



ACHIEVING AI AT SCALE: CEMEX'S LEARNING JOURNEY

Barbara H. Wixom, *Principal Research Scientist*

MIT Center for Information Systems Research (CISR)

Ida A. Someh, *Senior Lecturer in Business Information Systems*

UQ Business School at The University of Queensland, Australia

Cynthia M. Beath, *Professor Emerita*

University of Texas, Austin

Recently, large established organizations have been growing business value by increasing the volume of AI models they have in production, an activity we call **scaling AI**.¹ For the past five years, MIT CISR researchers have followed more than fifty data monetization initiatives that have relied on machine learning to recognize patterns, draw inferences, and predict outcomes and thereby inform the scaling AI process.²

In our research, we observed that scaling AI is the result of a learning journey during which organizations learn how to generate value from AI models across the AI lifecycle. This lifecycle encompasses three phases: deploying AI models, proliferating AI models, and industrializing AI models.

A key finding from the research is that scaling AI without breaking the bank requires three facilitating elements:

- Highly liquid enterprise **data assets**: data assets that have been prepared and are widely available for easy reuse and recombination in value creation using AI
- An AI-savvy **workforce**: employees who can effectively participate in the AI lifecycle and cultivate AI solutions
- Prudent use of scarce and costly **AI resources**: the ability to economize on data science capabilities, tools, and expertise

Organizations must establish these elements to arrive at **AI at scale**, which we define as *the state at which an organization cost effectively manages large volumes of interconnected models in production*. The rewards for operating in this state are inspiring: maximized AI returns, AI-fueled business enablement, and AI-fueled competitive moves.

Ideally, organizations establish the three facilitating elements as part of their scaling AI learning journey. Table 1 summarizes how organizations can build data liquidity, develop workforce savviness, and leverage scarce resources as they learn how to deploy, proliferate, and industrialize AI models. Some organizations establish these elements sequentially, while in other cases they establish them concurrently. Regardless of the order of their learning journey, organizations must establish all three facilitating elements to achieve AI at scale.

In this briefing, we describe scaling AI at Cemex, one of our research sites. We focus on key practices that helped leaders at Cemex establish the elements to facilitate scaling AI and advance the company's learning journey, and we share Cemex's rewards for achieving AI at scale.

SCALING AI AT CEMEX

Cemex SAB de CV (Cemex)³ is a \$15.6 billion global construction materials company headquartered near Monterrey, Mexico. A decade ago, CEO Fernando A. González introduced a strategic emphasis on delivering superior customer service and directed investments to support a customer-focused digital transformation. In 2017, the company launched Cemex Go, the first end-to-end digital platform enabling the customer journey in the construction materials industry.

Learning to Deploy AI Models

In 2017, Cemex established a global center of excellence for data science called Global Data Science (GDS) to develop standard, scalable AI solutions for Cemex. Initially, GDS data scientists developed AI use cases as individual projects, collaborating in cross-functional teams with business domain experts to select data sources, prepare project data, choose algorithms, engineer features, train models, and validate results. Once a model was on track for deployment, GDS

1 Scaling AI is described in B.H. Wixom and C.M. Beath, "Pega: Driving Customer Engagement Using AI-Enabled Decision Making," MIT CISR Working Paper No. 449, June 2021, https://cizr.mit.edu/publication/MIT_CISR-wp449_PegaAIDecisionMaking_WixomBeath.

2 This research draws on a Q1 to Q2 2019 asynchronous discussion about AI-related challenges with fifty-three data executives from the MIT CISR Data Research Advisory Board; more than one hundred structured interviews with AI professionals regarding fifty-two AI projects from Q3 2019 to Q2 2020; an October 2021 survey of the fifty-two AI project teams; and eleven AI project narratives published by MIT CISR between 2017 and 2024.

3 This briefing case draws from I. A. Someh, B. H. Wixom, C. M. Beath, and R. W. Gregory, "The Cemex Journey to AI at Scale," MIT CISR Working Paper No. 463, July 2024, https://cizr.mit.edu/publication/MIT_CISR-wp463_CemexAIatScale_SomehWixomBeathGregory.

worked with Cemex’s Information Technology (IT) function to assemble a scrum team to ready the model for integration with existing process-enabling software and begin change management activities.

As GDS’s development teams and end users interacted with the AI solutions, they asked an ever-growing number of questions about how and why the AI model produced its results. To explain the AI model and help people trust it, GDS created what became known as the Magic Tools: a suite of visualizations and what-if simulations that shed light on model mechanics and outcomes. The Magic Tools analyzed data going into and coming out of an AI solution. The analyses allowed users, managers, deployment teams, and technical support people to identify for themselves the root causes of AI model problems they encountered, track model usage, and make model change recommendations.

During this phase of Cemex’s learning journey, GDS involved business domain collaborators in creating data assets for model training and in the evaluation of AI models, providing critical feedback to the data scientists. The process of model building helped the business understand AI and trust AI model results. Some of the helpful new trust-building and AI model deployment practices that emerged were baked into the self-service Magic Tools. This period of learning led to a wave of enthusiasm at Cemex about AI model usage.

Learning to Proliferate AI Models

To leverage rising employee interest in AI, IT offered the data, platforms, standards, governance, and training needed for AI model development. IT and data owners in Cemex’s business

areas cleaned and mastered data from core systems, establishing a data catalog to help people search for data assets they could trust. IT established a cloud-based central data lake; an enterprise portal that authorized employees could use to access a data view, use a business intelligence tool, or tap into an API; and temporary online experimentation spaces.

Cemex encouraged and captured data science use case ideas via the company’s formal employee innovation program as well as from IT, GDS, and the Cemex data science community at large. When the number of data science ideas mushroomed, leaders introduced a portfolio management framework they called the Speed to Value framework to assess the potential of novel use cases and make efficient use of model development resources. A wide range of ideas went into an innovation funnel, and when a use case proved to have global potential, it was assigned to GDS for global deployment. In these cases, GDS had to adapt core models and retrain them using local data; different countries often operated using different business rules, requiring changes to the AI model itself, its software context, or how it was used. GDS began developing “super models” with parameters that could be turned on and off to accommodate the needs of many local areas.

When the Speed to Value framework identified good ideas with only local potential, those use cases were assigned to leaders of the responsible business area for further development, with significant support from IT and GDS. For example, GDS began hosting data science communities of practice and offering advisory services, and an IT planning group helped teams establish metrics for tracking model value.

Table 1: Outcomes of Establishing Facilitating Elements During the AI Scaling Journey

	Learning to deploy AI models	Learning to proliferate AI models	Learning to industrialize AI models	AI at Scale
Data Assets	Evolve from source data to project data assets	Evolve to shared enterprise data assets	Evolve to highly liquid data assets	Enterprise data assets with high data liquidity
Workforce	Evolve from data scientists with AI expertise to project teams with AI expertise	Evolve to citizen data scientists and AI communities of practice	Evolve to an empowered AI-savvy workforce that deploys, uses, and sustains AI models	An AI-savvy workforce
AI Resources	Evolve from bespoke AI techniques to practices of repeatable routines	Evolve to shared enterprise capabilities	Evolve to AI work that leverages machines and non-data scientists	Prudent use of scarce and costly AI resources

During this phase, employees across the enterprise participated in AI model building. Cemex encouraged pervasive participation by making shared resources available for global and local teams. As a result, model reuse opportunities grew, and new AI models flourished. Many people participated in project teams, innovation forums, and data science communities of practice. This period of learning led to an explosion at Cemex of AI models at various stages of deployment.

Learning to Industrialize AI Models

Cemex leaders used the term “industrializing” to refer to the work involved in globally deploying AI models as well as that required for setting up processes for monitoring, supporting, and managing these models. Managing hundreds of industrialized AI models over time, in the face of change, was not simple. Initially, GDS manually monitored model health for drift and degradation of performance, to determine when models needed to be reparameterized or retrained using new data. As the number of AI models grew, this proved to be an overly resource-intensive task, and GDS leaders turned to automating data science work.

In 2022, GDS kicked off an internal project to implement MLOps⁴ practices. Fully embracing the MLOps platform functionality allowed GDS to reduce the time to deploy models from weeks or months to hours or days. Leveraging the company’s prepared and accessible data, data scientists used MLOps tools to compare team-developed models to models the platform suggested and to retrain models for new cities or markets.

IT installed and managed an MLOps platform as a part of its infrastructure management responsibilities. GDS had a separate team that took on the bulk of AI model production support, escalating problems to GDS’s data scientists only when absolutely necessary. Using the MLOps tools for data scientists’ tasks freed up those experts for other, more valuable activities.

During this phase, Cemex was exploiting a sophisticated capability to deploy, scale, and sustain interconnected AI models across the globe. When possible, the company automated routine tasks, such as drift management, offered self-service access to resources, and increasingly made use of MLOps tool functionality. Such activities improved data assets to become more comprehensive and standardized, and they increased shared enterprise data asset usage. This period of learning led both to broader participation by Cemex’s em-

powered, AI-savvy workforce in AI model management and to highly efficient use of data science expertise.

Managing AI at Scale at Cemex

Today Cemex Go incorporates a collection of AI solutions, some of which are interdependent—for example, a demand forecasting model informs an overbooking model, the overbooking model influences an order confirmation model, and a plant scheduling model relies on the order confirmation model results. The company’s globally scaled AI solutions account for an estimated value of \$30 million from a variety of efficiencies, such as reduced truck distances traveled, improved order availability, and lower energy consumption.

Up to 2022, Cemex primarily used AI solutions to enable an outstanding customer experience and to streamline operations. After making significant headway in this regard, Cemex leaders saw tremendous opportunities in leveraging the company’s digital capabilities, particularly in data science, beyond enterprise boundaries. For example, in mid-2022, Cemex launched a company called Arkik to sell the proprietary IT solution Cemex created for use by its Ready Mix businesses to competing ready-mix businesses in the form of software-as-a-service.

SCALING AI WITHOUT BREAKING THE BANK

Achieving AI at scale means your organization is using AI for business enablement efficiently and that it is positioned to pursue new competitive opportunities. The key is for leaders to establish data liquidity, workforce AI savviness, and the ability to leverage scarce resources as their organization masters the AI lifecycle.

The Cemex story suggests ways to keep AI learning journeys on track. First, leaders can surface novel practices that AI teams have invented to build trusted models, inventory them, understand why they help teams, and promote those that help move AI use cases from idea to solution faster, cheaper, and better. Second, leaders can get ready for model proliferation by investing in data literacy and making AI expertise, processes, and technologies accessible to employees throughout the organization. Connecting people to shared resources empowers them to exploit AI. Finally, leaders can look for ways to make the best use of their available data science expertise. As the number of AI models grows, models become interconnected, and change happens too fast for manual model management to keep up. At that point, organizations must leverage automation and focus the attention of their data science experts on the highest-priority work.

⁴ MLOps (Machine Learning Operations) refers to the overall process of building AI models using machine learning and managing those models in production environments over time.

MIT CENTER FOR INFORMATION SYSTEMS RESEARCH (CISR)

MIT CISR helps executives meet the challenge of leading increasingly digital and data-driven organizations. We provide insights on how organizations effectively realize value from approaches such as digital business transformation, data monetization, business ecosystems, and the digital workplace. Founded in 1974 and grounded in MIT's tradition of combining academic knowledge and practical purpose, we work directly with digital leaders, executives, and boards to develop our insights. Our consortium forms a global community that comprises more than seventy five organizations.

CISR RESEARCH PATRONS

AlixPartners
Avanade
Cognizant
Collibra
IFS
Pegasystems Inc.
The Ogilvy Group

Cemex (Mexico)
Cencora
Cochlear Limited (Australia)
Commonwealth Superannuation Corp. (Australia)
Cuscal Limited (Australia)
CVS Health
Dawn Foods
DBS Bank Ltd. (Singapore)

Nomura Research Institute, Ltd. Systems Consulting Division (Japan)
Novo Nordisk A/S (Denmark)
OCP Group
Pacific Life Insurance Company
Posten Bring AS (Norway)
Principal Life Insurance Company
QBE

Uniting (Australia)
USAA
Webster Bank, N.A.
Westpac Banking Corporation (Australia)
WestRock Company
Wolters Kluwer
Xenco Medical
Zoetis Services LLC

CISR SPONSORS

Alcon Vision
Amcor
ANZ Banking Group (Australia)
AustralianSuper
Banco Bradesco S.A. (Brazil)
Banco do Brasil S.A.
Bank of Queensland (Australia)
Barclays (UK)
BlueScope Steel (Australia)
BNP Paribas (France)
Bupa (Australia)
CarMax
Caterpillar, Inc.

Doosan Corporation (Korea)
Fidelity Investments
Fomento Economico Mexicano, S.A.B., de C.V.
Fortum (Finland)
Genentech
Gilbane Building Company
Johnson & Johnson (J&J)
Kaiser Permanente
King & Wood Mallesons (Australia)
Koç Holding (Turkey)
Mercer
Nasdaq, Inc.
NN Insurance Eurasia NV
Nomura Holdings, Inc. (Japan)

Ramsay Health Care (Australia)
Raytheon Technologies
Reserve Bank of Australia
Scentre Group Limited (Australia)
Schneider Electric Industries SAS (France)
Stockland (Australia)
TabCorp Holdings (Australia)
Telstra Limited (Australia)
Terumo Corporation (Japan)
Tetra Pak (Sweden)
Truist Financial Corporation
UniSuper Management Pty Ltd (Australia)

CISR ASSOCIATE MEMBERS

MIT CISR wishes to thank all of our associate members for their support and contributions.

MIT CISR is funded by our members, and we gratefully acknowledge their financial support and their many contributions to our work.

Membership and benefits: cisl.mit.edu/content/member-benefits

MIT CISR research publications: cisl.mit.edu/research-library



MIT Sloan School of Management
Center for Information Systems Research

245 First Street, E94-15th Floor
Cambridge, MA 02142

t 617-253-2348 | e cisl@mit.edu

cisl.mit.edu |  

Team | Isobela Byerly-Chapman, Margherita Di Pinto, Christine G. Foglia *Associate Director*, Dorothea Gray-Papastathis, Cheryl A. Miller, Ina M. Sebastian, Alan Thorogood, Nick van der Meulen, Austin Van Groningen *Engagement Director*, Peter Weill *Chairman*, Barbara H. Wixom, Stephanie L. Woerner *Director*