CONFIDENTIAL
INTERNAL USE
PUBLIC UPON APPROVAL

DOCUMENT OF THE INTER-AMERICAN DEVELOPMENT BANK
MULTILATERAL INVESTMENT FUND

REGIONAL

RESPONSIBLE ARTIFICIAL INTELLIGENCE FOR GROWTH IN LATIN AMERICA AND THE CARIBBEAN
RAI 4 GROWTH

(RG-T4463)

DONORS MEMORANDUM

This document was prepared by the project team consisting of: 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 Muente (IFD/ICS), Juan Pedeflous (GCL/FML), and Delfina Müller (DSP/DVF).

This document contains confidential information relating to one or more of the ten exceptions of the Access to Information Policy and will be initially treated as confidential and made available only to Bank employees. The document will be disclosed and made available to the public upon approval.

CONTENTS

PROJECT SUMMARY

REGIONAL
RESPONSIBLE ARTIFICIAL INTELLIGENCE FOR GROWTH IN LATIN AMERICA AND THE CARIBBEAN
RAI 4 GROWTH
(RG-T4463)

In the Latin American and Caribbean innovation ecosystem, adoption of artificial intelligence (AI), the technology that is becoming a fundamental element for economic growth, lags behind the global average. Although the region's private sector is starting to integrate AI into its business models, products, and services, the lack of adequate support may undermine society's trust in these systems. This represents a barrier to adoption, and consequently, to making the companies that develop and use AI more competitive. Likewise, if AI is not developed responsibly, it risks reproducing and amplifying inequalities on a large scale, as has already occurred in other more advanced regions of the world. A lack of equity, transparency, and accountability can perpetuate human prejudices and errors.

Ensuring, from the start if possible, that the development and use of AI align with international standards in sectors like financial inclusion, employment, education, health, and justice is fundamental for its adoption and long-term financial sustainability, as well as to guaranteeing that its benefits will reach all groups in society, especially poor and vulnerable ones. Along these lines, incorporating a "responsibility by design" approach in the creation and use of AI systems will enable companies to implement this technology's value proposition, which will benefit all strata of society, and in turn improve companies' market competitiveness through the development of sound, trustworthy products and services. In addition, the use of responsible AI by the most innovative businesses supports the positioning of companies that are key to the region's economic development, by facilitating the internationalization of startups and growing companies, thanks to their compliance (early in their countries) with AI risk mitigation regulations, which makes them into more competitive companies that will create better quality jobs. Furthermore, in these stages, the companies foster a culture that values diversity as an engine of creativity and innovation. Therefore, they are key actors for collaborating on risk mitigation in areas like employability and training.

RAI 4 Growth (Responsible AI for Growth), the proposed solution, is an integrated program that fosters the widespread application of responsible AI standards and tools among AI companies in Latin America and the Caribbean, as well as the adoption of AI tools by startups, micro-, small, and medium-sized enterprises (MSMEs), and investors in the region. This program also develops a set of measurement plans[1] so that AI developer and user companies can understand how the systems respond to the mitigation of risks associated with economic vulnerability and exclusion, the gender perspective, and sustainability. In this way, AI developers and users will both be able to benefit from this

[1] A software system measurement plan is a structured set of metrics and criteria designed to evaluate the system's performance, quality, and impact over its life cycle. For responsible AI systems, measurement plans are used to ensure adherence to ethical principles and responsibility standards like equity, transparency, and nondiscrimination.

powerful technology for greater efficiency and trust, while reducing their risks and becoming more competitive. The program is structured around three pillars:

- **Tools, certifications, and measurement** for proper implementation of responsible AI (RAI);
- **Skills and educational resources** for the responsible development of AI in the innovation ecosystem; and
- **Trust-building and outreach,** focused on informing the population and raising awareness of the critical use of AI products and services that positively impact daily life.

The project benefits from the expertise of the Spanish Association for the Digital Economy (Adigital), to be joined by collaborating experts in technology, regulation, ethics, and public outreach (individuals and organizations), thereby promoting the IDB Group fAIr LAC initiative, which fosters the ethical, responsible use of artificial intelligence to contribute to the region's inclusive socioeconomic development.

ABBREVIATIONS

Adigital	Asociación Española de la Economía Digital (Spanish Association for the Digital Economy)
AI	Artificial intelligence
LAC	Latin America and the Caribbean
MSMEs	Micro-, small, and medium-sized enterprises
OECD	Organisation for Economic Co-operation and Development
RAI	Responsible artificial intelligence
SDGs	Sustainable Development Goals
SMEs	Small and medium-sized enterprises
UNESCO	United Nations Educational, Scientific, and Cultural Organization

EXECUTIVE SUMMARY

REGIONAL
RESPONSIBLE ARTIFICIAL INTELLIGENCE FOR GROWTH
IN LATIN AMERICA AND THE CARIBBEAN
RAI 4 GROWTH
(RG-T4463)

Country and geographic location:	Regional, with initial actions in Chile, Mexico, Colombia, Uruguay, and Ecuador		
Executing agency:	Adigital, the Spanish Association for the Digital Economy, will be the executing agency for the IDB Lab financing of up to US$1.4 million and will provide counterpart funds. IDB Lab (Multilateral Investment Fund) will execute its own resources of up to US$300,000.		
IDB Lab vertical areas and crosscutting pillars:	Education, talent, and employment; Gender and diversity		
Coordination with other donors/Bank operations:	Coordination with the fAIr LAC initiative and operations involving the development of knowledge, tools, analytical frameworks, training, and analysis of responsible AI for the innovation ecosystem in the region.		
Project beneficiaries:	Investors, MSMEs, and startups, with an emphasis on ones whose products or services target vulnerable populations, women, or environmental benefits.		
Financing:	Technical cooperation:	US$1.7 million	53%
	Total financing from IDB Lab:	US$1.7 million	
	Counterpart:	US$1.5 million	47%
	Total project budget:	US$3.2 million	100%
Execution and disbursement periods:	36 months for execution and 42 months for disbursements.		
Conditions for the first disbursement:	For the first disbursement of the technical cooperation funds, the executing agency will submit the following to the Bank's satisfaction: (i) project procurement plan; (ii) appointment of the program manager; and (iii) appointment of the project coordinator.		
Environmental and social impact review:	This operation was analyzed and classified on 20 May 2024 in accordance with the IDB Environmental and Social Policy Framework (document GN-2965-21). Since the project's impacts and risks are limited, the proposed category is "C."		
Unit responsible for disbursements:	CCH – For funds to be executed by Adigital; LAB/DIS – For funds to be executed by IDB Lab		

I. THE PROBLEM

A. Description of the problem

1.1 The integration of artificial intelligence (AI) into the Latin America and Caribbean business sector and innovation ecosystem offers great potential for addressing the challenges posed by the 2030 Agenda for Sustainable Development.[2] This set of technologies can help improve the impact of economic development in critical sectors like health, agriculture, industry (due to its extensive value chain), education, and the provision of public services. Projections indicate that by 2030, AI may account for over 5% of regional GDP.[3] It may also help the regional market (for the supply and demand of AI-based solutions) grow 18.3% between 2023 and 2028.[4]

1.2 A recent survey conducted by NTT Data and MIT Technology Review showed that the percentage of Latin American companies interested in using AI increased from 58% in 2020 to 71% in 2023, and that 79% have already started using AI tools to some extent.[5] However, the same report also indicates that Latin America and the Caribbean still lag behind the global average: the region's overall effective AI implementation rate is 37% of companies, compared to a worldwide average of 42%. The business sector's primary motivations for adopting these tools have to do with the quest to improve competitiveness, operational efficiency, and quality of operations, and to reduce errors.

1.3 Although the region's private sector clearly intends to adopt AI, few companies are actually prepared to take advantage of its maximum potential or to properly manage the risks involved. For example, McKinsey's 2023 report on the state of AI indicates that just 21% of companies that report having adopted AI tools have established explicit policies governing their use or measures to mitigate risks such as a lack of transparency, inaccuracy, cybersecurity issues,[6] or other more serious risks like the possibility of generating harmful biases or algorithmic discrimination.[7] This is especially significant because the use of algorithms is increasingly affecting and transforming lives around the world—how populations work, play, and access essential services like healthcare, financial services, education, employment, social assistance, critical infrastructure, agriculture, food security, and even telecommunications.[8] Despite the scarcity of safeguards established to reduce the risks of using AI, a July 2023 Statista report indicates that 70% of citizens of the region feel that the technology's potential outweighs any of its drawbacks.[9]

---

2 Naciones Unidas. *Objetivos de Desarrollo Sostenible. La Asamblea General adopta la Agenda 2030 para el Desarrollo Sostenible*.

3 Forbes Centroamérica. *Cinco de las principales startups de Inteligencia Artificial de América Latina y el Caribe*.

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

5 Artificial intelligence in Latin America, 2023, NTT Data – MIT Technology Review.

6 McKinsey Global Survey. The State of AI in 2023: Generative AI's breakout year.

7 Gobierno de España. Ministerio para la Transformación Digital y de la Función Pública. *Invisibilización y discriminación algorítmica*.

8 GatesNotes. The Age of AI Has Begun.

9 Statista. *¿Apoyan los latinoamericanos el uso de la inteligencia artificial?*.

The combination of businesses' accelerated adoption of AI and citizens' positive views of it confirms a trend in which its rapid introduction will touch almost all areas of daily life.[10] Notably, this also significantly increases the degree of risk and impact of its potential negative effects.

1.4 Furthermore, while the productivity gap between MSMEs and large companies exists around the world, it is especially marked in Latin America and the Caribbean. In the region, the production value of MSMEs only accounts for 25% of total production value; MSMEs in Latin America and the Caribbean generate just half the wealth of their European counterparts.[11] It is in this context that automation and AI represent major opportunities, not just to boost the competitiveness of the region's MSMEs, but also for the ecosystem of startups that has evolved considerably over the past few years.[12]

1.5 The acceleration of and potential for autonomous decision-making by AI systems present not only technical risks, but also social ones. These risks could exacerbate inequality and perpetuate dynamics of discrimination and socioeconomic exclusion, due to the rapid adoption of AI systems, the lack of tools and education in the private sector for managing these risks, and the technology's large-scale impact. Some of the most significant risks are described below:

- The use of AI will affect the poor and vulnerable population. In Latin America and the Caribbean alone (the region with the world's highest level of socioeconomic inequality),[13] it could affect over 180 million people who do not earn enough to cover their basic needs, and 70 million people who do not have sufficient income to purchase a basic food basket.[14] Socioeconomic inequality and poverty are two aspects that may be severely impacted by the rapid development and deployment of AI if safeguards and strategies are not put in place to properly mitigate the risks of discrimination, income concentration, and the asymmetrical, unilateral capture of economic value.[15] AI will also impact labor markets due to automation and adversely affect the earnings of the lowest-income individuals.[16] Notably, these are not long-term problems, but rather challenges that must be addressed right away.[17]
- Another important factor to consider in the development of AI is the technology's impact on environmental sustainability: training one single advanced AI model can produce up to the same amount of carbon dioxide as

---

10 El País. *¿Qué queda al margen de la inteligencia artificial?*

11 Small businesses, big impacts: Supporting productive SMEs as an engine of recovery in LAC, United Nations Development Programme, 2021.

12 Economic Commission for Latin America and the Caribbean (ECLAC), 2003. *Empresas emergentes (startups) en América Latina y el Caribe*, 2023.

13 IDB. The Complexities of Inequality in Latin America and the Caribbean.

14 ECLAC. Social Panorama of Latin America and the Caribbean 2023: Labour Inclusion as a Key Axis of Inclusive Social Development.

15 The Economist. Just how rich are businesses getting in the AI gold rush?

16 IMF 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 Committee on Digital Economy Policy Ministerial meeting, 15 November 2022.

300 round-trip flights between New York and San Francisco,[18] and training foundational models will require around a gigawatt of energy—practically the same amount generated by a nuclear reactor.[19] Although few companies have the data and supercomputing capacity needed to train models, and most use pretrained models or train their models using cloud hyperscalers, it is important for companies that use AI to incorporate approaches that take into account the environmental footprint of doing so and that can measure the environmental impact of these systems. Companies that do train their own models have made appreciable progress on establishing offsetting policies (whether through compensation in cash, in kind, or in terms of the positive contribution that some of the models make to environmental sustainability).[20] Although there are currently no standards in place on measuring and mitigating the environmental effects of using AI, companies can become more transparent and start to take the first steps.[21]

- One of the main risks of using and developing AI algorithms that could affect the vulnerable population is algorithmic discrimination, meaning that the data used to train automized decision-making systems or AI models replicate, perpetuate, and/or amplify biases and inequalities due to factors such as gender, race, age, and their intersectionalities. For women in general, AI can pose grave threats to economic and personal security, as well as serious risks of discrimination and exclusion if the technology is not appropriately developed and used.[22] This makes it necessary for companies—regardless of their size—and investors to take mitigation measures,[23] and to explain and be accountable for how they are doing this.

1.6 These risks, though generalizable to any productive sector, are escalating in sectors where there is evidence of the prevalence of biases and vulnerabilities. The specific challenges in certain entrepreneurial and innovation sectors are discussed below.

- In the **fintech** sphere, the challenges of AI are closely associated with the need to evaluate and improve AI systems so to avoid major problems associated with equity and transparency. The mitigation/elimination of gender and racial biases in credit scoring tools and algorithms can help resolve the structural problem of women and ethnic minorities with little to no credit history. In all these cases, algorithmic transparency is crucial, as a lack of clarity on how credit scores are calculated can erode trust in AI systems. Lastly, privacy is a

---

18 Earth.org The Green Dilemma: Can AI Fulfil Its Potential Without Harming the Environment?

19 Energy, not compute, will be the #1 bottleneck to AI progress.

20 Fang, M., R. Li, and X. Zhao. 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.

21 OECD. AI Policy Observatory. Will business or laws and regulations ever prioritise environmental sustainability for AI systems?

22 The Brookings Institution. AI poses disproportionate risks to women.

23 El País. *AI regular la inteligencia artificial, debemos colocar raza y género en el centro del debate*.

major concern, due to the risk of personal financial data being exposed or misused, leading to breaches of trust.

- In the **health** sphere, transparency in diagnostic tools can help find specific errors in the diagnosis of conditions in women and in diagnostic precision for certain racial groups. In all these cases, algorithmic transparency is crucial; a lack of understanding of AI-driven diagnostic processes can erode patient trust in these systems. Furthermore, sensitive health data could be filtered or misused, endangering patient confidentiality.
- With regard to **employability**, the mitigation/elimination of gender and racial biases in recruitment and job-matching algorithms can help resolve the structural problem whereby men and certain racial groups are inadvertently favored due to biased training data. As for privacy, background verifications may unnecessarily expose personal information.
- In the **education** sector, biases in educational recommendation algorithms and tools can reinforce stereotypes and limit students' exposure to diverse fields, and must be addressed. The limitations of AI systems can also impact personalized learning algorithms, preventing students from accessing the information that is most relevant to their interests.

1.7 Although growth in the market for responsible artificial intelligence (RAI) services has been observed worldwide,[24] the significant gap in specialized capacity for startups, growing companies, and MSMEs in Latin America and the Caribbean prevents them from appreciating the benefits of AI. The limited access to AI resources, training, and education[25] widens these gaps, meaning the region needs to identify cost-effective formulas that will make it possible to develop tools, ensure their mass adoption, and provide training on RAI,[26] which should be especially accessible to MSMEs that typically find it difficult to adopt new technologies, whether due to a lack of knowledge and skills or unawareness of the types of support available.[27] Notably, although a wide variety and large amount of information on next-generation RAI is available, it is scattered among multiple repositories and there is no guidance for companies that want to implement it to indicate which lines of work are most technically, regulatorily, and ethically suitable for their specific use cases.[28]

1.8 The IDB Group fAIr LAC initiative, based on the OECD AI Principles for Trustworthy AI,[29] was designed specifically to disrupt this trend and support the

---

[24] Grand View Research. AI Governance Market Size and Trends.

[25] https://www.oecd-ilibrary.org/sites/e7a00fd6-en/index.html?itemId=/content/component/e7a00fd6-en.

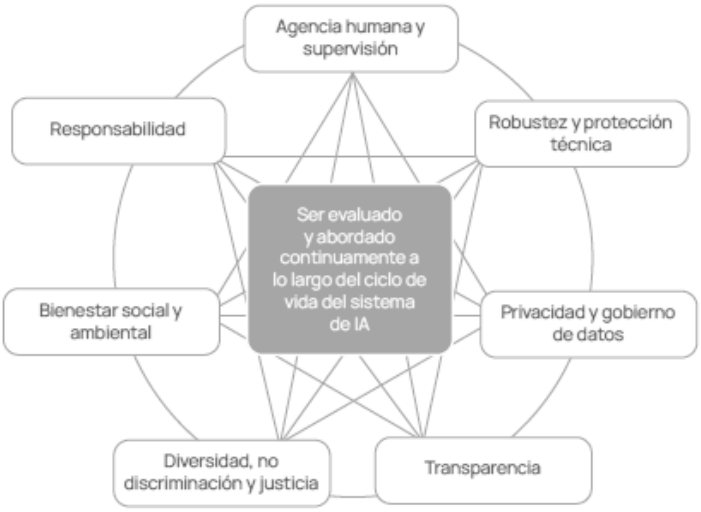
[26] Economist Impact. Investment in AI is booming in Latin America, but what will it mean for the region's economy?

[27] European Union. Boosting the Use of Artificial Intelligence in Europe's micro, small, and medium-sized Enterprises.

[28] For example, the Holistic.ai library is a good example of technical support, but it might be difficult for AI-user companies or new developers to understand which of the cases described are most applicable, or what the potential impact will be on the affected populations. Holistic AI. Measuring and Mitigating Bias: Introducing Holistic AI's Open-Source Library.

[29] OECD. AI Policy Observatory. OECD AI Principles Overview.

region in adopting a responsible approach to AI development and adoption that will be beneficial for all. With the momentum of pilot projects that have already been executed in Mexico, Chile, Colombia, and Costa Rica, which since 2019 have strengthened public and private sector stakeholders, entrepreneurs, and startups in the region, and the development of self-assessment tools for startups and investors, fAIr LAC has sought to build capacity among actors in the ecosystem to enable them to develop, adopt, or finance AI-based solutions and processes that incorporate ethics- and responsibility-based criteria. The main lessons learned from these experiences are that: (i) since AI is a groundbreaking topic for which there are no regulations or rules in place, or previous experience to refer to, its approach must be flexible and gradual, so that stakeholders can implement the principles within a framework of continuous improvement and early regulation; (ii) along these same lines, it is essential to continue building relevant skills among developers, users, and regulators as well as ecosystem stakeholders in general. This is because the idea of RAI can be interpreted in various ways, and it is important to incorporate not only the technological elements but also the humanistic view that considers the effects these systems have on organizations and on people's lives, thereby making it possible to appreciate AI's true potential for improving quality of life; (iii) it is important to establish diverse, multi-actor, participative processes for reflection and dialogue, in simple language, on the reasons for analyzing and improving AI systems based on ethics and responsibility; (iv) in working on innovation programs, it is very important to obtain the commitment of the entities' leaders and technical staff, to increase the probability that the pilot experiences will be executed and monitored; (v) for appropriate development and adoption of ethical AI, it is essential to raise awareness of the importance of including the monitoring of ethical considerations in project management tools; and (vi) in developing tools and products to help implement RAI, it is a good idea to consider participatory processes that will facilitate adoption.



OECD principles for responsible AI management

1.9 These challenges must be approached as opportunities. Humanity has already faced many other new technologies that entail both benefits and risks, and has addressed those risks in favor of societal progress. Accordingly, it is important to understand the opportunity to be found in fostering the responsible development and adoption of AI in MSMEs and startups. This is a nascent issue that the countries of the region have not yet incorporated into their current regulations.[30] Therefore, innovative companies can blaze the trail and lay the foundation for a technological market that allows for greater competitiveness and internationalization.[31] They can also take advantage of the creation of new market categories (such as the AI management and monitoring tools themselves), incentivizing inclusive and safe solutions, and leveraging the opportunities afforded by the technology while effectively managing its risks. It is essential to work with investors (e.g., multilateral banks, venture capital funds, and angel investors) in order to encourage companies that develop or use AI to adopt RAI approaches when they seek financing. This is particularly significant given that much of the venture capital in the region focuses on financing technological solutions in high social impact sectors such as financial services,[32] through modalities like alternative credit scoring.[33] These efforts must be leveraged with robust tools and action frameworks, together with greater capacity for measuring the desired results, which will make it possible to reach a larger audience of startups, growing companies, MSMEs, investors, and other ecosystem entities, thereby paving the way for the responsible, equitable adoption of AI to drive the regional technological AI innovation ecosystem, and position it on the international stage, while mitigating the risks and protecting the most vulnerable. Ultimately, the promotion of RAI by pioneering companies serves as a guiding light and mirror for the rest of the social and business community to implement AI systems with appropriate levels of control over privacy, equity, and transparency.

II. THE INNOVATION PROPOSAL

A. Project description

2.1 The use of AI must be responsible by design[34] to ensure that the competitiveness of the corporate sector is maintained.[35,36] While ultimately the entire ecosystem must adopt AI, in the first stage pioneering companies must act as models for other

---

30 Infobae. *¿Qué se está haciendo en Latinoamérica para regular la inteligencia artificial?*

31 The fourth edition of Adigital's Digital Economy in Spain (p. 34) shows how companies with greater compliance of sustainability and environmental, social, and governance policies have lower capital costs, which, together with their increased compliance, can positively affect their competitiveness.

32 Statista. *¿Qué industrias reciben más inversión de capital riesgo en IA en América Latina?*

33 Forbes Argentina. *El impacto de la Inteligencia Artificial en el acceso al crédito y la inclusión financiera*.

34 Responsibility by design: an approach that incorporates ethical and responsibility principles from the beginning of the AI system design and development stages. This principle ensures that equity, transparency, and accountability, among other characteristics, are built in, and thus minimizes biases, promotes explainability, and ensures that AI is used ethically and responsibly.

35 VTT Research. Ethical artificial intelligence as a competitive advantage.

36 Capgemini Research Institute. Why addressing ethical questions in AI will benefit organizations. This survey of 1,580 executives in 510 organizations and 4,400 consumers demonstrated the relationship between consumer trust and ethical interactions with AI systems.

companies to follow. Moreover, the provision of early support to companies with a direct impact on the vulnerable population serves as an example of how RAI can be adapted in critical settings. Lastly, the promotion of RAI in high-innovation environments, like the entrepreneurial community and its networks (such as that of the investment agents), creates an enormous opportunity to enhance the environmental, social, and governance culture as well as the impact vocation of these communities.

2.2 The project objective is to help improve the competitiveness of companies in Latin America and the Caribbean, especially startups and MSMEs, through the adoption of RAI standards for developing and using AI, and to foster an inclusive digital economy. The initiative is called **RAI 4 Growth**,[37] since it seeks to strengthen the regional ecosystem as a global leader in the development of ethical, responsible, and competitive AI, by empowering and strengthening the value propositions of companies (MSMEs and startups) and entities that support entrepreneurship (e.g., investment funds, incubators, accelerators, and business associations) with tools, resources, and skills. The project aims to ensure that the AI technology developed and used in Latin America and the Caribbean is not only at the forefront of innovation and productive transformation through digitization, but also fosters its equitable and accessible deployment, with a view to social welfare. To do so, it will set up a framework for RAI, which will include training, tools, certifications, and specialized measurement plans, depending on the type of solution, to benefit the most vulnerable population in particular.

2.3 The project will seek to identify MSMEs, startups, and investment agencies interested in using and developing RAI that will serve as leaders and examples for the rest of the ecosystem. In addition, RAI 4 Growth will hold specific activities to, using RAI, strengthen AI-based solutions that affect the climate or impact populations that are vulnerable or marginalized (e.g., due to gender).

2.4 RAI 4 Growth will have three main components (dialogue, training, and tools), and will initially be implemented in Chile, Colombia, Ecuador, Mexico, and Uruguay. Except for Ecuador, all of these countries already have national AI strategies, have implemented pilot fAIr LAC projects, and have created networks of agents committed to the issue of RAI, which they hope to expand and further through this project. In turn, Ecuador has an emerging group of AI-based startups,[38] and is therefore considered fertile ground for implementing and testing the tools developed in this context of incipient development.

2.5 To maximize its positive impact, the project will aim to ensure that the tools, resources, and skill development initiatives included in its activities—which are to be made available for the ecosystem as a whole—are adopted by companies and support agencies committed to improving their competitiveness through responsible-by-design AI, as part of their environmental, social, and governance objectives, as well as by enterprises working in areas that have a significant impact

---

[37] Responsible Artificial Intelligence for Growth.

[38] F6S. 11 top AI companies and startups in Ecuador in July 2024.

on vulnerable populations,[39] women, and minorities, so as to make it possible to verify the benefits of mitigating the risks of AI, especially with regard to biases, exclusion, and discrimination.

2.6 The initiative will seek to consolidate an ecosystem conducive to the responsible development and adoption of AI that will make businesses more competitive without jeopardizing the inclusion, equity, and rights of people, by delivering actionable tools, training, and knowledge, and raising awareness of the issue. The development will fundamentally be based on the tools that have already been tested by fAIr LAC and Adigital; the project will consolidate and improve integration of these tools and the user experience.

RAI 4 Growth – Components

2.7 **Innovation.** The project is noteworthy due to its innovative character as an integrated, unifying program established as a tool to promote the competitiveness of the region's companies and, in turn, raise awareness among the participating companies and organizations of avoiding biases, including ones that could currently be adversely affecting the marginalized, poor, and vulnerable groups that interact with AI technologies. The program design emphasizes a holistic approach that not only adopts AI technologies, but also encourages their ethical and sustainable use within the regional business ecosystem, as well as the measurement of results. The initiative also pioneers the coordination of scattered efforts to foster RAI practices, through a business support program aligned with international standards and regulatory trends, thereby preparing companies to compete on the global stage. A mix of virtual and in-person activities will be carried out to leverage the ecosystem's existing efforts and the adoption of the tools to be deployed.

2.8 The project will focus not only on implementing technology, but also on developing the human capital needed to manage it responsibly and effectively, through specialized training courses and the creation of resources to foster a culture of

[39] These include: (i) financial inclusion; (ii) educational platforms; (iii) health platforms; (iv) employment platforms; (v) climate sustainability platforms; and (vi) platforms that incorporate and validate bias-mitigation techniques.

innovation, with training sessions, audits, and certifications, adhering to democratic norms and values and promoting social and economic inclusion. The project will be structured into the following components:

Component I: Building trust (IDB Lab: US$332,920; Counterpart: US$178,090)

2.9 The objective of this component is to promote awareness in the regional technological innovation ecosystem of the importance of RAI for the development of fair, transparent, and reliable technologies, resulting in the development of technically robust and competitive products and knowledge of regulatory trends in the field. This will be achieved through activities conducted by the executing agency to spread information and foster exchange among the various ecosystem stakeholders, especially companies, investors, governments, and regulators, as well as civil society and academia. The component will pay special attention to making the regional corporate sector's startups and MSMEs more competitive. In addition, it will take actions to raise awareness of the adverse effects on vulnerable and poor groups, in line with existing sector policies and regulations on nondiscrimination, as well as regulations applicable in other regions that may be part of the market objectives of the participating startups and MSMEs. To that end, the component includes the following activities:

(i) RAI outreach and awareness campaigns that highlight the opportunities and risks involved in AI and introduce issues relevant to the Latin America and Caribbean regional context, with regard to supporting companies' competitiveness as well as aspects that affect vulnerable populations, gender considerations, and climate sustainability.

(ii) Preparation of reports (like the RAI index in the "fAIr Tech Radar" innovation ecosystem) and broad dialogue and reflection sessions (with the public and private sectors and civil society) to publicize best practices, the value of incorporating the responsible use of AI in businesses, and the impact on marginalized, poor, and vulnerable populations in sectors in which AI is becoming more important, such as: (a) AI applications in the health sector, the risks of excluding vulnerable populations, and mitigation measures; (b) the digital divide and financial system exclusion; and (c) education technologies and gender biases.

(iii) Monitoring actions to enhance the effectiveness of component activities, such as: (a) registration and attendance at work meetings and events; (b) participant evaluation of meeting quality; (c) impact in the media, on social networks, and downloads of the published reports; and (d) qualitative impact of meetings and reports based on talks with important industry, society, and public policy players.

2.10 As a result, entities in the regional ecosystem are expected to be better informed and aligned with RAI policies and national and international regulations; to be prepared to comply with regulations; and to identify opportunities for collaboration to develop and adopt a more inclusive, transparent AI that promotes trust and, therefore, business competitiveness. Specifically, the component is expected to:

(i) convene at least 5,000 participants in the various campaigns and events to be held; (ii) publish communications material with guidelines and practices for the responsible development and use of AI; (iii) systematize best RAI practices in Latin America and the Caribbean and the collaborative community around them; and (iv) improve public knowledge and perception of RAI in the target countries.

Component II: Ecosystem skills and talent development for responsible AI development (IDB Lab: US$374,210; Counterpart: US$504,480)

2.11 Companies that use AI and those that implement models and systems both need to understand what it means to implement RAI. In the same way that it took time and effort to grasp the minimum requirements for "privacy by design,"[40] implementing "responsibility by design" will require understanding the ethical, regulatory, technological, and governance foundations, complemented by use cases of companies that pioneer its implementation. The objective of this component is to build the main channel for delivering tools and certifications, emphasizing the creation of upskilling programs with specialized content on the risks of AI and RAI to train entrepreneurs, MSMEs, investors, coordinating agencies, and their teams. The project will achieve this objective through Adigital's previously developed knowledge offerings and fAIr LAC's module offering designed to incorporate RAI content into existing incubation or acceleration programs for AI ("AI-first or AI-enabled") companies. As part of this component, the executing agency will implement the following activities, using an iterative, continuous improvement approach:

(i) Development of RAI training modules to be implemented in collaboration with partner agencies that support entrepreneurship, and directly targeting startups, MSMEs, and investors. These modules may be in person, hybrid, massive open online courses (MOOCs), or specialized seminars. Topics include: introduction to RAI, biases in AI, explainable algorithms and transparency in AI, AI for the inclusion of persons in the socioeconomic fabric, RAI with a focus on the gender perspective, RAI for sustainable development, fundamentals of RAI for investors, AI governance and regulation for investors, assessment of risks and opportunities in AI startups, evaluation of AI providers for MSMEs, critical analysis of AI systems, and open-source AI tools for MSMEs. During implementation, other important topics may be identified, based on the markets' maturity level.

(ii) Training of trainers program, to promote the sustainability of the efforts to build capacities in startups, MSMEs, and investors, focused on key concepts, practical applications, and the critical evaluation of RAI solutions. The training modules will seek to balance theory and practice, and will include a pedagogical component to teach effective instructional techniques for adults. The call for participation in this training opportunity will be open and supported by RAI 4 Growth's

[40] Privacy by design: approach that incorporates the protection of personal data into the system from the start of the design and development stages, to ensure that privacy will be a built-in characteristic. This principle minimizes data collection, promotes anonymization, and guarantees user control over their own data.

collaboration networks. Ideally, participants should have prior training in technology development.

(iii) Specialized mentoring. Based on demand, teachers and experts participating in the training programs may mentor the participating entities to provide them with specific guidance on their path towards RAI, so as to properly prepare them for certification, auditing, and legal or regulatory compliance processes, and for using the tools/service platforms developed in Component III.

(iv) Pilot RAI training with startups and MSMEs. The component will carry out a pilot of the main RAI training programs to verify their quality, interest in them, and their impact on the organizations.

(v) Open offering of training programs based on improvements and lessons learned from the pilot experience, for which the project will: (i) contract a learning platform to supplement in-person training activities; (ii) hold marketing campaigns and events, directly or with the support of the academic and training collaborators; and (iii) maintain a section called the "RAI Academy" on the Latin American RAI platform to provide participants with technical and academic support.

(vi) Monitoring of training session results, in terms of lessons learned, the experience of participants, and the practical implementation of learning at 6 and 12 months.

(vii) Model for the sustainability and scaleup of the training programs through the "RAI Academy," which will offer training at reasonable prices to companies, institutions, and interested individuals. The platform will offer free modules (freemium), paid modules, and specialized mentoring plans for implementing RAI.

2.12 This component is expected to yield: (i) systematized content for RAI training that covers ethical AI principles, data governance, bias mitigation, environmental sustainability practices, and compliance with emerging regulations, and includes case studies and practical exercises; (ii) partnerships with academic entities, industry experts, and regulatory agencies for implementation of the curriculum; (iii) systematized guides for RAI curriculum trainers; (iv) certified trainers for the RAI curriculum; (v) RAI training pilot for startups and MSMEs led by trained trainers; (vi) systematization of improvements and continuous feedback for the program's institutional, technical, and financial sustainability; (vii) systematized reports on program participants' learning, attitudes, and RAI practices; and (viii) validated portfolio of advanced training topics for RAI experts.

Component III: Tools, services, and certifications—resources for RAI risk management (IDB Lab: US$853,010; Counterpart US$817,430)

2.13 The objective of this third component is to create and promote tools, services, and certifications that help companies and investors detect risks and implement simplified processes for measures to mitigate, track, and correct them for the responsible development and use of the technology. To that end, a multidisciplinary team will be formed to offer specific recommendations for mitigating and monitoring the identified risks, perform algorithmic tests, conduct

impact assessments, and develop tools and certifications for RAI use, wherever possible incorporating the mitigation of adverse impacts and achievement of positive ones for economically vulnerable and marginalized populations, as a result of deploying AI with a gender perspective and the ability to measure and offset the carbon footprint of systems that use AI models.

2.14 The project will start with the existing tools provided by IDB Lab and Adigital, which will be used to make RAI 4 Growth operative from day one. New services and certifications will be created through the activities described below and in coordination with the participating companies. However, merely creating tools is insufficient, and within RAI 4 Growth, Adigital is planning to launch an RAI framework that will help MSMEs, startups, etc., effectively implement RAI to make an impact.

2.15 The RAI framework to be developed up involves an advanced, comprehensive approach that, based on the steady relationship with industry, academia, society, and regulators, will ensure that AI systems are developed and used in an ethical, transparent, and equitable way. To that end, various resources and practices will be incorporated to ensure compliance with the highest standards of responsibility and ethics at every stage of the AI-development life cycle, thereby establishing a regional standard for startups and MSMEs that use and/or develop AI and a high-value-added tool for investment funds, banks, and any institution that needs to make decisions about using or financing AI.

2.16 The following are the main elements of the RAI 4 Growth responsible AI framework: (i) **Tools:** advanced technical resources and methodological guides to facilitate implementation of responsible practices at all stages of the AI-development life cycle, such as algorithms that promote transparency, software for detecting and mitigating biases, services to ensure equity in data, etc.; (ii) **Certifications:** certification programs that verify AI systems' compliance with best practices, regulations, and international ethics and responsibility standards, covering critical aspects such as transparency, explainability, data privacy, algorithmic equity, and accountability; (iii) **Methodologies and best practices:** based on research and expert consensus, in order to guarantee the responsible development and implementation of AI. These methodologies and best practices include ethical design guidelines, social impact assessment procedures, strategies for including interested parties and ensuring they participate, and protocols for managing risks and mitigating potential harm; (iv) **Evaluation and continuous auditing:** to ensure that the AI systems stay aligned with the principles of responsibility and ethics over time. This element includes periodic reviews, independent audits, and mechanisms for continuous feedback from users and other interested parties; and (v) **Measurement plan:** set of detailed, structured plans that establish metrics and key indicators to assess how an AI system applies RAI principles, and to measure their impact with regard to specific objectives like gender equity, support for vulnerable populations, and sustainability.

2.17 For this component, the executing agency will conduct the following activities:

(i) First tools and certifications aligned with the fAIr LAC tools (ethical self-assessment for entrepreneurs and risk-assessment tool for investors) and with the requirements of the Adigital transparency

certificate.[41] With the fAIr LAC tools and Adigital certificate, the component has the potential to create an aligned suite that will encompass everything from the self-assessment to the verified evaluation of the transparent use and development of AI systems. To that end, the component will strengthen technological support for the fAIr LAC tools and will integrate and standardize them with the Adigital certificate. Experts will carry out development and maintenance activities, and pilot tests will be used to verify the potential to form a representative group of startups and MSMEs that have already implemented these tools in high socioeconomic impact sectors.

(ii) Portfolio of services for companies. The component will create new complementary tools for RAI impact areas and to analyze sectors or spheres involving vulnerable populations and the gender perspective. Specifically, the component will develop: (a) a new gender bias analysis certification based on Adigital's algorithmic transparency certificate, grounded in the ethical principle of algorithmic justice/equity and state-of-the-art implementation of technological capacities; (b) a new carbon footprint analysis certification based on the Adigital algorithmic transparency certificate; (c) a gender bias mitigation measures toolbox; and (d) a toolbox for reducing AI systems' carbon footprint, to include considerations on the selection, training, and launch of models.

(iii) Third-party tools. In addition to the above, the component will identify a pool of existing tools that will progressively address the requirements arising from risk identification and analysis for RAI certification, based on the level of companies' maturity and their degree of technological development, as well as the risk level of their solutions. Accordingly, the component will include models geared towards facilitating: (a) conformity assessments and self-assessments; (b) audits and tests for algorithmic biases and precision; and (c) algorithmic impact assessments. The following table provides a nonexhaustive list of groups of tools and techniques that may be implemented in the project.

[41] Adigital. Algorithmic Transparency Certificate.

Table 1. Responsible AI tools

Responsible AI tool	Description
Conformity assessments and self-assessments	Initial assessments to identify risks and establish mitigation measures ▪ Data governance and privacy ▪ Transparency and explainability ▪ Justice and nondiscrimination ▪ Precision and performance of the models ▪ AI security, trustworthiness, and robustness ▪ Accountability
Audits and bias tests	Tools and in-depth processes to evaluate AI systems' performance and equity, to ensure they do not exhibit biases or undesired behaviors ▪ Technical evaluations of the implementation of AI/machine learning operations (in the training and inference stages) ▪ Database evaluations ▪ Data architecture and AI models ▪ Deployment and integration ▪ Detection of biases in the models ▪ Maintenance and updating ▪ Staff training and education
Algorithmic impact assessments	Measures to assess the social impact of using AI models and their influence on decision-making

(iv) Measurement plans (general as well as specific to AI systems, focused on vulnerable populations, gender bias mitigation, and sustainability). These plans set out: (a) the definition of key indicators that, based on the use of the RAI 4 Growth tools and certifications, will make it possible to evaluate aspects that are important for RAI systems, namely, efficacy, equity, transparency, the mitigation of different types of biases, and privacy; (b) procedures and specific tools to gather and analyze relevant data that reflect the system's performance in terms of responsibility; (c) the frequency with which the various measurements and evaluations should be carried out; (d) instructions for establishing clear targets and specific criteria that the AI system must fulfill to be considered responsible; (e) mechanisms for documenting the results of the measurements and communicating the findings to the interested parties in a transparent, accessible way; and (f) guidelines, demonstrations, and use cases based on real RAI 4 Growth experiences that support implementation of measurement plans at the beneficiary companies. The project provides for the following measurement plans:

- General measurement plan. Valid for any type of AI system, this plan indicates how to use the most common RAI tools and the most appropriate indicators.

- Gender bias mitigation measurement plan for AI systems. Also based on the general plan, this plan will determine if the system operates equitably and without (or is getting rid of) gender discrimination. This plan includes indicators that measure the presence and correction of gender biases in the system data and results, transparency in decision-making, and equity in the system's impact on the different genders based on the system's objectives.
- Measurement plan for sustainable AI systems. This plan is formulated to assess and guarantee measurement of the systems' environmental impact and how the system design will make it possible to minimize this impact and promote sustainable practices. As in the foregoing cases, the plan explains how to use the tools available in RAI 4 Growth for its correct implementation.

(v) Services. Based on their use of the tools, especially the self-assessments and conformity assessments,[42] the companies will receive a number of actionable recommendations for bringing their systems into line with RAI principles[43] or for implementing those principles, which in some cases will require expert support. The types of services may include: database reviews, data governance policies, selection and evaluation of AI models, training in and auditing of models, deployment and integration, and maintenance and updating. This portfolio of services will be made available through an RAI platform that will give exposure to the specialized agencies that have the technical and operational capacity to carry out monitoring and corrective measures. This will evolve into the creation of the "Latin American Responsible AI Platform," developed to adapt to the needs of this nascent market. The project will hold open, competitive, and transparent calls to identify the agencies that will participate in this platform; these calls will incorporate criteria for institutional integrity, operational and technical capacity, and a diversity of offerings. Local partnerships may be established to improve exposure and outreach. The technological platform will initially be established as a website offering information, training, and access to RAI resources, and it will evolve, together with the evolution and maturity of the market, into a platform for RAI services.

(vi) Communications strategy. In addition to implementing RAI assessment and certification processes in ecosystem coordinating entities (e.g., investors, accelerators, and incubators), startups, and MSMEs that develop or use AI systems, the component will implement a communications strategy to facilitate calls for participating agencies. This strategy will include network management, a website, supporting material, workshops, guidelines, and practical cases.

---

[42] Conformity assessments: processes that confirm AI system compliance with measures and/or requirements primarily for risk management, data governance, technical documentation, record-keeping, transparency and provision of information, human oversight, precision, robustness, and cybersecurity.

[43] OECD. AI Policy Observatory. OECD AI Principles Overview.

(vii) Pilots. The component will conduct pilots of the risk evaluation and subsequent certification processes in high-impact sectors for vulnerable and high-risk populations, validating the quality of each process as well as its impact on the participating organization. **The agencies that participate in the pilot tests will be selected** through open calls, agreements with business associations and guilds, or partnerships with entities looking to strengthen value chains. An important criterion will be their impact on RAI practices in startups and MSMEs whose users, customers, or beneficiaries include poor and vulnerable population segments, women, and minorities, or those whose AI-supported products and services are subject to the risks of bias and discrimination. These pilots will help form and validate the first versions of the RAI framework.

(viii) Design of the RAI 4 Growth sustainable business model, based on the portfolio of direct and indirect services piloted under this component, as freemium and paid products.

2.18 This component is expected to yield: (i) an RAI framework that includes various tools, guides, and measurement plans; (ii) aligned and validated assessment and certification tools; (iii) guidelines and instructional documents for startups, small and medium-sized enterprises (SMEs), and support agencies on the use of validated RAI tools; (iv) systematized protocols for the rollout of advisory services focusing on AI risk management; (v) measurement plans for different types of criteria; (vi) operational digital platform and engagement reports; (vii) 100 strengthened ecosystem support agencies that have incorporated RAI processes; (viii) 400 certifications/tests performed with startups and SMEs; (ix) report on AI risks in sectors that are sensitive for vulnerable populations in the Latin America and Caribbean region; (x) 3 case studies on the positive impact of certification; and (xi) business model and expansion strategy for the documented RAI services.

B. Project outcomes, measurement, monitoring, and evaluation

2.19 The project is designed to support startups, MSMEs, investors, and business support agencies that develop, use, and finance AI-based technological solutions, paying special attention to high-socioeconomic-impact sectors like: companies with high growth potential, health and care companies, employability and contracting, education and literacy, financial inclusion, as well as critical and essential infrastructure, justice and legal assistance, migration, agriculture, and food security.

2.20 The principal project outcomes will be: (i) 250 AI-developer companies and 1,000 MSMEs using AI risk mitigation and responsible assessment tools; (ii) 350,000 women using or benefiting from the products and services of participating companies that implement RAI measures to mitigate gender biases; (iii) 3,000 vulnerable persons using or benefiting from the products and services of participating companies that have implemented RAI measures; (iv) 150 startups and MSMEs that incorporate methodologies to measure and estimate their AI-associated carbon footprint, a third of which incorporate offsetting or mitigation strategies; and (v) 20% of participating companies that are seeking investments and explicitly apply RAI obtain investments from at least

one investor whose investment thesis includes the use and implementation of responsible technologies.

2.21 The executing agency is responsible for monitoring the project, and will do so through a dashboard that will contain the outcome indicators as well as the operational parameters for project execution. With this dashboard, the executing agency will be able to: (i) communicate and report the outcomes to IDB Lab based on the RAI 4 Growth project agreement, on the platform that exists for that purpose; (ii) keep appropriate information to perform strategic and operational analyses of the project and each of the project components; this information will serve as an input for informing government bodies about the initiative; and (iii) strategically communicate information to the various stakeholders and partners.

2.22 The executing agency will be responsible for compiling the data and reporting the results and achievements pursuant to the project results matrix. The executing agency will develop a monitoring plan early in the project to ensure follow-up and measurement of the indicators. It will also submit periodic project status reports to the Bank, as well as a final project status report on the project outcomes after completion.

2.23 In the third year of execution, the executing agency will evaluate the effects on the beneficiary populations of the selected pilots and the AI technologies that adopt ethics and responsibility criteria. These evaluations will assess if the incorporation of these criteria reduces bias and discrimination in beneficiaries and users. The specific terms of the evaluation will be formulated based on the issues addressed by the pilots and the AI solutions that form part of the program.

III. Alignment with the IDB Group, Scalability, and Risks

A. Alignment with the IDB Group

3.1 The project is aligned with the IDB Group Institutional Strategy 2024-2030 and its strategic objectives of reducing poverty and inequality and bolstering sustainable growth, by fostering private initiative and responsible technology adoption (document CA-631).

3.2 The project is also aligned with the country strategies of the first countries to be targeted, especially with regard to: (i) enabling the economy of the future, improving productivity, and increasing investment in research and development (document GN-3140, IDB Group Country Strategy with Chile 2022-2026); (ii) equitable access to services and human capital development (document GN-2982, IDB Group Country Strategy with Mexico 2019-2024); (iii) spurring innovation, business development, and the use of digital tools (document GN-2972, IDB Group Country Strategy with Colombia 2019-2022); (iv) contributing to the development of the productive sector and technological infrastructure (document GN-3103, IDB Group Country Strategy with Ecuador 2022-2025); and (v) fostering an entrepreneurial culture and increasing private sector innovation (document GN-3056, IDB Group Country Strategy with Uruguay 2021-2025).

3.3 Furthermore, the project is aligned with the IDB Lab Business Plan (document MIF/GN-264) and its strategy for the MIF IV capital replenishment (document GN-3204), which focuses on entrepreneurial innovation as a vehicle to improve quality of life, by making new technologies accessible and promoting their responsible development and adoption. Accordingly, the project complements the efforts that have already been implemented in the region through the fAIr LAC initiative pilot projects, in line with the vision of building a portfolio that includes "AI for Development" and laying the foundations for an RAI system in the region. IDB Lab has spearheaded the following projects for building the infrastructure for RAI use: (i) "Market Opportunities for Technology Firms – Public Procurement of Responsible, Ethical, and Transparent Algorithms" (operation CH-T1246), executed by the Universidad Adolfo Ibáñez, to promote ethical data management and the responsible development and implementation of algorithms, automated decision systems, and artificial intelligence in the public and private sectors; (ii) "Efficient Social Services for Vulnerable Populations in Jalisco through Responsible Artificial Intelligence," (operation ATN/ME-17550-ME), executed by the Instituto Tecnológico y de Estudios Superiores de Monterrey, to take advantage of the responsible, ethical adoption of AI to deliver social services in Jalisco, by bringing together academia, civil society, and the public and private sectors to develop capacities, enhance the social reach of AI, and leverage technology's impact for more efficient social services. As for using AI for the common good, IDB Lab launched a call for innovation in gender and AI in 2022, which gave rise to three projects that use AI to mitigate the gender gap in employment, financial inclusion, and the use of algorithms: (iii) "Jobecam: the Workplace Diversity Booster" (operation BR-G1021) is a platform to reduce biases in hiring; (iv) "Quipu: Promoting Financial Inclusion for Women Microentrepreneurs" (operation CO-G1050) uses fair, alternative credit scoring; and (v) "QuantilAI: Algorithmic Audit" (operation CO-G1049), for the development of RAI by automating bias evaluations of algorithms.

3.4 At the IDB Group level, IDB Lab, through the fAIr LAC initiative, has worked in close coordination with the public sector via the following technical-cooperation operations: (i) fAIr Costa Rica, through operation CR-T1242, "Support for the Guiding Framework for the Responsible and Ethical Use of Artificial Intelligence in Costa Rica," to build skills for responsible technology use, foster a dynamic ecosystem, and empower citizens to boost industry 4.0; (ii) operation RG-T3638, "Design and Execution of Pilot Projects under fAIr LAC Initiative;" and (iii) operation RG-T3450, "fAIr LAC – Responsible Artificial Intelligence for Efficient and Individualized Provision of Social Services for All," which spearheaded actions in Uruguay and Colombia. In Uruguay, the technical-cooperation operation was carried out to promote and strengthen the responsible use of artificial intelligence, especially in the public administration, through implementation of the AI Strategy for the Digital Government. The strategy was submitted for public consultation, and contributions were received, analyzed, answered, and in some cases incorporated into the final text of the current version of the AI Strategy. In Colombia, the initiative was executed to promote significant impact on the generation of public policies and digital transformation projects within the country's public agencies that use AI systems. The initiative also fostered access to and dissemination of knowledge for the implementation and deployment of AI in the country and region. As outcomes, this initiative developed the National Policy for Digital Transformation and AI, as

well as an Ethical Framework for AI in Colombia. Along these lines, the project aims to further the initiative's impact based on the knowledge and networks built, as well as by encouraging public-private cooperation for development of an RAI ecosystem.

3.5 Paris alignment: The project is considered to be aligned with BB1 (alignment with mitigation goals).

3.6 Sustainable Development Goals (SDGs). The project will contribute to: (i) SDG 5, on gender equality; (ii) SDG 8, on access to decent work and economic growth; (iii) SDG 9, on building resilient infrastructure, promoting inclusive and sustainable industrialization, and fostering innovation; and, by impacting the development and adoption of AI algorithms that mitigate potential biases and discrimination; and (iv) SDG 10, on reducing inequalities.

B. Additionality and scalability

3.7 From the design phase, the path to sustainability and scaleup of the project has been taken into account, through development of a platform of products, tools, and services that can be adopted as modules in accordance with each ecosystem's level of development. Specific value is added based on five pillars:

- Fostering trust: By focusing on transparency, explainability, and responsibility in AI operations, the project will help build trust in the technologies among consumers, users, investors, and regulatory agencies.
- Inclusion and equity: The initiative promotes adoption of RAI by startups and MSMEs in Latin America and the Caribbean, which will improve their competitiveness and directly benefit populations that are vulnerable or poorly represented, by mitigating the risks of biases and discrimination and promoting responsible data use, respect for people's rights, transparency, explainability, and accountability, and by mitigating the risks of automation-driven job displacement.
- Talent and skill development: Through the training and tool development programs, RAI 4 Growth will help close skills gaps in the region, enabling local companies to compete in a global market.
- Environmental sustainability: The project focuses on reducing the carbon footprint of AI technologies, raising awareness, and calling for action to include sustainability as one of the ethical principles to consider under the RAI approach.
- Collaboration among stakeholders and interested parties: Effective collaboration among various stakeholders in the ecosystem, including companies, investors, governments, regulators, civil society, and academia, is essential for scaling up the project's impact. By facilitating the exchange of knowledge, experiences, and resources, the project will create a regional climate conducive to encouraging the adoption of RAI. This strategic collaboration is structured through partnerships, specialized consulting services, the provision of technology infrastructure, and joint participation in training and research programs. Achieving synergy among these stakeholders will accelerate implementation of technological solutions and enable

distribution of their benefits, enhancing development and competitiveness in the region's various contexts and levels of technological maturity.

3.8 Scalability will also result from: (i) the project strategy of creating partnerships with strategic associates that add value in diverse critical areas. These collaborations may also take the form of consultations, in the evaluation, certification of responsibility, and risk mitigation stages; partnerships for the provision of technology infrastructure and tools; or participation in the training and research programs; (ii) future market demand for the services to be developed by the project, in a context in which AI is increasingly used and must be accompanied, whether due to regulations or in a voluntary fashion, by the adoption of tools to minimize risks; and (iii) the trend towards development of legal and regulatory frameworks that will accelerate the responsible adoption of AI, motivated, for example, by the existing regulatory convergence agreements among the Latin America and Caribbean region and other regions such as the European Union.

3.9 RAI 4 Growth's consolidation of tools, generation of capacities, and outreach will be key to further positioning the fAIr LAC initiative and the pilot projects as agencies for coordination and preparation to ensure RAI technologies serve as an engine to alleviate social and environmental problems. IDB Lab's additionality results from its experience with the fAIr LAC initiative, connections with key players in the regional ecosystem, and emphasis on piloting models with a socioenvironmental impact, in which the responsible development of technologies like AI is a basic pillar for progress on inclusion and equity issues.

3.10 The project's scalability can also be measured through the potential growth of AI (in terms of use and development in the region and in the countries in which RAI 4 Growth will operate in the next three years). The potential market of RAI 4 Growth project beneficiaries is estimated at 1.2 million MSMEs in the Latin America and Caribbean region,[44] and 500,000 in the countries where the RAI 4 Growth project is implemented. Moreover, there are expected to be a potential 2,000 to 5,000 AI-developer companies in the project countries in the next 10 years.[45] As noted above, RAI 4 Growth's objective is to have 1,000 AI-user companies, 250 AI-developer companies, and 100 investors participate in the project and implement RAI. In addition, the project will raise awareness and build skills (provide training) for over 5,000 public sector, private sector, and civil society stakeholders, laying the foundation for expansion to the rest of the ecosystem. Consequently, the average cost per user during the initial pilot stage is estimated to be around US$500; this amount will decrease once the technology and management infrastructure are fully operational.

C. Project and institutional risks

3.11 **Risk that the ecosystem is not conducive to the project:** If the target ecosystems are immature and unaware of the importance of RAI, the number of interested startups and MSMEs could be limited, which would impact the project's

---

[44] Conservative estimate that corresponds to 10% of the estimated number of MSMEs in the Latin America and Caribbean region (*Panorama Digital de las Micro, Pequeñas y Medianas Empresas (MiPymes) de América Latina 2021*, Maastricht University, Latin American and Caribbean Economic System).

[45] Adigital estimates based on secondary information and a market study conducted in the first half of 2024.

objectives. This risk will be **mitigated** through the outreach activities included in the project components, which will disseminate specific knowledge on how the RAI approach contributes to making businesses successful and achieving socially beneficial solutions. This is why the project will address business competitiveness as a means of raising awareness.

3.12 **Reputational risks and "ethics washing":** Some companies may be motivated to participate out of their commercial or reputational interests rather than a true commitment to RAI. As a **mitigating factor**, the project will establish rigorous evaluation and verification measures for obtaining certifications, as well as engage third parties to provide external validation of the companies' actions.

3.13 **Risk that the participating entities will fail to correct and mitigate the risks of AI:** Some companies could choose to not correct their risk mitigation models or incorporate transparency standards. As a **mitigating factor**, the project will create incentives to motivate the companies to effectively correct their models. These may include establishing conditions for certification, support, financing, advisory services, or access to services, and measuring the impact of the corrections.

3.14 **Risk of adverse algorithmic impact:** Stemming from the foregoing, some participating companies could choose not to correct environmental and social risks associated with AI-related activities. To **mitigate** this risk, the project will seek to incorporate participative, multi-actor design approaches in high-risk sectors, to make it possible to establish preventive monitoring and auditing measures.

3.15 **Deployment risk in the target countries:** The executing agency, headquartered in Spain, may not have the necessary local networks in the countries where the project will be executed, due to the geographic distance and its lack of prior experience in the region. As a **mitigating factor**, the project will leverage implementation of the activities through collaboration with the network of stakeholders built by IDB Lab and fAIr LAC. In addition, it will organize open calls for new stakeholders.

3.16 **Political and regulatory risk:** Across the Latin America and Caribbean region, there may be differences in the legal and regulatory frameworks for AI, as well as in national AI policies. This makes it difficult to standardize the services to be offered by the project. This risk will be **mitigated** by monitoring regulatory initiatives and consultation with local experts to understand significant regional differences while promoting compliance with accepted international standards that may be more restrictive than national regulations.

IV. BUDGET INSTRUMENT AND PROPOSAL

~~4.1~~ The total cost of the project is US$3.2 million, of which US$1.7 million (53%) will be provided by IDB Lab as nonreimbursable technical-cooperation funding and US$1.5 million (47%) will be provided by the counterpart. Adigital, Asociación Española de la Economía Digital, will be the executing agency for up to US$1,400,000 of the IDB Lab financing and will contribute counterpart resources. IDB Lab (Multilateral Investment Fund) will be the executing agency for up to US$300,000 of its own resources.

Indicative~~Summary~~ budget (US$)

Project components	IDB Lab Nonreimbursable technical cooperation		Counterpart		Total
	Executed by IDB	Executed by Adigital	In kind	In cash	
Component I: Building trust	50,000	~~332,920~~ 282,920	19,410	158,680	511,010
Component II: Ecosystem skills and talent development for responsible AI development	0	374,210	299,860	204,620	878,690
Component III: Tools, services, and certifications—resources for RAI risk management	250,000	~~853,010~~ 603,010	447,545	369,885	1,670,440
Translation	0	75,000	0	0	75,000
Contingencies*	0	64,860	0	0	64,860
Total	300,000	~~1,700,000~~ 1,400,00	766,815	733,185	3,200,000
% of financing	9%	~~53%~~44%	24	23	100

* The use of contingency funds requires the Bank's prior no objection.

V. Executing Agency and Implementation Structure

A. Description of the executing agency

5.1 As described in section B below, the project will be jointly executed by IDB Lab and Adigital.

5.2 Adigital, the Spanish Association for the Digital Economy, will be the principal executing agency for this project and will sign the respective agreement with the Bank. Adigital is a nonprofit organization formed by over 500 digital companies in Spain in key sectors of the digital economy, united by a shared objective: create an optimal climate for the development and growth of the digital economy, which will make it possible to achieve a more open, competitive, and sustainable society, with a special focus on digital ethics. This positions Adigital as a leader in responsible digital development in Spain and the European Union, and demonstrates its capacity for managing partnerships in the public and private sectors.

5.3 Adigital has experience successfully designing, building, and launching digital seals and certifications for the digital economy, and has adapted and responded to the evolution of the digital economy over the past 20 years. Its most notable programs are: (i) Confianza Online (Online Trust), which seeks to foster trust in online transactions and website browsing by awarding a seal to websites that have undergone an evaluation based on more than 30 criteria, including but not limited to legal compliance and ethical standards. This seal has been recognized for increasing user confidence and improving search engine rankings. It also entails adherence to of a code of ethics covering areas such as personal data protection and e-commerce; and (ii) the Robinson List, which allows individuals to reduce unwanted advertising by blocking the companies that carry out these campaigns,

which are required to consult this list, and may be penalized if they fail to comply. This service has been validated and recognized by the Spanish Data Protection Agency.

5.4 Regarding transparency and explainability of the use and development of AI, Adigital has launched several key initiatives to raise awareness and boost implementation in Spanish companies: (i) Adigital Academy, a recent winner of the 2024 Expo e-Learning award;[46] this academy provides training for members on RAI issues, including risk evaluation, transparency, explainability, model cards, and reinforcement learning from human feedback; (ii) the responsible technology repository, which includes courses, articles, short videos, and overviews of RAI products and other emerging trends; and (iii) the certification of algorithmic transparency for AI systems, in line with current international and European Union policy and regulatory developments; this certification evaluates how entities manage AI systems and their underlying models. This certification, built jointly with experts on technology, regulation, and public policies, was presented at UNESCO's Global Forum on the Ethics of AI,[47] and is included in the OECD catalog of trustworthy AI tools.[48]

5.5 To carry out the actions in the target countries, the project will identify agencies whose missions include addressing RAI issues and supporting the ecosystem of startups (e.g., investment, acceleration, and incubation) as well as smaller companies. Initially, the project will conclude cooperation agreements with entities that already have a relationship with fAIr LAC in these countries. It will then expand its reach through open calls, whose terms will be jointly established with the Bank and will include considerations for analyzing the entities' integrity.

5.6 **Integrity review:** The project team, with assistance from the Office of Institutional Integrity, performed a due diligence review of the project and identified integrity and reputational impact risk indicators, as well as mitigating factors. The integrity and reputational impact risks posed by this project are considered to fall within the IDB Lab level of tolerance for such risks. For more details, see the integrity annex.

B. Implementation structure and mechanism

5.7 To ensure that the project is appropriately coordinated with the fAIr LAC initiative, generates public knowledge goods, and coordinates with the rest of the IDB Group, IDB Lab will execute up to US$300,000 of its own funds for the implementation of project activities, ensuring monitoring of the operation as a whole, the quality of the progress reports prepared by Adigital, and proper project governance.

5.8 The activities to be executed by the Bank will be implemented by IDB Lab at Headquarters under the leadership of the Chief of the IDB Lab Discovery Unit, who may delegate coordination to the Team Leader or other staff members.

---

46 Expo e-Learning.

47 Adigital. Artificial Intelligence.

48 Adigital. *La OCDE incluye el Certificado de Transparencia Algorítmica de Adigital en su Catálogo de Herramientas y Métricas para la IA Confiable*.

5.9 The activities and resources financed by IDB Lab, to be executed by the executing agency, will include establishment of an execution unit that will take into account the structure needed to execute project activities and effectively and efficiently manage project funds. The executing agency will also be responsible for submitting progress reports on project implementation.

5.10 For effective governance, the project will establish several committees to provide technical, operational, and strategic support, as described below:

5.11 **Board of directors:** This board is the project's main decision-making body. It is made up of high-level representatives of all the key affiliated and associated agencies (at the outset, Adigital and IDB Lab, with companies and institutions that are brought into strategic partnerships to potentially be added later). The board's role is to establish and update the project vision and strategic direction, ensure achievement of results, approve major plans and budgets, and supervise project management at the macro level.

5.12 **Execution team:** This team acts on the direction of the board of directors and is formed by the project's operational leaders (Adigital as the executing agency, IDB Lab as co-executing agency and principal driver, strategic partnerships, and potentially and temporarily, nonstrategic partners that should be included in decision-making on specific components for strategic reasons). The team's primary role is to implement the strategies and policies established by the board of directors, manage day-to-day operations, and ensure that the project moves forward as planned in the results matrix.

5.13 **Advisory committee:** Made up of experts on AI, ethics, and regulation, as well as other representatives, this committee will advise the execution team and board of directors on specific questions regarding responsible implementation of the technology, ensuring that the project adheres to international standards and best practices in RAI.

VI. Compliance with Milestones and Special Fiduciary Arrangements

6.1 **Disbursement by results, fiduciary arrangements.**

6.2 Adigital will agree to comply with IDB Lab's standard results-based disbursement arrangements, procurement, and financial management policies applicable to the private sector, consistent with the 12 June 2019 version of the Financial Management Guidelines for IDB-financed Projects (OP-273-12) and the Guide for Milestones-based Management and Financial Supervision for IDB Lab and Social Entrepreneurship Program Technical Cooperation Projects.[49]

6.3 IDB Lab will follow the same results-based disbursement modality, procurement policies, and financial management mechanisms established by the Bank.

6.4 **Results-based disbursement.** The IDB Country Office in Chile will supervise the project, monitoring it in compliance with the results- and risk-management policies (compliance with milestones) established by IDB Lab in April 2008.

---

[49] Guide for Milestones-based Management and Financial Supervision for IDB Lab and Social Entrepreneurship Program Technical Cooperation Projects.

VII. ACCESS TO INFORMATION AND INTELLECTUAL PROPERTY

7.1 **Access to information.** In accordance with the Bank Access to Information Policy, this document will be made available to the public upon approval.

7.2 **Intellectual property.** The intellectual property of all project works and outcomes not based on pre-existing works of each of the parties will belong jointly to the Bank and to Adigital (project deliverables).

7.3 Each party may grant third parties a license to the project deliverables; this license will be nonexclusive, free, and for noncommercial purposes, to publicly use, copy, distribute, reproduce, exhibit, and execute the project deliverables.

7.4 If the project deliverables are used by third parties under a license granted by Adigital, their use and distribution will adhere to the Commons IGO 3.0 BY-NC-ND license, and they will display the following text:

7.5 "Copyright © [year of first publication] Spanish Association for the Digital Economy (Adigital) and Inter-American Development Bank (IDB). This work is protected under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 IGO (CC-IGO BY-NC-ND 3.0 IGO) license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode) and may be reproduced for any noncommercial use with credit given to Adigital and the IDB. Derivative works are not allowed. Any dispute over the use of RAI 4 Growth project works that cannot be settled amicably will be subject to arbitration in compliance with United Nations Commission on International Trade Law rules. Use of the RAI 4 Growth project name for any purpose other than attribution and use of the logotype will be subject to a separate licensing agreement and are not authorized under this CC-IGO license. Note that the URL link includes additional terms and conditions of this license."

7.6 The preexisting works of each of the parties, which are made available for use within the scope of the RAI 4 Growth project, and their improvements incorporated within the project framework, will remain under the ownership of each entity, as specified below:

IDB	Adigital
Self-assessment tool for entrepreneurs, 3S	Adigital algorithmic transparency regulations and certification
fAIr Venture investment fund risk assessment tool	Adigital Academy preexisting courses[50]

7.7 Both parties will guarantee that project execution does not and will not infringe the rights of third parties and agree to execute all activities necessary for both parties to exercise the rights set forth herein without limitation.

[50] The list of preexisting courses is available in the project's technical files.

7.8 Both parties agree to include assignment of the respective intellectual property rights, including copyrights, to both parties in all contracts entered into under the project with consultants involved in development of the project deliverables.

7.9 Both parties may disclose, reproduce, and publish any project-related information and include therein both parties' names and logotypes.