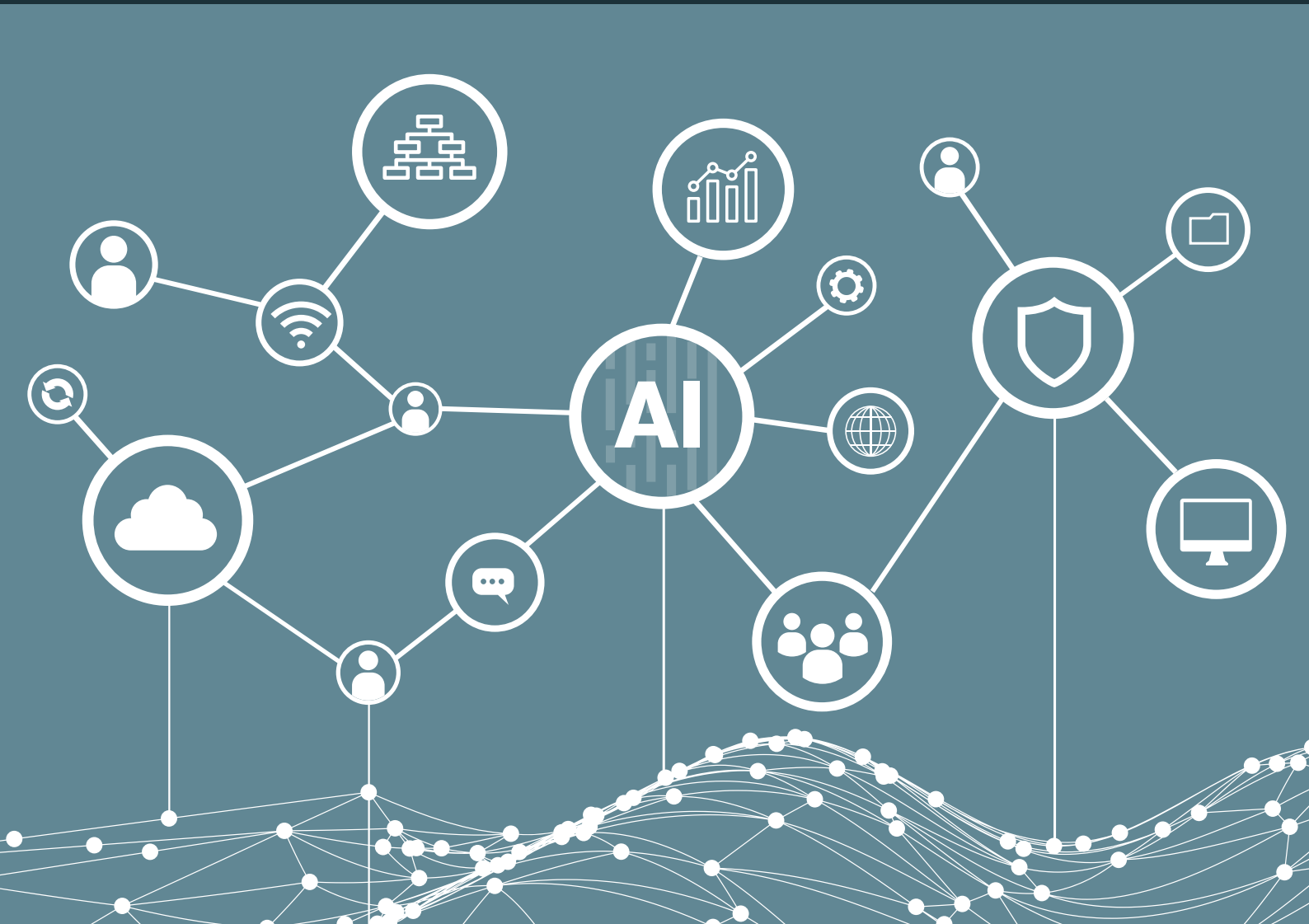
MIT Technology Review Insights

Produced in partnership with

databricks

How data and analytics leaders are delivering business results with cloud data and AI platforms

Building a high-performance data and AI organization



Preface

“Building a high-performance data and AI organization” is an MIT Technology Review Insights report sponsored by Databricks. To produce this report, MIT Technology Review Insights conducted a global survey of 351 chief data officers, chief analytics officers, chief information officers, and other senior technology executives. The respondents are evenly distributed among North America, Europe, and Asia-Pacific. There are 14 sectors represented in the sample and all respondents work in organizations earning $1 billion or more in annual revenue. The research also included a series of interviews with executives who have responsibility for their organizations’ data management, analytics, and related infrastructure. Denis McCauley was the author of the report, Francesca Fanshawe was the editor, and Nicola Crepaldi was the producer. The research is editorially independent, and the views expressed are those of MIT Technology Review Insights.

We would like to thank the following individuals for providing their time and insights:

Patrick Baginski, Senior Director Data Science, McDonald’s (United States)

Bob Darin, Chief Data Officer, CVS Health, and Chief Analytics Officer, CVS Pharmacy (United States)

Naveen Jayaraman, Vice President – Data, CRM & Analytics, L’Oréal (United States)

Michel Lutz, Group Chief Data Officer, Total (France)

Mainak Mazumdar, Chief Data and Research Officer, Nielsen (United States)

Andy McQuarrie, Chief Technology Officer, Hivery (Australia)

Sol Rashidi, Chief Analytics Officer, The Estée Lauder Companies (United States)

Ashwin Sinha, Chief Data and Analytics Officer, Macquarie Bank (Australia)

Don Vu, Chief Data Officer, Northwestern Mutual (United States)

CONTENTS

01 Executive summary

CxOs and boards recognize that their organization's ability to generate actionable insights from data, often in real-time, is of the highest strategic importance. If there were any doubts on this score, consumers' accelerated flight to digital in this past crisis year have dispelled them. To help them become data driven, companies are deploying increasingly advanced cloud-based technologies, including analytics tools with machine learning (ML) capabilities. What these tools deliver, however, will be of limited value without abundant, high-quality, and easily accessible data.

In this context, effective data management is one of the foundations of a data-driven organization. But managing data in an enterprise is highly complex. As new data technologies come on stream, the burden of legacy systems and data silos grows, unless they can be integrated or ring-fenced. Fragmentation of architecture is a headache for many a chief data officer (CDO), due not just to silos but also to the variety of on-premise and cloud-based tools many organizations use. Along with poor data quality, these issues combine to deprive organizations' data platforms—and the machine learning and analytics models they support—of the speed and scale needed to deliver the desired business results.

To understand how data management and the technologies it relies on are evolving amid such challenges, MIT Technology Review Insights surveyed 351 CDOs, chief analytics officers (CAOs; we refer to these and CDOs as "data leaders" at various points in the report) as well as chief information officers (CIOs), chief technology officers (CTOs), and other senior technology leaders. We also conducted in-depth interviews with several other senior technology leaders. Following are the key findings of this research:

- **Just 13% of organizations excel at delivering on their data strategy**. This select group of "high-achievers" deliver measurable business results across the enterprise. They are succeeding thanks to their attention to the foundations of sound data management and architecture, which enable them to "democratize" data and derive value from machine learning. The foundations ensure reduced data duplication, easy access to relevant data, the ability to process large amounts of data at high speeds, and improved data quality. The high-achievers are also advanced cloud adopters, with 74% running half or more of their data services or infrastructure in a cloud environment.