenible dentro del ecosistema empresarial regional y la medición de resultados. La iniciativa también es pionera en la articulación de esfuerzos dispersos en el impulso de prácticas de IA responsable, mediante un programa de apoyo empresarial alineado con estándares internacionales y tendencias regulatorias, preparando a las empresas para competir en un escenario global. Se desarrollará un mix de actividades virtuales y presenciales, buscando apalancar esfuerzos ya existentes en los ecosistemas y la adopción de las herramientas a ser desplegadas.

2.8 El proyecto no solo se concentrará en implementar tecnología, sino también en desarrollar el capital humano necesario para su manejo responsable y efectivo, mediante la formación especializada y la creación de recursos que fomenten una cultura de innovación, integrando capacitación, auditorías y certificaciones, respetando valores y normas democráticas, así como promoviendo la inclusión social y económica, estructurándose en los siguientes componentes:

Componente I: Generación de confianza (Contribución BID Lab US\$ 332.920, Contrapartida US\$ 178.090).

2.9 El objetivo del componente es promover la concientización del ecosistema de innovación tecnológica regional sobre la importancia de la IA responsable para el desarrollo de tecnologías justas, transparentes y fiables, que resulten en el desarrollo de productos técnicamente robustos y competitivos y conozcan las tendencias regulatorias en la materia. Esto se logrará a través de actividades llevadas a cabo por el OE que fomenten la difusión de información y el intercambio entre los diferentes actores del ecosistema, incluyendo principalmente a empresas, inversores, gobiernos y reguladores, así como la sociedad civil y la academia. Se pondrá especial atención en el aumento de la competitividad del tejido empresarial de MiPyMEs y startups de la región. Además, se realizarán acciones de concientización sobre las afectaciones a los grupos vulnerables o en condición de pobreza, en línea con políticas y regulaciones sectoriales existentes en materia de no discriminación, así como regulaciones aplicables en otras regiones que puedan ser parte de los objetivos de mercado de las empresas (startups y MiPyMEs) participantes. Para ello, se realizarán las siguientes actividades:

- (i) Campañas de sensibilización y concientización de la RAI resaltando las oportunidades y riesgos de la IA, introduciendo temas relevantes al contexto de la región LAC, en dimensiones tanto relativas al apoyo a la competitividad de las empresas como en aquellas que afecten a poblaciones vulnerables, consideraciones de género y sostenibilidad climática.
- (ii) Elaboración de Informes (como el índice de la IA responsable en el ecosistema de innovación “fAIr Tech Radar”) y sesiones de diálogo y reflexión amplia (sector público, privado y sociedad civil) para dar a conocer mejores prácticas, el valor de la incorporación del uso responsable de la IA en los negocios, impacto en poblaciones excluidas y en condiciones de pobreza y vulnerabilidad en sectores en los que la IA va cobrando relevancia, tales como por ejemplo: (a) aplicaciones de IA en el sector salud, riesgos de exclusión de poblaciones vulnerables, y medidas

de mitigación; (b) brecha digital y exclusión del sistema financiero; (c) tecnologías educativas y sesgos de género.

- (iii) Acciones de monitoreo que permitan relevar la efectividad de las actividades del componente, tales como (a) registro y asistencia a los eventos y sesiones de trabajo; (b) valoración de la calidad de las sesiones por los participantes; (c) impacto en medios, redes sociales y descargas de los informes publicados; (d) impacto cualitativo de sesiones e informes a partir de conversaciones con roles relevantes de industria, sociedad y política pública.

- 2.10 Como resultado, se espera que las entidades del ecosistema regional estén más informadas y alineadas con políticas de RAI, regulaciones nacionales e internacionales, preparadas para el cumplimiento regulatorio e identifiquen oportunidades de colaboración para el desarrollo y adopción de una IA más inclusiva, transparente y que fomente la confianza y, por tanto, la competitividad de las empresas. Específicamente, se espera: (a) convocar al menos 5000 participantes en las diversas campañas y eventos implementados; (ii) contar con material comunicacional con directrices y prácticas responsables de desarrollo y uso de la IA; (iii) sistematizar mejores prácticas RAI en LAC y comunidad colaborativa en torno a ellas; (iv) mejorar el conocimiento y percepción pública sobre la RAI en los países de intervención.

Componente II: Desarrollo de habilidades y talentos del ecosistema para el desarrollo responsable de la IA (Contribución BID Lab US\$ 374.210, Contrapartida US\$ 504.480).

- 2.11 Tanto las empresas que usan IA como las que implementan modelos y sistemas han de entender qué significa implementar IA responsable. De la misma manera que la "privacidad por diseño"⁴⁰ requirió tiempo y esfuerzo hasta que se entendieron cuáles eran los mínimos imperativos, la "responsabilidad por diseño" requiere comprender las bases éticas, regulatorias, tecnológicas y de gobernanza, complementándose con casos de uso de empresas pioneras en su implementación. El objetivo del componente es construir el principal canal de entrega de herramientas y certificaciones, enfatizando la creación de programas de upskilling con contenido especializado en riesgo de la IA e IA responsable para capacitar a emprendedores, MiPyMEs, inversionistas, entidades articuladoras y sus equipos. Esto se logrará a través de la oferta de conocimiento desarrollada ya por Adigital y la oferta modular de fAIr LAC diseñada para incorporar contenido de IA responsable en programas existentes de incubación o aceleración a empresas de IA ("AI-first o AI-enabled"). En el marco de este componente, el OE implementará las siguientes actividades, bajo un enfoque iterativo de mejora continua:

- (i) Desarrollo de módulos de capacitación en IA responsable a ser implementado en colaboración con entidades aliadas de apoyo al emprendimiento y de manera directa, dirigidos a startups, MiPyMEs e inversores. El formato podrá ser presencial, híbrido, MOOC o seminarios especializados, siendo las temáticas entre otras: Introducción a la Inteligencia Artificial Responsable, Sesgos en la Inteligencia Artificial, Algoritmos Explicables y Transparencia en IA, IA para la Inclusión de Personas en el Tejido Socioeconómico, IA Responsable con foco en Perspectiva de Género, IA Responsable para el Desarrollo Sostenible, Fundamentos de la IA Responsable para Inversores, Gobernanza y Regulación de la IA para Inversores, Evaluación de Riesgos y Oportunidades en Startups de IA, Evaluación de

⁴⁰ Privacidad por diseño: Enfoque que integra la protección de datos personales desde el inicio del diseño y desarrollo de sistemas, asegurando que la privacidad sea una característica predeterminada. Este principio minimiza la recolección de datos, promueve la anonimización y garantiza el control del usuario sobre su información.

Proveedores de IA para MiPyMEs, Análisis Crítico de Sistemas de IA, Herramientas de IA de Código Abierto para MiPyMEs. Durante la implementación se podrán identificar otras temáticas relevantes, en función del estado de madurez de los mercados.

- (ii) Programa de formación de formadores, orientado a dar sostenibilidad a los esfuerzos de generación de capacidades en startups, MiPyMEs e inversores, enfocándose en conceptos clave, aplicaciones prácticas y la evaluación crítica de soluciones de IA responsable. Los módulos formativos buscarán un equilibrio entre conceptos teóricos y su aplicación práctica, e incluirán un componente pedagógico para enseñar técnicas efectivas de enseñanza para adultos. La convocatoria de los participantes de esta oportunidad de formación se realizará de manera abierta, apoyado en las redes de colaboración de RAI 4 Growth. Será deseable que los participantes cuenten con formación previa en desarrollo de tecnología.
 - (iii) Mentoría especializada. Con base en la demanda, los docentes y expertos participantes en los programas de formación podrán realizar mentorías a las entidades participantes con el fin de darles orientación concreta respecto a su camino hacia la IA responsable, de manera que se preparen adecuadamente frente a procesos de certificación, auditoría, cumplimiento legal o regulatorio y el uso de las herramientas/plataforma de servicios desarrolladas en el componente III.
 - (iv) Piloto de capacitación RAI con startups y MiPyMEs. Se desarrollará un piloto de los programas centrales de capacitación de IA responsable que permita validar la calidad, interés e impacto en las organizaciones.
 - (v) Oferta abierta de los programas de capacitación sobre la base de las mejoras y aprendizajes de la experiencia piloto, para lo cual se: (i) contratará una plataforma de aprendizaje complementaria a las acciones de formación presencial; (ii) realizarán eventos y campañas de marketing directas o con el apoyo de los colaboradores académicos y de formación; y (iii) mantendrá un espacio denominado “RAI Academy” en la plataforma latinoamericana de IA responsable, en la cual se brindará soporte técnico y académico a los participantes.
 - (vi) Monitoreo de resultados de las formaciones tanto en cuanto a los aprendizajes logrados como de la experiencia de los participantes y de la implementación práctica de los aprendizajes a 6 o 12 meses.
 - (vii) Modelo de Sostenibilidad y escala de los programas de capacitación a través de “RAI Academy” que ofrecerá, a empresas, instituciones y personas interesadas, formación a costos razonables. Se definirán algunos módulos de libre acceso (freemium), módulos de pago y esquemas de mentoría especializada para la implementación de IA responsable.
- 2.12 Como resultado del componente se espera contar con: (i) contenido sistematizado para la formación RAI que incluya principios éticos de la IA, gobernanza de datos, mitigación de sesgos, prácticas de sostenibilidad ambiental y cumplimiento de regulaciones emergentes e incorpore estudios de casos y ejercicios prácticos; (ii) alianza con entidades académicas, expertos de la industria y entidades reguladoras para el despliegue del plan de estudios; (iii) guías sistematizadas para formadores de plan de estudios RAI; (iv) formadores del plan de estudios RAI certificados; (v) piloto de formación RAI dirigido a startups, pymes a cargo de formadores entrenados; (vi) sistematización de mejoras y retroalimentación continua para la sostenibilidad del programa tanto en sus aspectos institucionales, como técnicos y financieros; (vii) reportes sistematizados de aprendizaje, actitudes y prácticas RAI de los

participantes en el programa; (viii) portafolio validado de temáticas de formación avanzada para expertos RAI.

Componente III: Herramientas, Servicios y Certificaciones: Recursos para la gestión de riesgos de IA responsable (Contribución BID Lab US\$ 853.010, Contrapartida US\$ 817.430).

- 2.13 El objetivo de este tercer componente es crear e impulsar herramientas, servicios y certificaciones que ayuden a las empresas e inversores a detectar riesgos e implementar procesos simplificados de medidas que los mitiguen, rastreen y corrijan para el desarrollo y uso responsable de la tecnología. Para ello se contará con un equipo multidisciplinar que ofrecerá recomendaciones concretas para la mitigación y monitoreo de riesgos identificados, realizará pruebas algorítmicas, evaluaciones de impacto, y desarrollará herramientas y certificaciones para el uso responsable de la IA, incorporando siempre que sea posible la mitigación de impactos negativos y obtención de impactos positivos para las poblaciones vulnerables (desde los puntos de vista económico y de exclusión), fruto del despliegue de la IA, con perspectiva de género y capacidades para la medición y compensación de la huella de carbono de los sistemas que utilizan modelos de IA.
- 2.14 El proyecto comienza con herramientas existentes provistas por BID Lab y Adigital, utilizadas para que RAI 4 Growth sea ejecutivo desde el primer día. Mediante las actividades descritas más abajo y de manera coordinada con las empresas participantes, se crearán nuevos servicios y certificaciones. Sin embargo, crear herramientas per se no es suficiente. Adigital, dentro de RAI 4 Growth, plantea la puesta en marcha de un marco (framework) de IA responsable que ayude a las MiPyMEs, startups, etc., a implementar la IA responsable de manera efectiva y con impacto.
- 2.15 El marco de IA responsable a desarrollar considera un enfoque integral y avanzado que, a partir de la relación constante con la industria, el entorno académico, la sociedad y los reguladores, asegurará que los sistemas de IA se desarrollen y utilicen de manera ética, transparente y equitativa. Para ello, se integrarán diversos recursos y prácticas que garanticen, en cada etapa del ciclo de vida del desarrollo de IA, el cumplimiento con los más altos estándares de responsabilidad y ética, generando de esta forma un estándar regional para startups y MiPyMEs que usan y/o desarrollan IA y una herramienta de alto valor agregado para fondos de inversión, banca y cualquier institución que ha de tomar decisiones sobre el uso o financiamiento de la IA.
- 2.16 Los principales elementos del marco de IA responsable de RAI 4 Growth son los siguientes: (i) **Herramientas**: Recursos técnicos avanzados y guías metodológicas para facilitar la implementación de prácticas responsables en todas las etapas del ciclo de vida del desarrollo de IA, tales como algoritmos que promueven la transparencia, software para la detección y mitigación de sesgos, servicios para asegurar la equidad en los datos, etc.); (ii) **Certificaciones**: Programas de certificación que validan la conformidad de los sistemas de IA con mejores prácticas, regulaciones y estándares internacionales de ética y responsabilidad y que abarcan aspectos críticos como la transparencia, la explicabilidad, la privacidad de datos, la equidad algorítmica o la rendición de cuentas; (iii) **Metodologías y Mejores Prácticas**: Basadas en la investigación y el consenso de expertos para garantizar el desarrollo y la implementación responsable de IA. Incluye guías de diseño ético, procedimientos para la evaluación de impacto social, estrategias para la inclusión y participación de las partes interesadas, y protocolos para la gestión de riesgos y mitigación de daños potenciales; (iv) **Evaluación y Auditoría Continua**: para asegurar que los sistemas de IA se mantengan alineados con los principios de responsabilidad y ética a lo largo del tiempo. Esto incluye revisiones periódicas, auditorías independientes y

mecanismos para la retroalimentación constante de usuarios y otras partes interesadas; (v) **Plan de Medición:** Conjunto de planes estructurados y detallados que establecen métricas e indicadores clave para evaluar cómo un sistema de IA aplica los principios de IA Responsable y miden su impacto en relación con objetivos específicos como la equidad de género, el apoyo a poblaciones vulnerables y la sostenibilidad.

2.17 Para el desarrollo del componente, el OE llevará a cabo las siguientes actividades:

- (i) Primeras herramientas y certificaciones alineadas con las herramientas de fAIr LAC (Autoevaluación Ética para emprendedores y Herramienta de Valoración de Riesgo para inversionistas) y con los requerimientos del Certificado de Transparencia de Adigital⁴¹. Las herramientas de fAIr LAC y el certificado de Adigital tienen el potencial de componer una suite alineada que englobe desde la autoevaluación hasta la evaluación evidenciada del uso y desarrollo transparente de sistemas de IA. Para ello, se fortalecerá el soporte tecnológico de las herramientas de fAIr LAC y se realizará una integración y normalización de éstas con el Certificado de Adigital. El desarrollo y mantenimiento se realizará de la mano de expertos y se validará con pruebas piloto que podrán contar con un grupo representativo de startups y MiPyMEs que ya han implementado estas herramientas en sectores de alto impacto socioeconómico.
- (ii) Portafolio de servicios a empresas. Se crearán nuevas herramientas complementarias en áreas de IA responsable de impacto o que permitan analizar sectores o ámbitos relacionados con poblaciones vulnerables y perspectiva de género. Específicamente se desarrollarán: (a) nueva certificación de análisis de sesgos de género a partir del certificado de transparencia algorítmica de Adigital, con base en el principio ético de equidad/justicia algorítmica y al estado del arte de la implantación de capacidades tecnológicas; (b) nueva certificación de análisis de huella de carbono a partir del certificado de transparencia algorítmica de Adigital; (c) caja de herramientas de medidas de mitigación de sesgos de género; (d) caja de herramientas para la disminución de huella de carbono de sistemas de IA que incluyan consideraciones respecto a la selección de modelos, su entrenamiento y puesta a disposición.
- (iii) Herramientas de terceros. Sumado a lo anterior, se identificará un pool de herramientas existentes que atiendan de forma gradual los requisitos emanados de la identificación de riesgos y análisis para la certificación RAI, en función del nivel de madurez de las empresas, su desarrollo tecnológico y el nivel de riesgo de sus soluciones. De esta forma se contará con modelos orientados a facilitar: (a) Evaluaciones de conformidad y autoevaluaciones, (b) Auditorías y pruebas de sesgo y precisión de los algoritmos y (c) Evaluaciones de impacto algorítmico. A continuación, se detalla una lista no exhaustiva de grupos de herramientas y técnicas que podrán ser desplegados por el proyecto

Herramienta de IA Responsable	Descripción
Evaluaciones de conformidad y autoevaluaciones	Valoraciones iniciales para la identificación de riesgos y el establecimiento de medidas de mitigación. <ul style="list-style-type: none"> - Gobernanza de datos y privacidad - Transparencia y explicabilidad - Justicia y no discriminación

⁴¹ <https://www.transparenciaalgoritmica.es/>

	<ul style="list-style-type: none"> - Precisión y desempeño de los modelos - Seguridad, confianza y robustez de la IA - Rendición de cuentas
Auditorías y pruebas de sesgo	<p>Herramientas y procesos en profundidad para evaluar el rendimiento y la equidad de los sistemas de IA, asegurando que no presenten sesgos o comportamientos no deseados.</p> <ul style="list-style-type: none"> - Valoraciones técnicas del despliegue de operaciones de IA/ML (en sus fases de entrenamiento e inferencia) - Valoraciones de las bases de datos - Arquitectura de datos y modelos de IA - Despliegue e integración - Detección de sesgos de los modelos - Mantenimiento y actualización - Entrenamiento y formación de personal
Evaluaciones de impacto algorítmico	<p>Medidas de evaluación del impacto social del uso de modelos de IA y su influencia en la toma de decisiones.</p>

Tabla 1. Herramientas de IA responsable

(iv) Planes de medición (general y específicos para sistemas de IA centrados en población vulnerable, mitigación de sesgos de género y sostenibilidad). Estos planes se componen de (a) definición de indicadores clave que, a partir del uso de las herramientas y certificaciones provistas en RAI 4 Growth, permitan evaluar aspectos relevantes para sistemas de IA responsable relacionados con la eficacia del sistema, pero también con la equidad, la transparencia, la mitigación de diferentes tipos de sesgos o la privacidad; (b) procedimientos y herramientas específicas para recopilar y analizar datos relevantes que reflejen el desempeño del sistema en términos de responsabilidad; (c) periodicidad con la que se deben realizar las diferentes mediciones y evaluaciones; (d) pautas para el establecimiento de metas claras y criterios específicos que el sistema de IA debe alcanzar para ser considerado responsable; (e) mecanismos para documentar los resultados de las mediciones y comunicar los hallazgos a las partes interesadas de manera transparente y accesible; (f) guías, demostraciones y casos de uso basados en experiencias reales de RAI 4 Growth que apoyen la implantación de planes de medición en las empresas beneficiarias. Algunos planes de medición previstos son:

- Plan de medición general. Válido para cualquier tipo de sistema de IA, indica cómo utilizar las herramientas más comunes de IA responsable y aquellos indicadores más adecuados.
- Plan de medición de sistemas IA de mitigación de sesgos de género. También a partir del plan general, se determina si el sistema opera de manera equitativa y libre (o en proceso de liberación) de discriminación de género. Incluye indicadores que miden la presencia y la corrección de sesgos de género en los datos y en los resultados del sistema, la transparencia en la toma de decisiones, y la equidad en el impacto del sistema sobre diferentes géneros en base a los objetivos del sistema.
- Plan de medición de sistemas de IA sostenibles. Definido para evaluar y garantizar que miden su impacto ambiental y cómo su diseño permite la minimización de su impacto ambiental y la promoción de prácticas

sostenibles. Al igual que en los casos anteriores, el plan explica cómo hacer uso de las herramientas disponibles en RAI 4 Growth para su correcta implementación.

- (v) **Servicios.** Derivado del uso de las herramientas, sobre todo de las autoevaluaciones y evaluaciones de conformidad⁴², las empresas recibirán una serie de recomendaciones accionables para el alineamiento de sus sistemas con principios de IA responsable⁴³ o la operacionalización de dichos principios que en algunos casos requerirán de apoyo experto para llevarse a cabo. Algunos ejemplos de la tipología de servicios podrán ser: revisión de bases de datos, políticas de gobernanza de datos, selección y evaluación de modelos de IA, entrenamiento y auditoría de modelos, despliegue e integración, mantenimiento y actualización, entre otros. Este portafolio de servicios será disponibilizado mediante un panorama de IA responsable que visibilizará las entidades especializadas que cuentan con capacidad técnica y operativa para llevar a cabo medidas de monitoreo y corrección, el cual evolucionará a la creación de la “Plataforma Latinoamericana de IA Responsable”, desarrollada de manera que se adapte a las necesidades de este naciente mercado. Para la identificación de las entidades participantes de esta plataforma, se realizarán llamados abiertos, competitivos y transparentes, incorporando criterios de integridad institucional, capacidad operativa y técnica y de diversidad de oferta y para los cuales se podrán establecer alianzas locales con el fin de mejorar su visibilidad y difusión. La plataforma tecnológica corresponderá inicialmente a una web que ofrezca información, formación y acceso a recursos de IA responsable, evolucionando, junto a la evolución y madurez del mercado, a una plataforma de servicios de IA responsable.
- (vi) **Estrategia de comunicación** Además de la implementación propiamente tal de los procesos de certificación y evaluación de IA responsable en entidades articuladoras del ecosistema (inversionistas, aceleradoras, incubadoras, etc.), startups y MiPyMEs que desarrollan o utilizan sistemas de IA, se implementará una estrategia de comunicación que facilite la convocatoria de las entidades participantes que incluirá la gestión de redes, sitio web, material de apoyo, realización de talleres, guías y casos prácticos.
- (vii) **Pilotos.** Se desarrollan pilotos de los procesos de evaluación de riesgos y posterior certificación en sectores de alto impacto para la población vulnerable o de alto riesgo, realizando tanto una validación de calidad de cada proceso, como el impacto que tiene en la organización participante. La **selección de las entidades participantes en las pruebas piloto** se realizará mediante convocatorias abiertas, acuerdos con gremios y asociaciones empresariales o alianzas con entidades que buscan fortalecer cadenas de valor, manteniendo como un criterio importante para ello la incidencia en prácticas de IA responsable en startups y MiPyMEs cuyos usuarios, clientes o beneficiarios incluyen a segmentos de población pobre y vulnerable, mujeres y minorías o cuyos productos y servicios apoyados en IA sean susceptibles de riesgo de sesgo y discriminación. Estos pilotos ayudarán a conformar y validar las primeras versiones del Marco de IA responsable.

⁴² Evaluaciones de conformidad: Procesos que validan el cumplimiento de los sistemas de IA con medidas o requisitos de gestión de riesgos, gobernanza de datos, documentación técnica, mantenimiento de registros, transparencia y provisión de información, supervisión humana, precisión, robustez y ciberseguridad, principalmente.

⁴³ <https://oecd.ai/en/ai-principles>

- (viii) Diseño de modelo de negocios sostenible de RAI 4 Growth, basado en el portafolio de servicios, directos e indirectos, piloteados en el marco de este componente, bajo un modelo *freemium* y *de pago*.
- 2.18 Como resultado de este componente se espera contar con (i) un Marco de IA responsable compuesto por diferentes herramientas, guías y planes de medición; (ii) herramientas de evaluación y certificación alineadas y validadas; (iii) guías y documentos de orientación para startups, pymes y entidades de apoyo sobre el uso de herramientas de IA responsable validadas; (iv) protocolos sistematizados para el despliegue de servicios de asesoramiento centrados en la gestión de riesgos de la IA; (v) planes de medición para diferentes tipologías; (vi) plataforma digital operativa y reportes de tracción de la misma; (vii) 100 entidades de apoyo del ecosistema fortalecidas y con procesos de IA responsable incorporados; (viii) 400 certificaciones/pruebas realizadas con startups y pymes; (ix) reporte de riesgos de IA en sectores sensibles para la población vulnerable de ALC; (x) 3 estudios de caso sobre el impacto positivo de la certificación; (xi) modelo de negocios y estrategia de expansión de los servicios RAI documentados.
- B. Resultados, medición, seguimiento y evaluación del proyecto**
- 2.19 El proyecto está diseñado para apoyar startup, MiPyME, inversor y entidad de apoyo empresarial que desarrolla, utiliza y financia soluciones tecnológicas basadas en IA, con especial atención en sectores de alto impacto socioeconómico tales como: empresas con alto potencial de crecimiento, empresas de salud y cuidados, empleabilidad y contratación, educación y alfabetización, inclusión financiera, además de infraestructuras críticas y esenciales, justicia y asistencia legal, migración, agricultura o seguridad alimentaria.
- 2.20 Los principales resultados del proyecto serán: (i) 250 empresas desarrolladoras de IA y 1000 MiPyMEs utilizan las herramientas de valoración responsable y mitigación de riesgos de IA; (ii) 350.000 mujeres usuarias o beneficiarias de los productos y servicios de empresas participantes que implementan medidas de IA responsable para la mitigación de sesgos de género; (iii) 3.000 personas en situación de vulnerabilidad usuarias o beneficiarias de los productos y servicios de empresas participantes que implementan medidas de IA responsable; (iv) 150 startups y MiPyMEs que incorporan metodologías de medición y estimación de su huella de carbono relacionada con sus sistemas de IA, de las cuales un tercio incorpora estrategias de compensación o mitigación; (v) un 20% de las empresas participantes que buscan inversión y que aplican de manera explícita la IA responsable, obtienen inversión con al menos un inversor cuya tesis de inversión incluye el uso e implementación de tecnologías responsables.
- 2.21 El seguimiento del proyecto, a cargo del OE, se realizará mediante un cuadro de mandos que contendrá los indicadores de resultados, además de parámetros operativos de ejecución del proyecto que permita al organismo ejecutor: (i) Informar y reportar los resultados a BID Lab según el convenio del proyecto RAI 4 Growth, en la plataforma existente para esos fines; (ii) Mantener la información adecuada para realizar análisis estratégico y operativo del proyecto y cada uno de sus componentes, que sirva como insumo para informar a los órganos de gobernanza de la iniciativa; (iii) Informar de manera táctica a los diferentes colaboradores y aliados.
- 2.22 El OE será responsable de recopilar los datos e informar sobre los resultados y logros según la matriz de resultados del proyecto. El OE desarrollará un plan de monitoreo al inicio del proyecto que garantice el seguimiento y la medición de los indicadores. Además, el OE informará al Banco a través de un PSR periódico (Project Status Report) y presentará un PSR Final sobre los resultados del proyecto tras su finalización.

- 2.23 Durante el tercer año de ejecución, se realizará una evaluación sobre los efectos de los pilotos seleccionados y de las tecnologías de IA que adoptan criterios de ética y responsabilidad sobre las poblaciones beneficiarias, en los cuales se valorará si la incorporación de estos criterios disminuye sesgos o discriminaciones en beneficiarios o usuarios. Los términos específicos de la evaluación se formularán en función de las temáticas abordadas por los pilotos y las soluciones de IA que formen parte del programa.

III. ALINEACIÓN CON EL GRUPO BID, AMPLIACIÓN DE ESCALA Y RIESGOS

A. Alineación con el Grupo BID

- 3.1 El proyecto se alinea con la Estrategia Institucional del Grupo BID 2024 - 2030 y sus objetivos estratégicos de reducir la pobreza y desigualdad e impulsar el crecimiento sostenible, mediante el impulso de la iniciativa privada y la adopción responsable de tecnología (documento CA-631).
- 3.2 El proyecto se encuentra alineado con las estrategias de los países de intervención inicial, especialmente en lo referido a: (i) Habilitar la economía del futuro, la mejora de productividad y la inversión en I+D (GN-3140, EBP Chile); (ii) Acceso equitativo a servicios y desarrollo de capital humano (GN-2982, EBP México); (iii) estimular la innovación, el desarrollo empresarial y el uso de herramientas digitales (GN-2972, EBP Colombia); (iv) contribuir al desarrollo del sector productivo y la infraestructura tecnológica (GN-3103, EBP Ecuador); y (v) el fomento a la cultura emprendedora y aumentar la innovación por parte del sector privado (GN-3056, EBP Uruguay).
- 3.3 Asimismo, el proyecto está alineado con el Plan de Negocios de BID Lab (MIF/GN-264) y su estrategia para la IV reposición de capital (GN-3204), que pone la innovación emprendedora en el centro como vehículo para mejorar la calidad de vida de las personas, haciendo accesibles las nuevas tecnologías y promoviendo su desarrollo y adopción responsables. En ese sentido, el proyecto complementa los esfuerzos ya desplegados en la región a través de proyectos piloto de la iniciativa fAIr LAC en línea con la visión de construir un portafolio consistente tanto en “AI for Development” como en la construcción de las bases para un ecosistema de IA responsable en la región. Los proyectos que han sido impulsados por BID Lab son los siguientes: a nivel de la construcción de la infraestructura para el uso responsable de la IA se encuentra el proyecto (i) Algoritmos Éticos CH-T1246, “Compras públicas de algoritmos responsables, éticos y transparentes”, ejecutado por la Universidad Adolfo Ibáñez con el propósito de promover la gestión ética de los datos y el desarrollo e implementación responsables de algoritmos, sistemas de decisión automatizada e inteligencia artificial, tanto en el sector público como en el mundo privado; (ii) fAIr Jalisco ATN/ME-17550-ME, “Servicios sociales eficientes a través de IA responsable”, ejecutado por el Tecnológico de Monterrey para aprovechar la adopción responsable y ética de la Inteligencia Artificial para el bien social en Jalisco, a través de la articulación de academia, sociedad civil, sector público y sector privado para desarrollar capacidades, potencializar el alcance social de la IA y aprovechar el impacto de la tecnología para servicios sociales más eficientes. En materia del uso de la IA para el bien común, la Convocatoria en Género e IA lanzada por BID Lab en 2022 dio origen a tres proyectos que utilizan IA para mitigar la brecha de género en el ámbito laboral, la inclusión financiera y el uso de algoritmos: (iii) Jobecam BR-G1021 “IA para impulsar la diversidad en el mundo laboral”, es una plataforma para reducir sesgos en las contrataciones; (iv) Quipu CO-G1050, “IA para la inclusión financiera de mujeres vulnerables” a través de un scoring crediticio alternativo y justo; y (v) Quantil CO-G1049, “Auditorías algorítmicas para

el desarrollo de IA responsable”, a través de la automatización de las evaluaciones de sesgo de los algoritmos.

- 3.4 A nivel Grupo BID, BID Lab, a través de la iniciativa fAIr LAC ha coordinado estrechamente su trabajo con el sector público a través de las siguientes cooperaciones técnicas: (i) fAIr Costa Rica a través del proyecto CR-T1242 “Apoyo al marco orientador para el uso responsable y ético de la Inteligencia Artificial en Costa Rica” para la generación de capacidades en materia de uso responsable de la tecnología, el fomento de un ecosistema dinámico y el empoderamiento ciudadano para impulsar la industria 4.0; (ii) RG-T3638 “Diseño y ejecución de Proyectos Piloto en sectores sociales bajo la Iniciativa fAIr LAC”; y (iii) RG-T3450 “Fair LAC - Inteligencia Artificial Responsable para una Prestación más Eficiente e Individualizada de Servicios Sociales para Todos”, a través de la cual se impulsaron acciones en Uruguay y Colombia. En el caso de Uruguay, se llevó a cabo la cooperación técnica para promover y fortalecer el uso responsable de la inteligencia artificial, particularmente en la administración pública a través de la implementación de la Estrategia de IA para el Gobierno Digital. La estrategia fue presentada a consulta pública, se recibieron aportes, se analizaron y fueron respondidos, y en algunos casos incorporados al texto final de la versión vigente de la Estrategia IA. En el caso de Colombia, la iniciativa se desarrolló con el objetivo de promover un alto impacto en la generación de políticas públicas y proyectos de transformación digital dentro de las entidades públicas del país en las que se utilizan sistemas de IA. Asimismo, se fomentó el acceso y divulgación de conocimiento para la implementación y despliegue de IA en el país y la región. Como resultado de esta iniciativa, se desarrolló la Política Nacional para la Transformación Digital e IA, así como un Marco Ético de IA en Colombia. En este sentido, el proyecto pretende seguir contribuyendo al impacto de la iniciativa con base en el conocimiento y las redes forjadas, así como mediante el fomento de la colaboración público-privada para el desarrollo de un ecosistema de IA responsable.
- 3.5 Alineación con el Acuerdo de París: se considera que el proyecto está alineado con BB1.
- 3.6 Objetivos de Desarrollo Sostenible (ODS). El proyecto contribuirá al: ODS 5, sobre igualdad de género; (ii) ODS 8, sobre acceso a trabajo decente y crecimiento económico; (iii) ODS 9, sobre construir infraestructuras resilientes, promover la industrialización inclusiva y sostenible, y fomentar la innovación. Asimismo, al incidir en el desarrollo y adopción de algoritmos de IA que mitigan potenciales sesgos y discriminación; y (iv) OSD 10, sobre reducción de desigualdades.

B. Adicionalidad y Escalabilidad

- 3.7 El proyecto considera, desde su diseño, la ruta para la sostenibilidad y la escala, mediante el desarrollo de una plataforma de productos, herramientas y servicios que podrán ser adoptadas de manera modular de acuerdo al grado de desarrollo de cada ecosistema, agregando un valor distintivo basado en cinco pilares:
- Fomento de la confianza: Al centrarse en la transparencia, la explicabilidad y la responsabilidad en las operaciones de IA, el proyecto ayudará a construir confianza en las tecnologías entre consumidores, usuarios, inversores y organismos reguladores.
 - Inclusión y equidad: la iniciativa promueve la adopción de la IA responsable entre startups y MiPyMEs en ALC, lo cual mejora su competitividad y beneficiará directamente a poblaciones vulnerables o de baja representatividad al mitigar riesgos de sesgos y discriminación, el uso responsable de datos, el respeto a los derechos de las personas, la transparencia y explicabilidad, la rendición de cuentas, y los riesgos de sustitución laboral por causa de la automatización.

- Desarrollo de talento y capacidades: A través de los programas de formación y desarrollo de herramientas, RAI 4 Growth contribuirá a cerrar brechas de habilidades en la región, permitiendo a las empresas locales competir en un mercado global.
 - Sostenibilidad ambiental: El proyecto tiene un enfoque en la reducción de la huella de carbono de las tecnologías de IA, dando visibilidad y realizando un llamado a la acción al incluir la sostenibilidad como uno de los principios éticos a considerar bajo el enfoque de IA Responsable.
 - Colaboración entre actores y partes interesadas: La colaboración efectiva entre diversos actores del ecosistema, incluyendo empresas, inversores, gobiernos, reguladores, sociedad civil y academia, es fundamental para escalar el impacto del proyecto. Al facilitar el intercambio de conocimientos, experiencias y recursos, se crea un entorno regional propicio para impulsar la adopción de la IA responsable. Esta colaboración estratégica se estructura mediante alianzas, consultorías especializadas, provisión de infraestructura tecnológica y participación conjunta en programas de formación e investigación. La consecución de la sinergia entre estos actores acelera la implementación de soluciones tecnológicas, y permite la distribución de los beneficios, potenciando el desarrollo y la competitividad en diferentes contextos y niveles de madurez tecnológica en la región.
- 3.8 La escalabilidad estará también dada por: (i) la estrategia de generación de alianzas del proyecto con socios estratégicos que aporten valor en diferentes áreas críticas. Estas colaboraciones podrán tomar la forma de consultorías, en las fases de evaluación, certificación de responsabilidad o mitigación de riesgos, alianzas para la provisión de infraestructura y herramientas tecnológicas, participación en los programas de formación o de investigación; (ii) la demanda de mercado futura por los servicios a ser desarrollados por el proyecto, en un contexto en que la IA se adopta de manera creciente y que debe ir acompañada, ya sea por tendencias regulatorias o voluntarias, de la adopción de herramientas de minimización de riesgos; (iii) la tendencia de desarrollo de marcos jurídicos y regulatorios que acelerarán la adopción responsable de la IA, motivados por ejemplo, por los acuerdos de convergencia regulatoria ya existentes entre la región ALC y otras regiones como la Unión Europea.
- 3.9 La consolidación de las herramientas, generación de capacidades y concientización de RAI 4 Growth será pieza clave para seguir posicionando la iniciativa fAIr LAC, y los proyectos piloto impulsados, como instancia de articulación y habilitación para que las tecnologías de IA responsable actúen como motor en favor de la disminución de problemáticas sociales y ambientales. La adicionalidad de BID Lab está dada por la experiencia ya desarrollada en el marco de la iniciativa fAIr LAC, las conexiones con actores clave del ecosistema regional y el énfasis en el pilotaje de modelos con impacto socioambiental, en los cuales el desarrollo responsable de tecnologías como la IA es un pilar básico para el avance en temas de inclusión y la equidad.
- 3.10 La escalabilidad del proyecto se puede medir también a partir del crecimiento potencial de la IA (en uso y desarrollo en la región y en los países donde operará RAI 4 Growth en los próximos tres años). El mercado potencial de beneficiarios en la región ALC más allá del proyecto RAI 4 Growth se estima en 1.200.000 de MiPyMEs⁴⁴. Si el foco es en los países donde opera el proyecto RAI 4 Growth, el número estimado es de 500.000. Desde el punto de vista de empresas que desarrollan IA, se espera un potencial de 2.000 a 5.000 empresas

⁴⁴ Estimación conservadora que corresponde al 10% del número estimado de MiPyMEs en la región LAC (Panorama Digital de las Micro, Pequeñas y Medianas Empresas (MiPyMEs) de América Latina 2021. Universidad de Maastricht, Sela).

de desarrollo de IA en los próximos 10 años y en los países del proyecto⁴⁵. RAI 4 Growth, tal y como se indica anteriormente, tiene un objetivo de conseguir que 1.000 empresas usuarias de IA, 250 empresas desarrolladoras de IA y 100 inversores participen en el proyecto e implementen IA responsable. Además se sensibilizarán y generarán capacidades (formación) en más de 5.000 actores del sector público, privado y sociedad civil, sentando las bases para la expansión al resto del ecosistema. De esta manera, durante la fase inicial de pilotaje se estima un costo medio por usuario en torno a los USD 500, monto que será decreciente una vez que la infraestructura tecnológica y de gestión se encuentre operando en régimen.

C. Riesgos del proyecto e institucionales

- 3.11 **Riesgo de un ecosistema poco favorable para el proyecto:** Si los ecosistemas objetivo carecen de madurez y conciencia sobre la importancia de la IA responsable, estos podrían limitar el número de startups y MIPYMES interesadas lo que impactaría los objetivos del proyecto. Este riesgo se **mitiga** a través de actividades de concientización, incluidas en los componentes del proyecto, al distribuir conocimiento específico de cómo el enfoque de IA responsable contribuye al éxito de los negocios y la consecución de soluciones socialmente beneficiosas. Es por ello por lo que el proyecto incidirá en la competitividad empresarial como factor de concientización.
- 3.12 **Riesgos reputacionales y “ethics washing”:** Algunas empresas podrían motivar su participación basándose en intereses comerciales o reputacionales en lugar de compromiso real con la IA responsable. Como **mitigante**, se establecerán medidas rigurosas de evaluación y verificación para la obtención de certificaciones, así como el involucramiento de terceras partes para la validación externa de las acciones de las empresas.
- 3.13 **Riesgo de falta de corrección y mitigación** de riesgos de la IA por parte de las entidades participantes: Algunas empresas podrían optar por no corregir sus modelos de mitigación de riesgos o incorporar estándares de transparencia. Como **mitigante**, se generarán incentivos que motiven a las empresas a hacer una corrección efectiva de sus modelos, tales como el condicionamiento de certificación, apoyo, financiamiento, asesoría o acceso a servicios, y medir el impacto de dichas correcciones.
- 3.14 **Riesgo de impacto algorítmico negativo:** Derivado de lo anterior, algunas empresas participantes podrían optar por no corregir riesgos ambientales y sociales asociados con sus actividades relacionadas con la IA. Para **mitigarlo**, se buscará incorporar enfoques de diseño participativo y multiactor en sectores de alto riesgo, que permitan el establecimiento de medidas de monitoreo y auditoría de forma preventiva.
- 3.15 **Riesgo en el despliegue en los países de ejecución:** La entidad ejecutora con sede en España podría carecer de las redes locales necesarias en los países donde se llevará a cabo la ejecución del proyecto debido a la distancia geográfica y falta de experiencia previa en la región. Como **mitigante**, el proyecto apalancará el despliegue de actividades mediante la colaboración con las redes de actores construidas por BID Lab y fAIr LAC. Además se organizarán convocatorias abiertas para sumar nuevos actores.
- 3.16 **Riesgo político y regulatorio:** La región de ALC puede experimentar diferencias en su marco legal y regulatorio de IA, así como en sus políticas nacionales de IA, dificultando la estandarización de los servicios ofrecidos por el proyecto. Este riesgo se **mitiga** mediante el monitoreo de las iniciativas regulatorias y la consulta con expertos locales que permitan entender las diferencias regionales significativas a la vez que se promueve el cumplimiento

⁴⁵ Estimaciones de Adigital en base a información secundaria y estudio de mercado realizado el primer semestre de 2024.

de estándares internacionales aceptados en ocasiones más restrictivos que las regulaciones nacionales.

IV. INSTRUMENTO Y PROPUESTA DE PRESUPUESTO

El proyecto tiene un costo total de US\$3.200.000, de los cuales US\$1.700.000 (53%) serán aportados por BID Lab bajo la forma de Cooperación Técnica no Reembolsable y US\$1.500.000 (47%) por la contraparte. [Adigital, Asociación Española de la Economía Digital, será Organismo Ejecutor de financiamiento BID Lab por hasta US\\$ 1.400.000 y aportará recursos de contrapartida. BID Lab \(Fondo Multilateral de Inversiones\) será el ejecutor de sus propios recursos por hasta US\\$300.000.](#)

Presupuesto- **Indicativo-Resumido** (US\$)

Componentes del proyecto	BID Lab (CTNR)		Contrapartida		Total
	Ejecutado por el BID	Ejecutado por Adigital	En Especie	En Efectivo	
Componente 1: Generación de confianza	50.000	332 282.920	19.410	158.680	511.010
Componente 2: Habilidades y talentos del ecosistema	0	374.210	299.860	204.620	878.690
Componente 3: Herramientas y certificaciones	250.000	853 603.010	447.545	369.885	1.670.440
Traducciones	0	75.000	0	0	75.000
Imprevistos*	0	64.860	0	0	64.860
Total	300.000	1.709 400.000	766.815	733.185	3.200.000
% de financiamiento	53 9%	44%	24%	23%	100%

(*) La utilización de recursos Imprevistos deberá contar con la no objeción ex ante del Banco.

V. ORGANISMO EJECUTOR Y ESTRUCTURA DE IMPLEMENTACIÓN

A. Descripción del organismo ejecutor

- 5.1 Tal como se describe en la sección B del presente apartado, el proyecto será ejecutado de manera conjunta por BID Lab y Adigital.
- 5.2 Adigital, asociación española de la economía digital, será el Organismo Ejecutor (OE) principal de este proyecto y firmará el convenio con el Banco. Adigital es una organización sin ánimo de lucro formada por más de 500 empresas digitales en España de sectores clave de la economía digital unidas por un objetivo común: crear un entorno óptimo para el desarrollo y crecimiento de la economía digital, que permita alcanzar una sociedad más abierta, competitiva y sostenible, con especial atención a la ética digital. Lo anterior, posiciona a Adigital como uno de los referentes del desarrollo digital responsable en España y la Unión Europea y evidencia su capacidad de gestión de alianzas tanto en el ámbito público como privado.
- 5.3 Adigital tiene experiencia en idear, construir y lanzar con éxito certificaciones y sellos digitales para la economía digital, adaptándose y reaccionando a la evolución de la economía digital durante los últimos veinte años, siendo los más destacados: (i) Confianza Online, que busca fomentar la confianza en las transacciones en línea y la navegación en

sitios web, otorgando un sello a los sitios web que se han sometido a una evaluación basada en más de 30 criterios que incluyen, entre otras materias, cumplimiento legal y estándares éticos. Este sello ha sido reconocido por aumentar la confianza de los usuarios y mejorar el posicionamiento en los motores de búsqueda, incluye además la adhesión a un código ético que abarca áreas como la protección de datos personales y el comercio electrónico; (ii) Lista Robinson, permite a los particulares reducir la publicidad no deseada mediante el bloqueo de las empresas que realizan estas campañas, las que están obligadas a consultar la Lista Robinson, con posibles sanciones en caso de incumplimiento. Este servicio ha sido validado y reconocido por la Agencia Española de Protección de Datos.

- 5.4 Respecto a transparencia y explicabilidad del uso y desarrollo de IA, Adigital ha puesto en marcha algunas iniciativas clave para incrementar tanto el conocimiento como la implementación en las empresas españolas: (i) Adigital Academy, recientemente ganadora del premio ExpoE Learning 2024⁴⁶, que brinda a sus miembros formación en temas de IA Responsable tales como evaluación de riesgos, transparencia, explicabilidad, tarjetas modelo, RLHF (aprendizaje por refuerzo con retroalimentación humana), etc.; (ii) repositorio de “Tecnología Responsable” que incluye cursos, artículos, videos cortos, y panoramas de productos de IA responsable y otras tendencias emergentes; (iii) certificación sobre transparencia algorítmica para sistemas de IA, en línea con la política y los desarrollos regulatorios actuales de la UE (Unión Europea) e internacionales, que evalúa cómo las entidades manejan los sistemas de IA y sus modelos subyacentes. Esta certificación ha sido construida junto con expertos en tecnología, regulación y políticas públicas, se ha presentado en el Foro Global de Ética e IA de UNESCO⁴⁷ y es parte del Catálogo de herramientas confiables de IA de la OCDE⁴⁸.
- 5.5 Las acciones de proyecto en los países de intervención se realizará identificando entidades que tengan como parte de su misión los temas de la RAI y de apoyo al ecosistema de startups (inversión, aceleración, incubación, entre otros) o a las empresas de menor tamaño. Inicialmente se establecerán convenios de colaboración con entidades que ya tienen relación con fAlr LAC en estos países para luego escalar el alcance mediante convocatorias abiertas, cuyos términos serán co-construidos con el Banco e incluirán aspectos para el análisis de integridad de las entidades.
- 5.6 **Revisión de Integridad:** El Equipo del Proyecto, con la asistencia de OII, llevó a cabo una diligencia debida de integridad sobre el Proyecto e identificó indicadores de riesgo de integridad e impacto reputacional, así como factores mitigantes. Los riesgos de integridad e impacto reputacional presentados por este proyecto se consideran dentro de la tolerancia de BID Lab para tales riesgos. Para obtener más detalles, consulte el Anexo de Integridad.

B. Estructura y mecanismo de implementación

- 5.7 Con el fin de asegurar la adecuada articulación del proyecto con la iniciativa fAlr LAC, la generación de bienes públicos de conocimiento y coordinación con el resto del Grupo BID, BID Lab ejecutará sus propios recursos por hasta \$ 300.000 para la implementación de actividades del proyecto, velando por el seguimiento de la operación en su conjunto, la calidad de los reportes de avance preparados por Adigital y la adecuada gobernanza del proyecto.

⁴⁶ <https://expoelarning.com/>

⁴⁷ <https://www.youtube.com/watch?v=zFdVid2gev0&list=PLi0Q8J3Ksqb7BI-UN2L4Sq3YKWpwyR5N&index=31>

⁴⁸ <https://www.adigital.org/actualidad/la-ocde-incluye-el-certificado-de-transparencia-algoritmica-de-adigital-en-su-catalogo-de-herramientas-y-metricas-para-la-ia-confiable/>

- 5.8 Las actividades ejecutadas por el Banco, serán implementadas por BID lab en la Sede bajo el liderazgo del Jefe de la Unidad Discovery de BID Lab, quien podrá delegar su coordinación en el Jefe de Equipo u otros miembros del Staff.
- 5.9 Las actividades y recursos financiados por BID lab a ser ejecutados por el OE, contemplarán el establecimiento de una unidad de ejecución que considere la estructura necesaria para ejecutar las actividades del proyecto y gestionar los recursos con eficacia y eficiencia. El OE también será responsable de presentar informes de avance acerca de la implementación del proyecto.
- 5.10 Para su gobernanza efectiva, el proyecto contará con un diferentes comités para el acompañamiento técnico, operativo y estratégico como se describe a continuación:
- 5.11 **Consejo Directivo:** Este consejo es el órgano principal de toma de decisiones, compuesto por representantes de alto nivel de todas las entidades asociadas y afiliadas clave (inicialmente Adigital y BID Lab, potencialmente añadiendo empresas o instituciones que conformen alianzas estratégicas). Su función es la de establecer/actualizar la visión y dirección estratégica del proyecto, velar por el logro de resultados, aprobar los grandes planes y presupuestos, y supervisar la gestión del proyecto a nivel macro.
- 5.12 **Equipo Ejecutivo:** Este equipo actúa bajo la dirección del Consejo Directivo y está compuesto por líderes operativos del proyecto (Adigital como agente ejecutor, BID Lab como coejecutor e impulsor principal, alianzas estratégicas y, potencialmente y de manera temporal, partners no estratégicos pero que tácticamente deban ser incluidos en la toma de decisiones de componentes concretos). Su rol principal es el de implementar las estrategias y políticas establecidas por el Consejo Directivo, gestionar las operaciones diarias y asegurarse de que el proyecto avance según lo planificado en la matriz de resultados.
- 5.13 **Comité Asesor:** Compuesto por expertos en IA, ética, regulación y otros representantes, su función será asesorar al Comité Ejecutivo y al Consejo Directivo en cuestiones específicas relacionadas con la implementación responsable de la tecnología, asegurando que el proyecto cumpla con los estándares internacionales y las mejores prácticas en IA responsable.

VI. CUMPLIMIENTO DE HITOS Y ACUERDOS FIDUCIARIOS ESPECIALES

- 6.1 **Desembolsos por Resultados, Acuerdos Fiduciarios.**
- 6.2 Adigital se compromete con los arreglos estándar de desembolsos por resultados de BID Lab, las políticas de contrataciones y gestión financiera aplicables al sector privado, en consonancia con la Guía de Gestión Financiera para Proyectos Financiados por el BID (OP-273-12) versión del 12 de junio de 2019 y según se especifica en la "Guía de Gestión por Hitos y Supervisión Financiera para proyectos de Cooperación Técnicas de BID Lab y Programa de Empresariado Social"⁴⁹.
- 6.3 BID Lab se atendrá a la misma modalidad de desembolso por resultados, las políticas de adquisiciones y mecanismos de gestión financiera establecidos por el Banco.
- 6.4 **Desembolsos basados en resultados.** La Oficina de País de Chile supervisará el Proyecto. El seguimiento se llevará a cabo de acuerdo con las políticas de gestión de resultados y riesgos (cumplimiento de hitos) establecidas por el BID Lab en abril de 2008.

⁴⁹ [Guía de Gestión por Hitos y Supervisión Financiera para proyectos de Cooperación Técnicas de BID Lab y Programa de Empresariado Social](#)

VII. ACCESO A INFORMACIÓN Y PROPIEDAD INTELECTUAL

- 7.1 **Acceso a información.** De acuerdo con la Política de Acceso a Información del Banco, este documento es de acceso público una vez aprobado.
- 7.2 **Propiedad intelectual:** La propiedad intelectual de todos los trabajos y resultados del proyecto que no se basen en obras preexistentes de cada una de las partes, pertenecerá al Banco y a Adigital de manera conjunta (los “Entregables del Proyecto”).
- 7.3 Cada parte podrá otorgar a terceros una licencia sobre los Entregables del Proyecto licencia que será no exclusiva, gratuita y con fines no comerciales para usar, copiar, distribuir, reproducir, exhibir y ejecutar públicamente los Entregables del Proyecto.
- 7.4 En caso de que los Entregables del Proyecto sean usados por terceras partes bajo una licencia otorgada por Adigital, el uso y distribución respeterá la licencia Commons IGO 3.0 BY-NC-ND, exhibiendo la siguiente leyenda:
- 7.5 “Copyright © [año de la primera publicación] Asociación Española de la Economía Digital (Adigital) y Banco Interamericano de Desarrollo (BID). Esta obra está bajo una licencia Creative Commons IGO 3.0 Reconocimiento-NoComercial-SinObraDerivada (CC-IGO BY-NC-ND 3.0 IGO) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando crédito a Adigital y BID. No se permiten obras derivadas. Cualquier disputa relacionada con el uso de las obras del proyecto RAI 4 Growth que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI. El uso del nombre del proyecto RAI 4 Growth para cualquier fin que no sea para la atribución y el uso del logotipo, estará sujeta a un acuerdo de licencia por separado y no está autorizado como parte de esta licencia CC-IGO. Notar que el enlace URL incluye términos y condicionales adicionales de esta licencia”.
- 7.6 Las obras preexistentes de cada una de las partes, que sean puestas a disposición para su uso en el ámbito de RAI 4 Growth, y sus mejoras incorporadas en el marco del proyecto, permanecerán bajo titularidad de cada entidad, como se detalla a continuación:

BID	Adigital
Herramienta de autoevaluación para emprendedores, 3S	Reglamento y certificación de transparencia algorítmica de Adigital
Herramienta de valoración de riesgos para fondos de inversión, fAIr Venture	Cursos preexistentes de Adigital Academy ⁵⁰

- 7.7 Ambas partes garantizarán que la ejecución del Proyecto no infringe ni infringirá derechos de terceros y se obligan a llevar a cabo todas las actividades necesarias para que ambas partes puedan ejercer los derechos aquí previstos sin limitaciones.
- 7.8 Ambas partes se obligarán a incluir en todos los contratos que celebren con consultores bajo el Proyecto involucrados en el desarrollo de los Entregables del Proyecto, la cesión a favor de ambas partes de los respectivos derechos de propiedad intelectual, incluyendo los derechos de autor.
- 7.9 Ambas partes podrán divulgar, reproducir y publicar cualquier información vinculada al Proyecto e incluir en dicha información el nombre y logotipo de ambas partes.

⁵⁰ La lista de cursos preexistentes está disponible en los archivos técnicos del proyecto