

**2024 GUIDE**

# **PREPARING YOUR DATA FOR GENERATIVE AI**

**FUTUREPROOF YOUR  
DATA MANAGEMENT**



# TABLE OF CONTENTS

## INTRODUCTION

pg 3

AI requires robust data management to be effective and the studies are in. Your data quality can set the stage for adoption of generative AI.

## NEXT STEPS

pg 4

- 1** *CONSOLIDATE YOUR DATA*
- 2** *BUILD A SEMANTIC LAYER*
- 3** *FOSTER DATA LITERACY*
- 4** *CLEARLY DEFINE METRICS*

## CHALLENGES

pg 9

Integrating disparate data sources into a coherent, AI-ready model will require sophisticated data engineering skills.

## SOLUTIONS

pg 10

Blue Margin's managed data service (MDS) handles the challenges of modern data management with a scalable, subscription-based fractional data team. The blend of technology and broad expertise allows your organization to make data-driven decisions faster, stay agile as needs change, and be prepared for upcoming generative AI analytics tools.

## RESOURCES

pg 11

# INTRODUCTION

In today's rapidly evolving business landscape, robust data management is increasingly critical for successful AI integration. Salesforce's "State of Data and Analytics" report underscores the untapped potential of data and its pivotal role in AI-driven decision-making, stressing that AI's benefits hinge on a well-engineered data foundation.

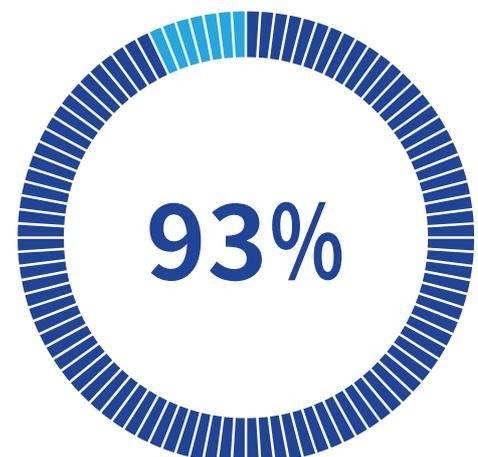
AWS's "CDO Agenda 2024: Navigating Data and Generative AI Frontiers" report echoes these insights, spotlighting generative AI's impending impact and the necessity of precise use cases and high-quality data for maximizing AI's capabilities. It affirms that an effective data strategy is crucial for generative AI success.

Wavestone's "2024 Data and AI Leadership Executive Survey" also reinforces this perspective, with a significant 87.9% of executives ranking data and analytics investments as a top priority, and 62.3% viewing generative AI investments as vital. Yet, only 15.9% believe the industry has sufficiently addressed data and AI ethics, as key aspects of responsible AI usage.

Collectively, these studies present a compelling narrative: for businesses to harness the power of AI, getting their data in order is not just an option, it's a prerequisite. The journey towards AI maturity starts with a solid data foundation, strategic data governance, and a culture that embraces data-driven decision-making. In this era where AI is boosting our abilities, it is vital that data strategy aligns with business goals, not just for gaining a competitive edge, but as a basic business requirement.

**"93% of CDOs believe that a practical data strategy is essential for organizations to succeed with generative AI."**

*CDO Agenda  
AWS 2024*



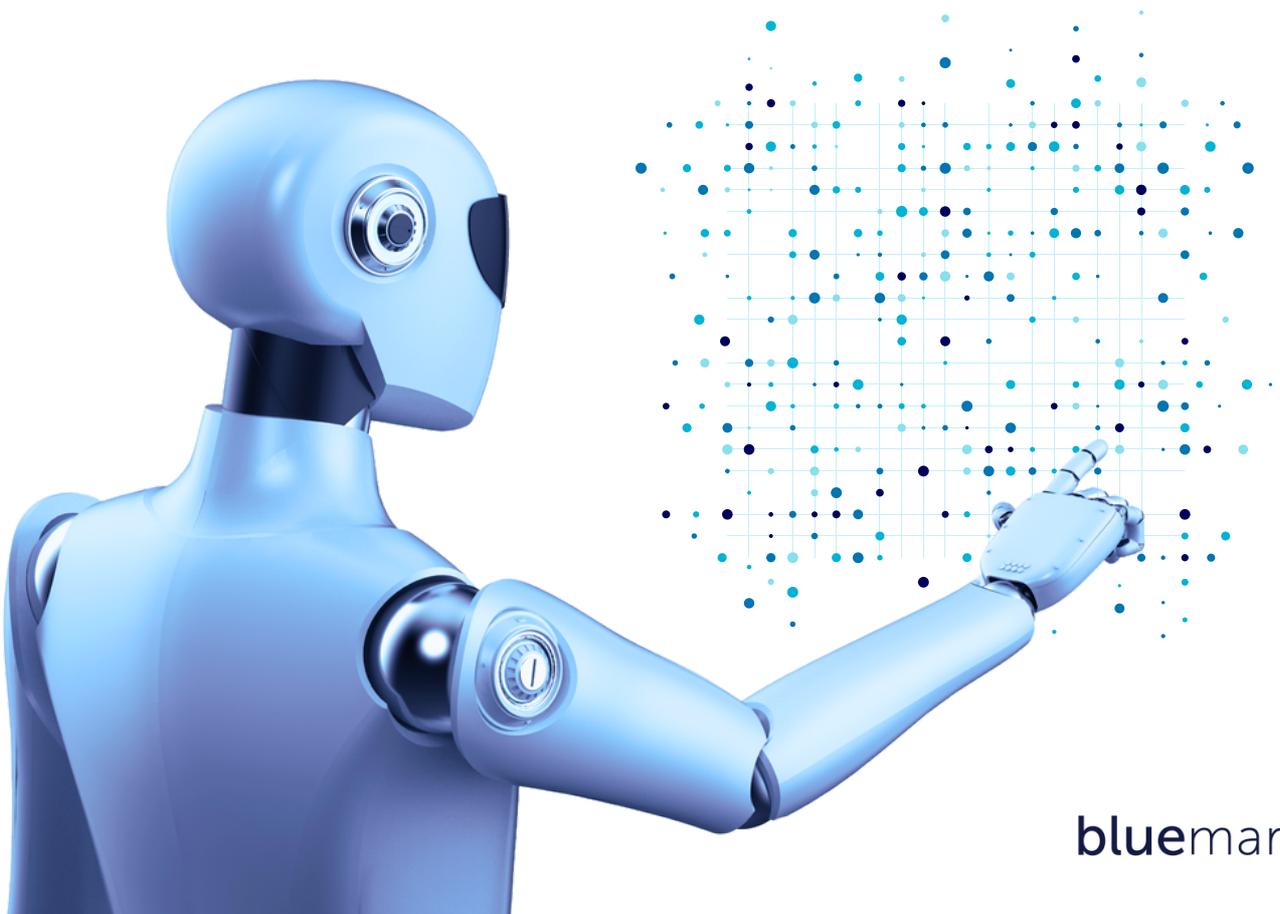
## NEXT STEPS

Businesses are starting to see how important data is for making the most of artificial intelligence.

Now, the focus is shifting from just knowing why data matters to figuring out how to use it effectively. It's understood that organized, quality data is key for using AI well, but knowing exactly how to get there isn't as clear.

The next step involves getting into the details of making and carrying out a data strategy. This isn't just about getting ready for AI; it's about using AI to change how businesses work, make better decisions, and offer amazing experiences to customers.

We're here to guide businesses through the tricky parts of managing data in the AI era, helping them move from just understanding the theory to taking actions that really make a difference.



Getting all your data together in a central location is the first step for using AI effectively. In many businesses, data is scattered across different places (think multiple business tools, Excel files, ERP systems, and other databases), and this scattered setup makes it hard for AI to do a good job analyzing and finding useful insights.

When you bring all your data together, AI can access what it needs to get the full picture of how your business works and surface opportunities for improvement. Plus, having all your data in one place means less duplication, better data quality, and easier management for humans, too.

Data consolidation is also crucial for getting valuable AI results. AI and machine learning get better as they deal with larger quantities of data. The bigger the data set, the better they can spot complex patterns and learn the subtleties.

If your data isn't consolidated in one place, it might hold back AI's ability to fully analyze data and could lead to models that aren't as accurate as they might be. In short, bringing your data together is imperative to making sure you can fully benefit from what AI has to offer.

**“82.2% of CDO/CDAO and data/AI leaders stated that their organizations were increasing their investment in data and analytics”**

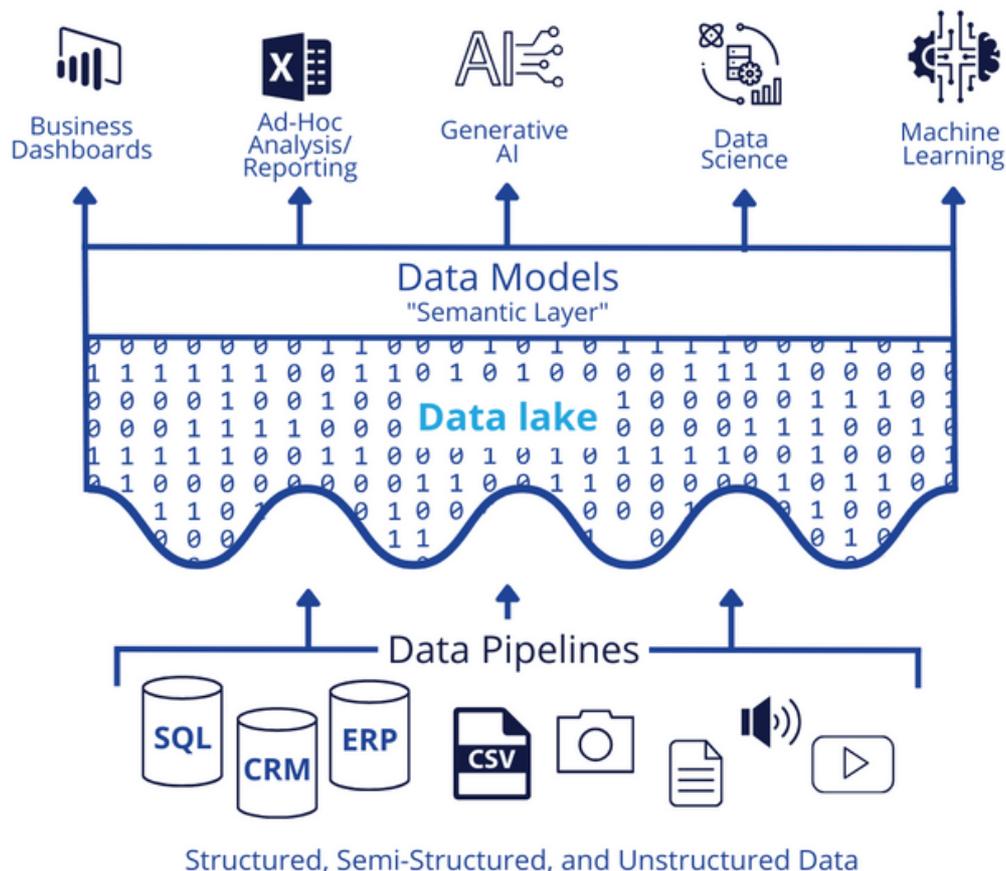
*Data and AI Leadership Executive Survey  
Wavestone 2024*



Beyond making complete and integrated sets of data accessible to AI tools, it must also be organized - which is where the semantic layer comes in.

Think of the semantic layer as a middleman between the raw data in various data stores (e.g., data lakehouse, etc.) and the users or AI that need to make sense of this data. This “layer” provides a clear and consistent way to see data definitions and relationships, helping AI tools and human users to better understand and use the data. It's like providing a map to make sure the human and AI users are able to query the data the way they want to.

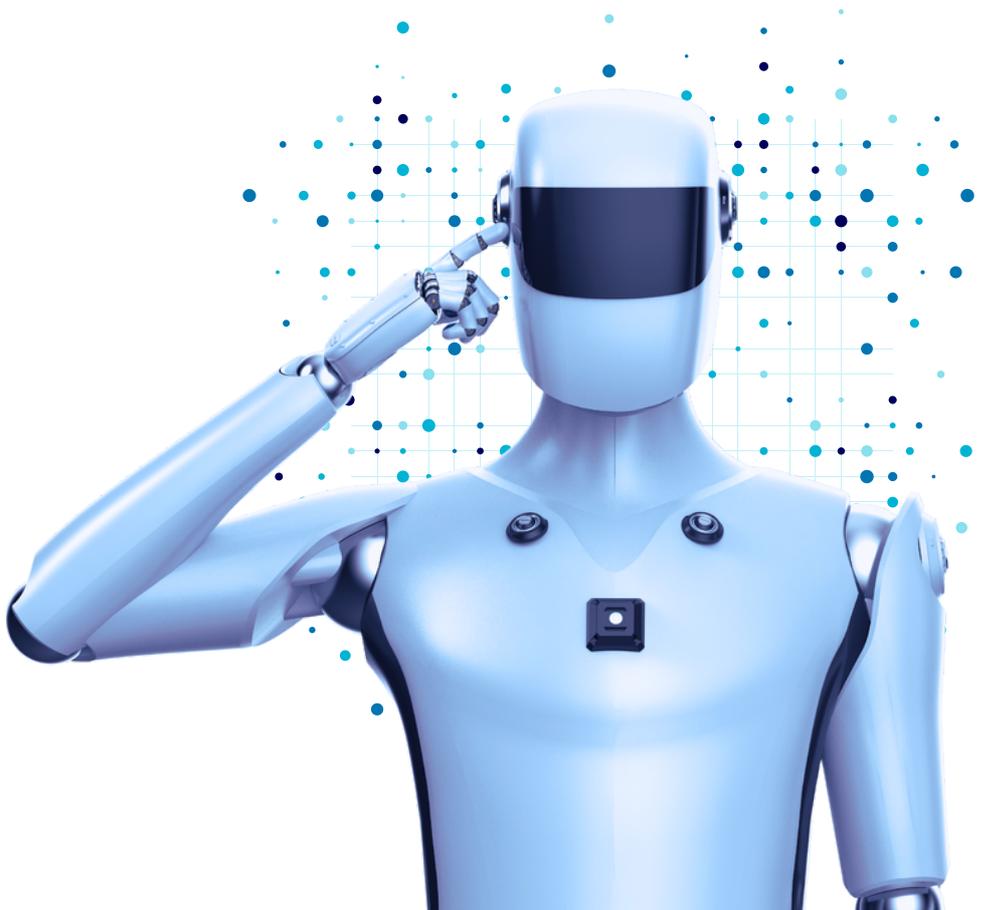
The semantic layer is key to making AI models accurate and user-friendly, even for people who aren't tech experts. It simplifies how users ask for and interact with data, which is crucial when you're putting AI into applications that people will use. Additionally, it helps with management of data governance and security by determining how access and use of sensitive data.



In order to harness and oversee AI-generated insights, it will be crucial to increase data literacy across an organization in order to build a data-savvy workforce that can work effectively with AI.

Key steps for boosting data literacy include:

- **Assess Competency and Create Learning Materials Around Data Literacy:** Start by identifying the current levels of data literacy within your team and develop targeted learning programs/materials to fill the gaps. Bring in outside expertise to help, when needed.
- **Use Diverse Learning Formats:** Adopt a variety of educational methods such as short workshops, online tutorials, and practical projects to cater to different learning preferences.
- **Equip:** Encourage a culture where exploring data and learning from both successes and failures is valued. Provide tools and resources for data analysis and interpretation to all knowledge workers.
- **Fully Integrate Data into Decision-Making:** Require data-driven insights to at least contribute to strategic planning and decision-making at all levels of the organization. This reinforces the importance of data literacy in achieving business objectives.



If you plan to leverage AI analytics tools, having clear definitions of your key performance metrics will be critical. Well-defined (and documented) metrics will produce more trustworthy results from AI, and end users will be able to verify that the outputs they're getting from AI tools are what they requested.

Key metric definition involves the following steps:

- **Identify Strategic Objectives:** Begin with a clear understanding of your business objectives. Identify the metrics that will indicate an objective is being met. Also identify metrics that provide leading indicators toward the objective.
- **Involve Stakeholders:** Engage with stakeholders from across the company. A collaborative approach ensures that the selected metrics are comprehensive and also that there is agreement on how they are defined.
- **Simplify and Prioritize:** Resist the temptation to track too many metrics. Edit down to a concise set of key performance indicators (KPIs) that truly reflect the health and progress of your initiatives. Simplicity aids focus and actionability.
- **Select Actionable KPIs:** Choose metrics that not only provide insights but also drive action. Each metric should intuitively point to clear steps that can be taken to improve performance or capitalize on opportunities.
- **Regular Review and Adaptation:** Objectives evolve; your metrics should too. Regularly review your KPIs to ensure they remain aligned with your strategic goals and make adjustments as necessary.
- **Educate and Communicate:** Ensure that everyone in the organization understands what each metric means, how it's defined and why it's important. This will help foster a data-driven culture and align efforts across teams.

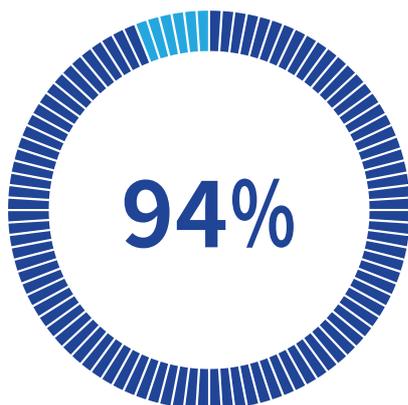
## CHALLENGES

Preparing data for AI implementation is a challenging and complex undertaking. One of the primary hurdles is ensuring data quality and integrity. Data that originates from varied sources can lead to inconsistencies, duplicates, and errors. Integrating disparate data sources into a coherent, AI-ready model will require sophisticated data engineering skills. This often involves data quality initiatives and sometimes updating system use and processes. Companies must also navigate the intricacies of data privacy and security needs, ensuring that AI implementations comply with company and regulatory standards.

Each of these challenges underscores the importance of having an experienced data team, but staffing even a small internal team of three or four people can cost upwards of \$500,000 a year. Add to that the following challenges:

- **Talent acquisition/retention:** The current competitive market for data professionals can cause delay, compromise, and disruption to your projects.
- **Limited knowledge pool:** Teams are confined to their knowledge and experience, which can limit innovation and slows troubleshooting.
- **Capacity limitations:** High workload periods and competing initiatives may overload the team, causing delays and inefficiencies.

Companies looking to increase their data capabilities and keep pace with rapid advances without increasing employee overhead can right-size their approach with a managed data service. ([See our pricing.](#))



**“94% of business leaders feel their organization should be getting more value out of it’s data.”**

*State of Data and Analytics  
Salesforce 2023*

## WHY BLUE MARGIN

With data and AI poised to redefine the midmarket landscape, the time to act is now. Leveraging data well demands a strategic approach, underpinned by robust data management practices.

Regardless of whether you use in-house or external talent, the next step to data maturity involves consolidating company data into a data lakehouse (a process that often takes less than 2 weeks) to provide rapid access to centralized, trustworthy data for analysts, report writers, machine learning, data science, and emerging AI tools.

Blue Margin's managed data service (MDS) handles the challenges of modern data management with a scalable, subscription-based fractional data team. The blend of technology and broad expertise allows your organization to make data-driven decisions faster, stay agile as needs change, and be prepared for upcoming generative AI analytics tools.

As we look ahead, the importance of efficient, effective data management will only grow. Emerging technologies are bringing new standards and best practices, and a fractional BI team is well-positioned to help you keep up. By enabling rapid data consolidation, integration, data modeling, reporting, and (soon) natural language query capabilities, MDS is a better way for many companies to manage and access their data.



*[Blue Margin's Managed Data Service] gives me the ability to plan out the year, gather consensus on initiative priorities, and better predict the budget I need to secure.*

Pedro Renteria, Director of FP&A | SimonMed

# LEARN MORE

## PODCAST EPISODES

[How to Prepare Your Data for AI Advances](#)

[Data is Gold in AI's Wild West: Insights from Salesforce's State of Data & Analytics Report](#)

[Data Quality Sets the Stage for Generative AI: AWS's CDO Agenda for 2024](#)

## ARTICLES

[How the Mid-Market is Fast-Tracking Data Integration with Data Lakes](#)

[In the Age of AI, You Need a Data Lake\(house\) in the First 100 Days](#)

[Private Equity Strategy: Is Your Portfolio Ready for Generative AI?](#)

[Managed Data Service: A Turnkey, Fractional Approach to Data Management](#)