

AI IMPACT SCALING PROGRAM

Project Proposal

Predictive Analytics & Scenario Simulation Engine for Education Policy in Colombia

Organization: Fundación Empresarios por la Educación (ExE)

Country: Colombia

Focus Area: Education — AI for Social Impact

Date: March 2026

Program: AI Impact Scaling Program — Tech To The Rescue

1. Executive Summary

Fundación Empresarios por la Educación (ExE) is a Colombian non-profit organization that has spent over a decade building one of the country's most comprehensive educational data platforms — the Educational Management Observatory (Observatorio a la Gestión Educativa). The platform currently integrates dozens of national and international data sources and enables descriptive analysis and visualization for policymakers, education authorities, and private sector actors.

The next — and most critical — step in this platform's evolution is to move from describing what has happened to predicting what will happen and simulating what could happen. ExE is seeking support from Tech To The Rescue's AI Impact Scaling Program to build a Predictive Analytics and Scenario Simulation Engine that will empower decision-makers to:

- Forecast baseline trajectories of key education indicators (dropout, learning outcomes, equity gaps) over a 3–5 year horizon.
- Simulate how those trajectories would shift under different policy or investment decisions — helping governments and private funders allocate resources where they will have the greatest impact.
- Interact with projections and scenarios through a natural language AI interface, making insights accessible to non-technical users.

The potential impact is significant: Colombia spends billions of pesos annually on education programs and initiatives. Even marginal improvements in the quality of resource allocation decisions — guided by evidence-based simulation rather than intuition — could redirect investment toward the students and territories who need it most.

Support Requested from Tech To The Rescue

Development of AI/ML predictive models and scenario simulation tools
Technical advisory from data science and machine learning experts
Mentorship and capacity-building for ExE's internal data team

2. About the Organization

Fundación Empresarios por la Educación (ExE) was founded with the mission of improving the quality and equity of education in Colombia through evidence-based decision-making, cross-sector collaboration, and the strategic use of data and technology to strengthen public policy.

ExE operates at the intersection of the public sector, private enterprise, and civil society — convening stakeholders from all three to coordinate action on education. Its work is grounded in the

belief that improving education outcomes at scale requires not just resources, but better information about where and how to deploy those resources.

2.1 The Educational Management Observatory

Over the past several years, ExE has built the Educational Management Observatory — a digital data platform that integrates large volumes of structured data from multiple national and international sources, including:

- National administrative education records
- Standardized test results (Saber 11 and others)
- Enrollment, retention, and dropout indicators
- Socio-economic and contextual variables by territory
- Data on education initiatives and programs across Colombia

The platform currently supports descriptive analytics and visualization, allowing users to explore historical trends, compare performance across territories, and monitor system outcomes. It has established ExE as a trusted data partner for policymakers at national and subnational levels.

2.2 Current AI Capabilities

ExE has taken early steps into AI by integrating a natural language interface that allows users to query the platform conversationally. This early-stage AI agent provides a foundation for the more advanced AI capabilities proposed in this project.

Existing Strengths	Current Gaps
• Robust multi-source data infrastructure • Longitudinal datasets covering 10+ years • Descriptive analytics & visualization tools • Early NLP agent on the platform • Internal team with statistical expertise • Strong relationships with MEN and territorial secretariats	• No predictive or forecasting models • No scenario simulation capabilities • No prescriptive analytics • Limited internal ML/AI expertise • No resource allocation optimization tools • AI agent not yet connected to forward-looking analytics

3. Problem Statement

Colombia's education system faces persistent structural challenges: high dropout rates in rural and low-income areas, persistent learning outcome gaps between territories, and inequitable distribution of public and private investment. Decision-makers at all levels — from national

ministries to local secretariats to private companies funding education programs — must regularly make high-stakes allocation decisions with limited forward-looking information.

The current state of the art, even at well-resourced organizations like ExE, is primarily descriptive: decision-makers can see what has happened historically, but they cannot yet see where the system is heading or how their decisions today would alter that trajectory. This creates a critical gap:

The Core Gap

Decision-makers are allocating billions of pesos in education resources based on historical snapshots — without the ability to project where those resources will have the greatest future impact, or to simulate the consequences of choosing one intervention over another.

This is not a data problem. ExE has the data. It is an AI and analytics capability problem.

Concretely, policymakers, territorial education authorities, and private sector investors currently cannot answer questions such as:

- If current trends continue, which departments will see the largest increases in dropout rates over the next 3–5 years?
- If a private company allocates COP \$500M to an education program, which territories and intervention types are most likely to improve Saber 11 results based on historical patterns?
- How would a policy change in school feeding programs or teacher deployment affect retention rates in rural municipalities over the next 4 years?
- Which territories are at the highest risk of widening equity gaps if current trajectories continue?

Without tools to answer these questions, resource allocation remains reactive and suboptimal. Funding goes to territories that are already visible or politically salient, rather than to those where investment would produce the greatest measurable change.

4. Proposed Project

4.1 Project Vision

ExE proposes to build a Predictive Analytics and Scenario Simulation Engine — an AI-powered capability layer on top of its existing Educational Management Observatory — that transforms the platform from a historical data viewer into a forward-looking decision-support system.

The engine will serve two primary use cases, with the second as the central anchor of the project:

#	Use Case	Value Delivered
1	Predictive Indicator Forecasting Baseline projections of key education indicators under current trends	Enables early warning systems and evidence-based prioritization of at-risk territories
2	Budget or Public Policy Scenario Simulation AI-guided simulation of how different investment decisions and policy choices would alter education outcomes	Empowers governments and private funders to allocate education resources or change public policies based on projected impact rather than historical inertia.

4.2 Target Indicators (Phase 1 Focus)

The first phase of the project will concentrate predictive and simulation capabilities on three indicators, selected for their policy relevance, data availability, and direct connection to resource allocation decisions:

Indicator	Policy Question Answered	Decision-Maker
School Dropout Rate	Which territories will see the sharpest increases in dropout if current trends continue? Where should retention programs be deployed?	MEN, Territorial Secretariats
Learning Outcomes (Saber 11)	If a private company invests in teacher training in a given territory, what improvement in results could be projected in 3 years?	Private sector funders, MEN
Territorial Equity Gaps	Which municipalities are at risk of falling further behind the national average? Where would investment have the highest equity return?	All decision-maker segments

4.3 Technical Architecture/Expected Outputs

The proposed engine will be built as a capability layer on top of ExE's existing data infrastructure, structured in three functional components:

Component	Description
-----------	-------------

Forecasting Layer	Time-series and regression models trained on ExE's longitudinal datasets to project baseline trajectories of target indicators at national and territorial levels over a 3–5 year horizon. Techniques may include ARIMA, gradient boosting, and neural network architectures depending on data structure.
Scenario Simulation Engine	A parameterized simulation framework that allows users to define policy or investment inputs (e.g., 'allocate X resources to teacher training in departments A, B, C') and receive projected impact on target indicators based on historical patterns and causal modeling. This is the core anchor of the project.
AI-Enabled Decision Interface	Evolution of ExE's existing NLP agent (a conversational AI advisor) to allow users to interact with the forecasting and simulation layers through natural language queries. When a user asks where to focus efforts or how to allocate a budget, the agent does not simply return indicators, it synthesizes the simulation outputs and historical patterns to recommend specific territories, intervention types, and sequencing of actions, always grounded in evidence and accompanied by the data reasoning behind each recommendation

4.4 Implementation Roadmap Proposal

The project is structured in four phases over a 12-month horizon:

Phase	Activities & Deliverables
Phase 1	Data Audit & Architecture Design - Audit of existing datasets for completeness, quality, and suitability for predictive modeling - Definition of data pipelines and feature engineering requirements for target indicators - Technical architecture design for the forecasting and simulation layers - Deliverable: Data readiness report + technical architecture blueprint
Phase 2	Predictive Model Development - Development and validation of forecasting models educational indicators - National-level baseline projections (3–5 year horizon) - Initial territorial disaggregation of models for priority departments - Deliverable: Validated forecasting models with documented methodology
Phase 3	Scenario Simulation Engine - Design of the scenario parameterization framework (policy and investment inputs) - Development of causal and simulation models linking inputs to indicator outcomes - Integration with ExE's existing platform and NLP agent - User testing with past scenarios

	Deliverable: Functional scenario simulation engine + integrated AI interface
Phase 4	Capacity Building & Sustainability - Training program for ExE's internal data team in predictive modeling and ML - Documentation of all models, pipelines, and methodologies - Definition of long-term maintenance and model refresh protocols - Deliverable: Trained internal team + sustainability roadmap

5. Target Users

The engine is designed to serve four distinct user segments, each with different analytical needs and decision-making contexts:

User Segment	Primary Use	Key Decision Supported
Ministry of Education (MEN)	National policy scenario modeling; monitoring trajectory of systemic indicators	National education budget priorities and policy design
Territorial Secretariats	Department and municipal-level projections; local risk identification	Local program deployment and resource allocation
Private Sector Funders	Investment scenario simulation for education programs and Corporate Social Responsibility (CSR) initiatives	Where and how to invest for maximum measurable impact
Researchers & Academia	Access to validated predictive models for evidence generation and policy analysis	Research agenda setting and policy evaluation

6. Support Requested from Tech To The Rescue

ExE is requesting three types of support through the AI Impact Scaling Program:

Type of Support	What We Are Asking For
AI/ML Development	Engineering support to build the predictive models, scenario simulation engine, and AI interface layer (conversational AI Advisory). This includes data scientists and ML engineers who can work with ExE's datasets to develop, validate, and deploy production-ready models. This is the primary and most critical ask.

Technical Advisory	Expert guidance from data scientists and ML practitioners on model selection, causal inference methodology, scenario parameterization design, and platform architecture. ExE's team will participate actively in all development work to maximize knowledge transfer.
Mentorship & Capacity Building	Structured mentorship program for ExE's internal data team covering machine learning fundamentals, predictive modeling workflows, model maintenance, and AI system design. The goal is to ensure ExE is not dependent on external resources to operate and evolve the system after the project concludes.

7. Impact - Why This Project

Colombia is at an inflection point in education policy.

After years of investment in data infrastructure, the country now has the raw material to make evidence-based decisions at a level of sophistication that was not previously possible. What it lacks is the AI capability to translate that data into forward-looking, actionable insight.

ExE sits at the center of this opportunity. With its established platform, its trusted relationships across the education ecosystem, and its commitment to building internal AI capabilities, it is uniquely positioned to deploy a scenario simulation and predictive analytics engine that reaches decision-makers at every level — from the Ministry of Education to a private company deciding where to invest its CSR budget.

With Tech To The Rescue's support, ExE can close the gap between data and decision and help redirect education investment toward the students and territories who need it most.