eBook

Business Intelligence Meets AI

Self-service analytics that finally work

databricks

tdwi

Contents

FOREWORD

written by tdwi

For years, organizations have been trying to democratize analytics, i.e., to make analytics more accessible to business users. At TDWI, we've seen that self-service analytics — easy-to-use solutions for nontechnical users who do not require coding and do not require IT to set up data access, queries, visualizations and preparation — has been the ***top priority*** for organizations we survey. This had been the top response in our surveys for more than five years, and now in 2025, self-service analytics came in a close second, behind only generative AI.

There are numerous benefits to this democratization via self-service. Democratizing business intelligence (BI) empowers organizations by enabling faster, more informed decision-making. Specifically, with self-service access to data and insights, business users can make real-time decisions without waiting for IT and data analysts, increasing agility and responsiveness. As more employees engage with data in their daily workflows, organizations can foster a strong culture of data-driven decision-making, leading to more strategic and informed actions. Additionally, by shifting routine reporting and analysis to end users, IT and data teams can focus on higher-value initiatives such as data governance, advanced analytics and AI, ultimately driving greater innovation and efficiency.

Yet, while the value of democratization is clear, business use has remained somewhat anemic. In TDWI research, we've seen that business users aren't widely adopting self-service analytics tools, which have yet to reach mainstream adoption. This may explain why organizations are not that happy with their ability to derive insights from their data. For instance, in 2024 and 2025 TDWI surveys, only about half of respondents were satisfied with their access to data and over 40% remain either dissatisfied or on the fence in terms of their organization's ability to derive insights from data. If the business users who are closest to decision-making can't use company data to drive decisions easily, that ultimately hurts the business.

To make BI tools easier to use, vendors have been working for years to infuse AI into their platforms, to perform tasks such as surfacing insights or suggesting queries. AI can recommend the most relevant charts, graphs and dashboards based on user queries and dataset structures. It can help with automated forecasting. Vendors have been providing early conversational interfaces that use natural language processing techniques to make it easier for business users to ask questions of their data. TDWI has referred to these systems as AI-infused BI. However, it is the advent of generative AI — a subset of artificial intelligence that includes systems designed to generate images, music, text and many other forms of media based on its training data — that has vendors and organizations alike excited about its potential to transform BI and truly democratize it.

The idea is very appealing. Business users can ask their data questions in a natural way, essentially having a conversation with the data through a prompt-like interface that many are now using or experimenting with. In fact, in a 2024 TDWI survey that asked how organizations are planning to use generative AI, close to 25% said "to act as a front end for analyzing company data." What we've seen at TDWI is something we're calling phase 1 generative AI analyses. This is where companies are taking their unstructured data (think call center notes, trouble tickets) and using generative AI to summarize or classify this data. For instance, generative AI might classify the kinds of problems that customers are complaining about and the classes of troubles they are facing. This can be a very useful kind of analysis.

Some organizations are taking analysis and turning it into an application (one that continually analyzes issues, for instance) by operationalizing it using frameworks such as retrieval augmented intelligence (RAG). RAG can provide data in context to foundation models such as large language models without the need for extensive retraining, which can be compute-intensive. RAG combines specific information with prompts to enhance the relevance of the model's output. It can be fairly simple to build an application using retrieval augmented generation if there is just one data source. We heard a lot about RAG in 2024 in the context of building AI applications such as chatbots. Analytics against traditional structured data in a warehouse or lakehouse, however, is more complex.

For democratization, organizations want to use generative AI to analyze traditional company data, including the structured data stored in various platforms across the enterprise that until now has been the domain of BI. This is the data users are familiar with for answering their business questions around sales and revenue, customer behavior and churn, operational efficiency, financial performance, and marketing effectiveness. Business users want to be able to naturally ask questions and get answers based on this data without having to learn a tool or be proficient in SQL. In some cases, business users are using tools provided by their CRM or ERP vendors to do pieces of this. However, democratizing analytics using the data stored in their data warehouse or data lakehouse is more complex. You can't just slap a generative AI front end onto your data warehouse and start asking questions.

Consider an example where a company is concerned about bottlenecks in its supply chain. Its data warehouse may contain information about production times, inventory levels and supplier performance. However, the foundation model it is using to help answer questions is a general foundation model; it does not contain specific information about how this company defines supplier performance, what specific inventory it has and so on. In other words, the generative AI foundation model does not understand the meaning and context of the company's data and how that particular business operates. The organization may think to provide that data in context through a RAG system. However, that alone may not solve the problem; there are too many variations in the kinds of questions a company can ask of its data. It also opens the company to issues, including hallucinations, where the system will provide incorrect answers.

What does your company need to do? In this eBook, you'll learn about **data intelligence** — an important concept where AI has learned the details of your company's data through a comprehensive understanding of both metadata and system knowledge. Data intelligence also uses rules about how your company handles this data. This provides context, structure and data definitions for a generative AI system to use for analysis — it provides the unique aspects of your company's data. It also enables the system to ask questions when it isn't sure of an answer to get feedback from the user and continue to learn.

You'll also hear more about **AI agents** and what Databricks calls compound AI — building on the concept of agentic AI. Agentic AI refers to AI systems that are designed to act autonomously and proactively toward specific goals and objectives, often without needing continuous direct human oversight. Agentic AI's potential to revolutionize industries by automating processes and optimizing outcomes has sparked significant interest and investment. This eBook describes how specialized agents can be used to handle distinct tasks in the analytics lifecycle, including data retrieval, data analysis, results validation and results visualization to further support democratization.

This eBook will guide you through the key concepts, strategies and technologies that can make AI-driven BI a reality. It provides excellent, easy-to-understand, real-world examples of how AI can be infused into BI to make self-service more available to everyone and how the Databricks Platform supports this. I learned some new things about the next generation of analytics from this book, and trust that you will too.

Fern Halper, Ph.D., TDWI VP, Research, and Sr. Research Director for Advanced Analytics
February 19, 2025

CHAPTER 1

Redefining Self-Service Analytics: A New Era in Business Intelligence

For the past two decades, self-service analytics has promised to make data insights accessible to everyone — even those without technical expertise. But why hasn't this vision been fully realized? How would your business change if anyone within it could ask their data questions and get meaningful insights instantly without having to wait on a data team?

> "Through 2026, knowledge workers will still spend the majority of their time in analytics processes searching for and accessing the correct data to include in their analyses, resulting in lost productivity."
>
> — David Menninger, Executive Director, Software Research, ISG Software Research

Remember logging in to an analytics platform and being faced with a wall of dashboards? That's the reality for most business teams today. The scenario is that marketing teams are overwhelmed by hundreds of dashboards and still can't find the specific insight they need, like campaign performance by region.

Take Sarah's marketing team at a midsize retailer. They needed to understand which Q3 campaigns drove the most revenue in their Western region. Simple question, right?

Not quite. Sarah had to:

1. Search through 47 different dashboards
2. Export data from three separate reports
3. Merge everything in Excel
4. Build pivot tables to analyze performance
5. Double-check her calculations

Total time spent? Four hours. And she still wasn't fully confident in her numbers.

This scenario plays out thousands of times daily across organizations. Business users have no lack of questions. But they can't get straight answers from their data. Instead, they face a familiar choice: either slowly navigate the complex dashboard jungle or wait weeks for their data team to build something that works.

But what if Sarah could simply ask: "Which Q3 campaigns had the highest revenue in the Western region?" and get an accurate answer in seconds?

Let's look at another example. A sales director needs to understand "What's the revenue impact of platinum customer churn in Q1?"

The traditional path would be:

- Search existing dashboards
- Request a new dashboard from the business intelligence (BI) team
- Wait for an analyst to become available
- Review dashboard requirements
- Wait for development and testing
- Final delivery two to three weeks late

The new approach would follow this kind of give-and-take:

- **User:** "What's the revenue impact of platinum customer churn in Q1?"
- **AI:** "How do you define platinum customers?"
- **User:** "Customers spending over $50K annually."
- **AI:** "Q1 platinum customer churn resulted in $2.7M lost revenue, a 12% increase from Q4."

Time to insight? Thirty seconds.

This isn't about replacing existing BI tools. It's about making them work better for knowledge workers who need quick answers to drive decisions. When people can ask questions naturally and get trusted answers quickly, they make better choices because they have timely access to the data they need.

The three challenges of traditional BI

Ask any business user about their dashboards and you'll hear a common lament: "I know the data is in there somewhere, but I can't find it." Even if their data detective skills allow them to locate the necessary info, they're not always sure if they can trust it. How would they even know if it can be trusted?

Let's break down why this keeps happening:

1. THE RIGIDITY PROBLEM

A marketing VP spots a drop in customer engagement. The dashboard shows *what* happened but not *why*. Each answer leads to three new questions: Was it a specific region? A particular customer segment? A pricing change?

The dashboard can't adapt to this natural flow of questions. Users get stuck, forced to request new reports or export data to Excel.

2. THE EXPERT BOTTLENECK

A sales manager notices an unusual trend in their territory. To investigate, they need:

- A new dashboard from the BI team
- Custom filters for their specific region by rep
- Comparative analysis against other territories
- Integration with CRM data

Timeline? Two to three weeks — if they're lucky. By then, the opportunity might be gone.

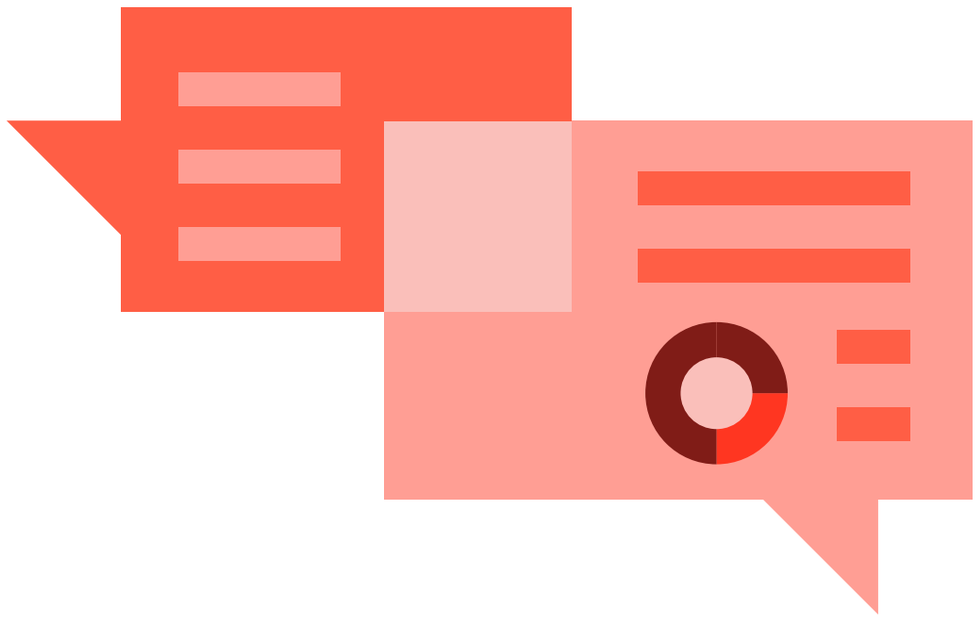
3. DASHBOARD OVERLOAD

Picture a company that has 2,000+ dashboards. Because of “unique requirements,” each department creates its own versions. Finance has one view of customer revenue, Sales another and Marketing a third. Which one is right? Users spend more time searching for data and arguing over the output in meetings than analyzing it.

Think about your own experience. When was the last time you:

- Found exactly what you needed in an existing dashboard?
- Got an answer without exporting it to Excel?
- Didn’t have to wait for the data team?

Here’s the real question: What if your BI tools could understand and answer questions as a data analyst would? Not just show pre-built charts but actually help explore data naturally and follow your train of thought wherever it led?



CHAPTER 2

The Evolution of BI: From Dashboards to Discovery

The dream was simple. Give everyone the power to find answers in their data. For over 20 years, BI innovators tried to deliver on this promise. Let's see what worked, what didn't and why we're still not quite there.

EARLY 2000s: THE DASHBOARD ERA

Remember Cognos and BusinessObjects? These tools brought us our first interactive dashboards. Users could filter data and drill into basic details — a big step up from static PDF reports. But getting answers still meant submitting a ticket to the BI team and waiting days or weeks for new reports. There was often complex mapping of the underlying data architecture to the semantic layer that needed to happen in the BI system before any work could begin on reports. It was time-consuming and onerous and required expertise.

THE 2010s: THE DISCOVERY WAVE

Then came Qlik and Tableau, which changed how we interacted with data. Users could:

- Click through visualizations
- Filter across multiple charts
- Create their own views

But there was a catch. While the tools were more flexible, they needed experts to:

- Set up data models
- Create initial dashboards
- Maintain data connections
- Train users

2015+: THE SEARCH REVOLUTION

Tools like ThoughtSpot tried something new: search data in ways similar to searching the web. Type “sales by region Q3” and get instant results. This made exploring data and visualizing basic metrics easier, but users hit walls when they needed deeper insights or had to connect data from multiple sources.

Each generation of BI tools made data more accessible. But look closely at the organization of today:

- How many people can actually answer their own data questions?
- How long does it take to get new insights?
- How much does the data team still handle?

The pattern is clear: BI keeps getting better at giving users what they think they need. But it struggles to adapt when they need something different or they come up with good follow-up questions.

Why? Because traditional BI works from a fixed set of assumptions:

- We know which questions users will ask
- We can build dashboards for every scenario
- Users understand how their data is structured

Real business questions don’t work that way. They’re messy and evolving and often need context from multiple sources.

CHAPTER 3

The GenAI Revolution in BI

Generative AI burst onto the scene with a bold promise: just ask data questions in plain language and get instant answers. No more complex queries, no more waiting on dashboards. Sounds perfect, right? Well, it sounds so compelling that every legacy BI vendor suddenly bolted on a GenAI bot to assist the user at every stop on their analysis journey.

Let's look at what GenAI brings to BI:

Natural language queries that work

Before, users would need to specify something like:

```
SELECT SUM(revenue) FROM sales WHERE region = 'West' AND date BETWEEN …
```

Now, they can simply query: "What's our revenue in the Western region this quarter?"

INSTANT INSIGHTS

A sales leader asks: "Why did we miss our Q3 targets?" GenAI can:

- Analyze performance data
- Spot key trends
- Highlight unusual patterns
- Suggest focus areas

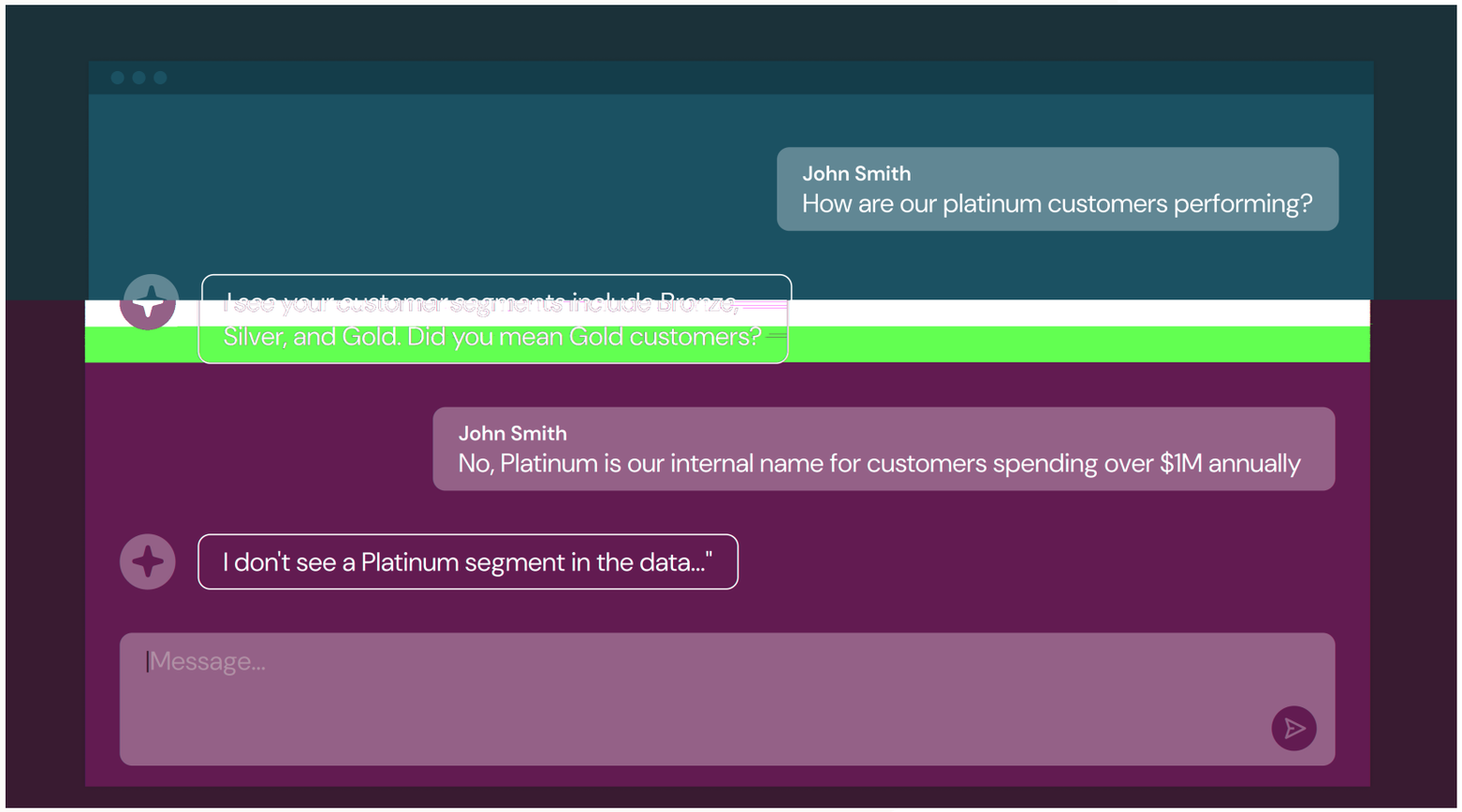
BUSINESS CONTEXT

When a user asks about "churn in enterprise accounts," GenAI understands:

- Their company's definition of enterprise
- Which metrics matter for churn
- Relevant comparison periods
- Industry benchmarks

But here's the catch. Most of today's bolt-on GenAI BI tools are like smart parrots — they're great at rephrasing questions and presenting data, but they lack critical business understanding.

An *example*



This gets to the heart of the problem. GenAI needs more than natural language skills — it needs to understand business context. And bolt-on GenAI does not offer that.

Nearly two-thirds of data decision-makers expect a democratization of access to insights in 2024, with 84% of survey respondents believing GenAI will help their organization access insights faster.[1]

[1]Joey Frazee and Pratyarth Rao. "Introducing New Training Courses for SQL Analytics and BI, With AI-Powered Self-Service Analytics." Databricks, December 4, 2024. https://www.databricks.com/blog/new-training-courses-sql-analytics-and-bi?itm_data=bimeetsaiebook

CHAPTER 4

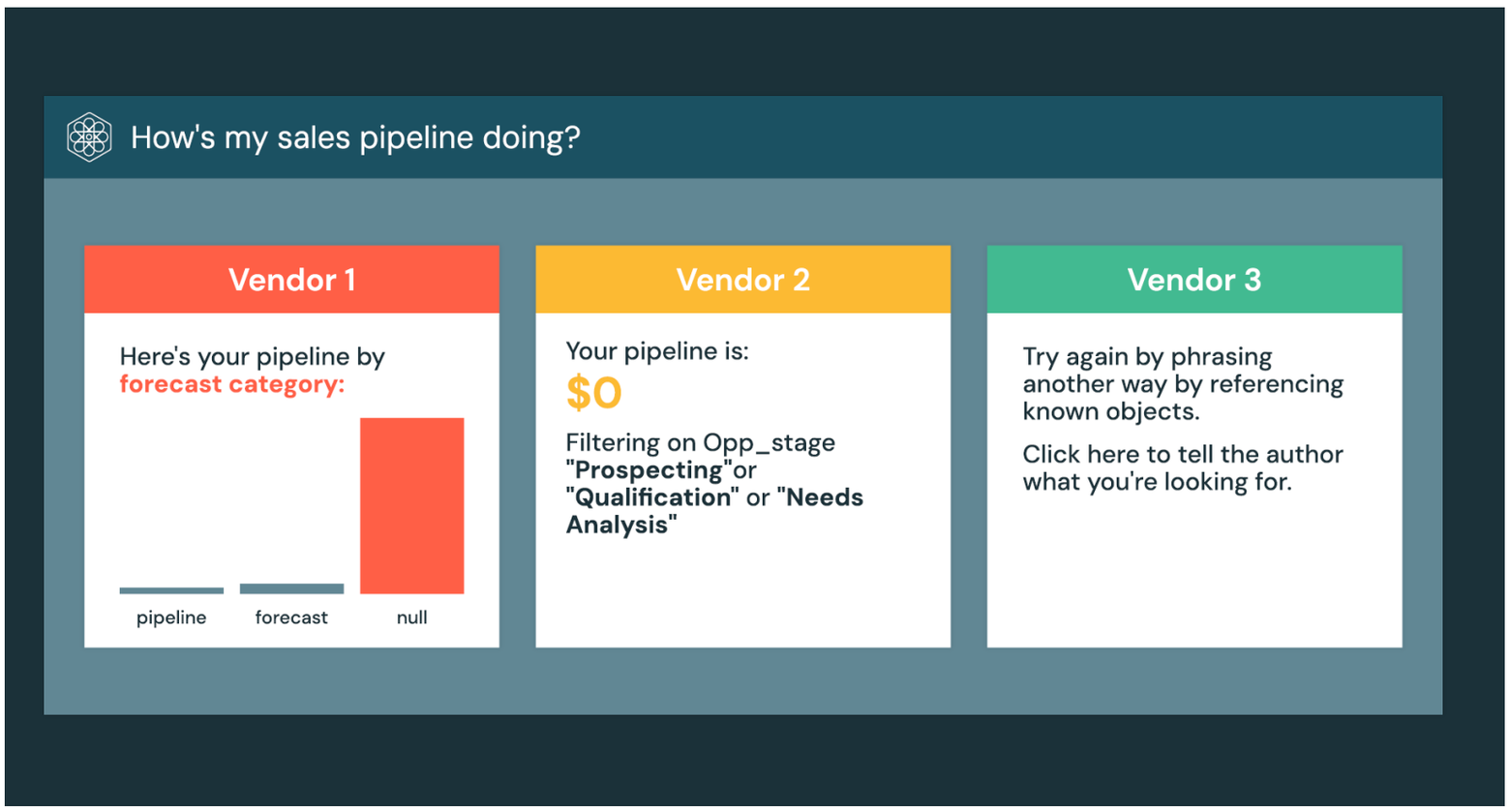
The Limits of Bolt-On GenAI in BI

We recently tested several natural language analytics solutions from leading BI vendors. While these systems showed promise, they also revealed a common flaw: simply bolting on GenAI to existing BI tools fails to address an organization's business context. The following real-world scenario illustrates why organizations need more than just a bolt-on AI upgrade to navigate complex data and deliver accurate insights.

A tale of three responses

User: "How's my pipeline?"

Responses were as follows:



When we tested a realistic sales opportunity dataset, we asked a simple question: **"How's my pipeline?"** Yet, the responses we received from three leading vendor solutions were far from useful:

- **Vendor A** displayed a series of null values. It was technically "truthful," but offered no meaningful insight.
- **Vendor B** insisted there was no pipeline at all, because the sales stages did not match its expected format, leading it to the wrong conclusion
- **Vendor C** claimed it did not understand the term "pipeline" because it was never explicitly defined in the semantic model

None of these tools provided the answer we needed or expected. Instead, they highlighted the inherent challenges of bolt-on GenAI systems, which often lack a deep contextual understanding of specific businesses and their data.

Three core limitations

1. DATA CONFUSION

Real-world data is rarely pristine, which often confuses bolt-on AI systems. They don't understand the nuances your analysts, data scientists and sales ops use to reconcile missing columns or contradictory fields in their notebooks. For example, your CRM may list 1,000 active customers, while the billing dashboard shows 950 because of different status definitions. Meanwhile, support tickets might apply an entirely separate categorization system. Without embedded logic to handle these discrepancies, bolt-on AI either ignores them or guesses incorrectly, leading to inconsistent or misleading results.

2. MISSING BUSINESS CONTEXT

When a user asks, "How's my pipeline?" a bolt-on AI often lacks the specific stage definitions that form your unique sales pipeline definition, which includes "Prospecting," "Qualification" and "Needs Analysis." As a result, it may generate queries that miss entire categories of deals or rely on generic assumptions that don't match how your sales team tracks progress. Lacking this core context, the AI returns incomplete or misleading answers instead of reflecting your true pipeline status.

3. STATIC UNDERSTANDING

Many bolt-on AI solutions rely on fixed semantic models, which can only answer questions that were anticipated and "pre-modeled" in advance. In our example, the tool used by Vendor 3 had no concept of "pipeline" because it wasn't explicitly defined. This rigid approach fails to reflect the fluid nature of real businesses, where new product lines launch, marketing segments evolve and reporting requirements shift faster than the model can be updated. It also isn't realistic to model everything before users ask their questions — especially since they themselves may not know what questions will arise.

Bolt-on AI solutions struggle because they rely on rigid, outdated semantic models and can't adapt to real-world changes. As a result, they fail to account for unique business definitions (like "pipeline" stages), become confused by inconsistent data sources, and cannot keep pace when new product lines or customer definitions emerge. These limitations leave organizations with incomplete or inaccurate answers that undermine trust and hinder decision-making. Without built-in awareness of evolving metrics or the ability to reconcile multiple data silos, bolt-on AI cannot deliver the continuous, context-rich insights modern businesses require.

The cost of getting it wrong

Consider a revenue forecast carried out by bolt-on AI:

- Pulls raw numbers without understanding seasonality
- Misses critical deal stages unique to the sales process
- Overlooks recent changes in how revenue is recognized

The result? Important business decisions based on flawed insights.

The next section shows how data intelligence solves these problems by building understanding from the ground up.

CHAPTER 5

Why Data Intelligence Is Essential for BI Success

Data intelligence is fundamentally different from general AI. While both use artificial intelligence, data intelligence is specialized — it understands an organization's specific data and business concepts.

Think of it as the difference between talking to someone who just joined a company versus a colleague who's worked there for years. While both can have a conversation, the colleague who has been there for years understands the company's unique terminology and context.

Industry innovation

Recent advances in AI and machine learning have enabled new approaches to understanding enterprise data context. Key developments include:

- Semantic layer innovations
- Metadata management evolution
- Natural language processing with business context

When we talk about data intelligence in BI, we're describing AI that's learned the details of an organization's data — from how it's structured to what different terms mean in specific business contexts. This is very different from general AI models, which have broad knowledge but don't understand a company's particular way of defining and measuring things.

For example, if you ask a general AI system about "customer churn," it might return a textbook definition. But data intelligence knows exactly how your company calculates churn, including specific time frames, customer segments and revenue thresholds that matter to your business.

The key difference is that data intelligence is built specifically to work with enterprise data. It understands the data's structure, knows how different pieces connect, and most importantly, recognizes the business rules and definitions unique to an organization.

How data intelligence works

Data intelligence becomes useful through three main approaches: comprehensive knowledge, standardized instructions and continuous learning. Let's break these down in practical terms.

First, think about how data intelligence learns a data platform. It doesn't just look at individual pieces — it studies everything from metadata to how data is used and flows through the system. This gives it a complete picture of a data ecosystem, similar to how an experienced analyst understands not just the numbers but where they come from and how they connect.

The second part involves "gold standard instructions." These are like playbooks that teach the AI specific business rules. For example, when calculating something like net revenue, the AI learns from certified examples that show exactly how a company handles different scenarios — for example, maybe you exclude certain types of returns or treat international sales differently.

Finally, data intelligence works a lot like a careful analyst who isn't afraid to ask questions. When unsure about something, it asks for clarification. As users provide feedback and corrections, it gets better at its job. For instance, if someone points out that a particular calculation should exclude trial users, the AI remembers this for future queries.

HOW DATA INTELLIGENCE LEARNS AND DELIVERS VALUE

Data intelligence operates effectively by mastering three key pillars of knowledge:

1. **Learning data:** It examines data structures, usage, relationships and lineage to understand how information flows through a system.
2. **Learning from expert instructions:** It applies business-approved definitions — your "gold standard" rules — to ensure that every calculation aligns with company-specific standards.
3. **Learning from real-time feedback:** It refines its approach based on user input. If someone clarifies how a metric should be calculated, data intelligence updates its logic — making future analyses more accurate.

DATA INTELLIGENCE IN ACTION

Learning from data

The user asks: "What's the enterprise churn rate this quarter?"

- **Data intelligence knows:** It produces a precise churn rate that reflects how the company defines active and inactive accounts and the appropriate time window.
- **How it knows:** By analyzing metadata, understanding the data lineage and reviewing historical queries, it recognizes which accounts to include, how to aggregate them and what time frame to use — just like a well-informed analyst who understands not just the numbers but their origins and connections.

Learning from expert instructions

The user asks: "What's the revenue impact of platinum customer churn?"

- **Data intelligence knows:** It immediately applies the correct definition of platinum customers (e.g., those spending > $1M annually), understands churn as both cancellations and downgrades, and factors in contract terms to deliver a contextually accurate revenue impact figure.
- **How it knows:** These gold standard rules — certified by an organization's finance and analytics teams — tell the system exactly how to interpret "platinum," what counts as churn and how to calculate revenue impact. It follows these approved guidelines to ensure every insight aligns with a logic unique to the business.

Learning from real-time feedback

The user requests: "Show me monthly recurring revenue (MRR) growth."

- **Data intelligence knows:** Having previously learned from a user correction, it knows to exclude trial users automatically when calculating MRR growth.
- **How it knows:** From real-time feedback. Initially, the system might have included trial users in MRR. But after receiving clarification from users, it adjusted its approach, remembering this refinement for all subsequent analyses — demonstrating how user feedback continually improves the system's accuracy.

INDUSTRY EXAMPLE: GLOBAL GAMING COMPANY

Challenge

SEGA Europe faced difficulties in fully leveraging their extensive data resources, which included processing 50,000 events per second from over 40 million players across more than 100 video games. This challenge hindered their ability to enhance the player experience, improve customer retention, and optimize sales operations and forecasting.

Solution

Implementation of a data intelligence platform that:

- Consolidated data into a unified architecture, integrating data across global databases
- Enabled data democratization, allowing various departments to access data in plain language
- Developed an asset generation model to expedite the video game design process

Results

Achieved up to a 40% increase in player retention through real-time insights and enhanced player sentiment analysis.

https://www.databricks.com/customers/sega?itm_data=bimeetsaiebook

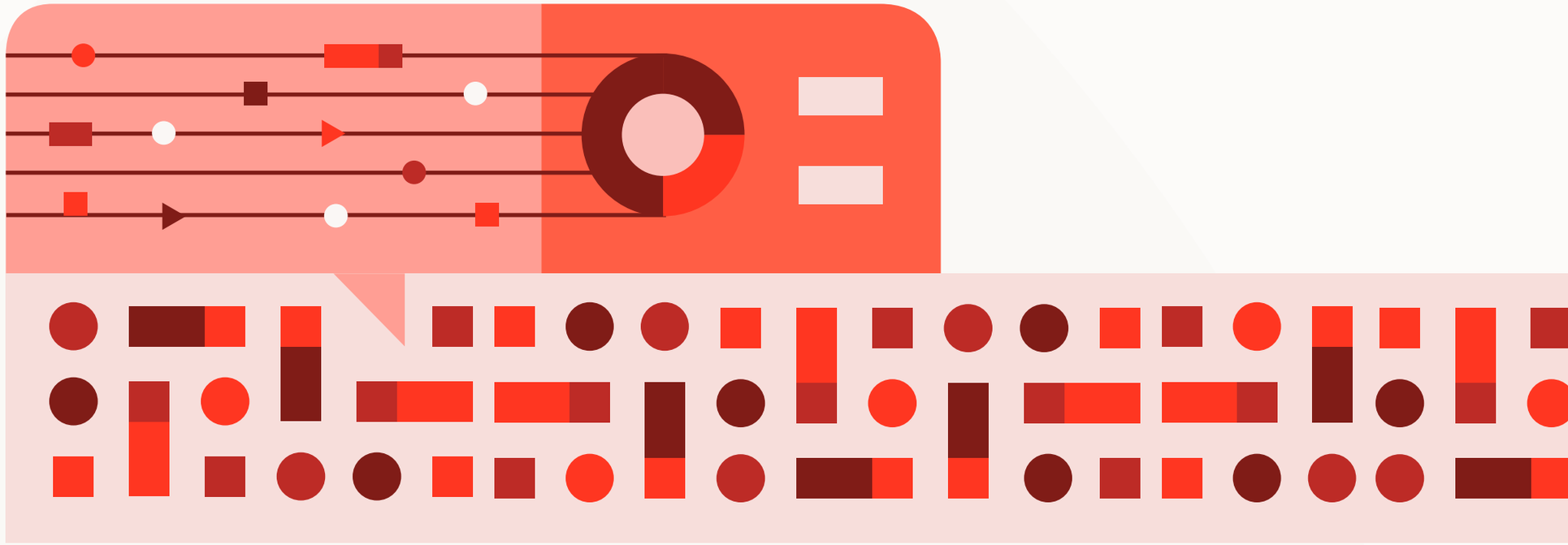
Foundation for compound AI

Data intelligence sets the foundation for building more advanced AI systems that can work effectively with business data. This understanding of an organization's data makes it possible for users to get reliable self-service analytics.

Think of data intelligence as the translator between your business data and AI applications. Because it understands your specific context, definitions and rules, it helps ensure that any AI-driven analysis stays accurate and relevant to business needs.

In the next section, we'll examine how compound AI systems use this foundation and how they deliver real-time insights that adapt based on both user needs and an organization's specific context.

This is why data intelligence isn't just another layer of AI — it's the essential component that makes all other AI capabilities more reliable and useful in a business setting. With this foundation in place, organizations can build more sophisticated AI systems while maintaining accuracy and relevance.



CHAPTER 6

Compound AI: The Key to Unlocking Data Intelligence

Picture a team of expert analysts, each specializing in a different aspect of understanding and using data. The business analyst excels at interpreting business questions and clarifying complex terms. An analytics or data engineer knows precisely which data sources and tables to query. The domain expert applies business rules and verifies outputs against historical patterns. And the reporting specialist is adept at turning raw outputs into clear, actionable insights.

Now imagine these experts working in perfect harmony, collaborating instantly to answer your questions. This is the promise of compound AI — not a single, all-purpose tool but a coordinated team of specialized AI agents working together to solve complex business problems.

Evolution of AI architectures in analytics

- Traditional single-model approaches forced all responsibilities onto one system
- Specialized AI agents now handle distinct tasks
- Modern multi-agent systems coordinate these specialists
- Compound AI uses data intelligence to ensure meaningful business outcomes

Coordinated expertise in action

Consider Sarah again, the marketing director we met earlier. She wants to understand why enterprise churn is rising. In a traditional scenario, even a sophisticated AI might struggle, forcing Sarah to break her query into technical steps. With compound AI, it's different:

- **Sarah:** "What's lost revenue from increased churn in enterprise accounts?"
- **Question interpreter agent:** Looks for a "certified" SQL example but can't find a relevant one, leading it to ask Sarah for clarification: "Can you clarify churn? I don't think I have your specific definition."

- **Sarah:** "Churn means anyone who generated more than $1,000,000 in any given fiscal quarter but not in the next."
- **Question interpreter agent:** [Learns rule] and describes SQL logic to data retriever agent.

```sql
WITH MonthlySales AS (
  SELECT
    Customer_FullName,
    EXTRACT (MONTH FROM OrderDate) AS SaleMonth,
    EXTRACT (YEAR FROM OrderDate) AS SaleYear,
    SUM(TotalDue) AS MonthlySales
  FROM
    product_sales_gold
  GROUP BY
    Customer_FullName,
    EXTRACT (MONTH FROM OrderDate),
    EXTRACT (YEAR FROM OrderDate)
  HAVING SUM(TotalDue) AS MonthlySales > $1000000 --NOTE: definition of enterprise account
),
ChurnedCustomers AS (
  SELECT
    a.Customer_FullName
  FROM
    MonthlySales a
    LEFT JOIN MonthlySales b ON
      a.Customer_FullName = b.Customer_FullName
      AND a.SaleMonth = 2
      AND a.SaleYear = 2024
      AND b.SaleMonth = 3
      AND b.SaleYear = 2024
  WHERE
    b.Customer_FullName IS NULL
    AND a.MonthlySales > 0 --NOTE: definition of churned customer
)
SELECT
  Customer_FullName
FROM
  ChurnedCustomers;
```

- **Explanation agent:** Explains to Sarah, "This query identifies enterprise accounts (accounts > $1M revenue in one month) that churned in February 2024 by comparing their sales in February and March 2024. Churned customers are those who had sales in February 2024 but not in March 2024."
- **Results agent:** Returns a table of data to Sarah for churned enterprise accounts.
- **Visualization agent:** Plots a bar chart for the data.

Compound AI refers to AI systems that integrate multiple specialized models, tools and external APIs into a cohesive framework to perform complex tasks. These systems combine foundational AI models with additional capabilities like retrieval augmented generation, database querying or task-specific fine-tuning, which enables them to function as versatile, end-to-end solutions rather than stand-alone models.[2]

Within seconds, Sarah receives a contextually accurate and trustworthy result. The coordinated agents confirm which enterprise accounts meet her new churn definition and present a clear breakdown of lost revenue. Instead of juggling spreadsheets or writing complex queries, she now — at a glance — sees which accounts stopped generating over $1 million per quarter, how that drop impacted overall revenue and how these changes evolved month over month. With this immediate, data-backed clarity, Sarah can move quickly from problem identification to strategy adjustment, confident in the depth and precision of the analysis.

[2]Matei Zaharia, Omar Khattab, Lingjiao Chen, Jared Quincy Davis, Heather Miller, Chris Potts, James Zou, Michael Carbin, Jonathan Frankle, Naveen Rao, and Ali Ghodsi. "The Shift from Models to Compound AI Systems." The Berkeley Artificial Intelligence Research Blog, February 18, 2024. https://bair.berkeley.edu/blog/2024/02/18/compound-ai-systems/.

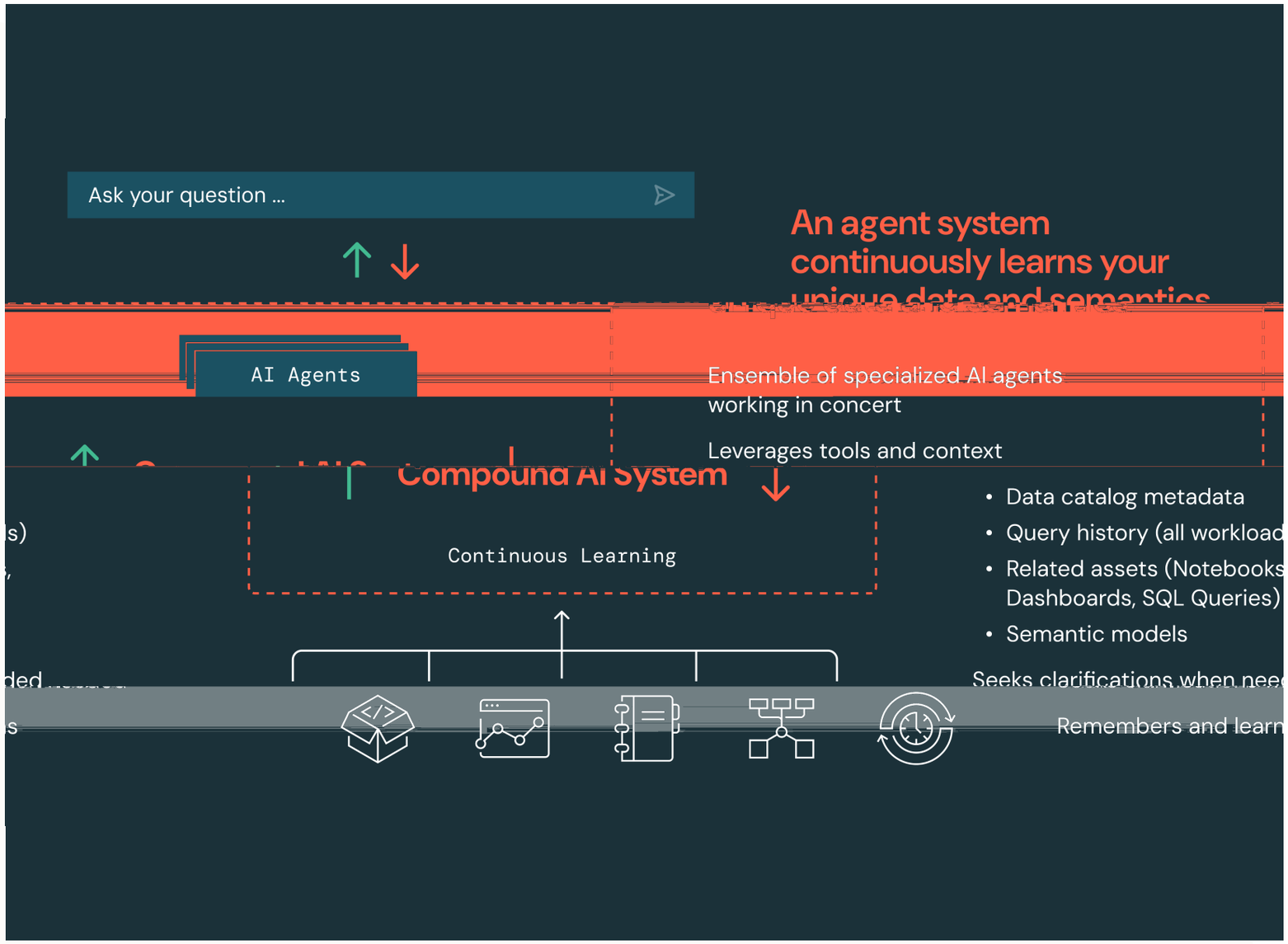


Figure 6.1: Compound AI systems use agents and continually learn

The interplay of agents demonstrates the power of compound AI. Each agent contributed its expertise — writing a SQL query, applying domain rules, validating against norms and explaining its results in business language — and was part of a coordinated system response. The result wasn't just faster; it was more thorough and trustworthy than what a single AI model could produce on its own.

Beyond simple questions

Compound AI truly shines when faced with complex, multi-layered questions. Imagine a business analyst asking: "How have our recent pricing changes affected customer retention across different segments, and what's the estimated revenue impact next quarter?"

In a single-model world, this question might be too broad. With compound AI, the system springs into action:

- **Question interpreter agent:** Breaks the question into parts such as "recent pricing changes," "customer retention," "segments" and "estimated revenue impact." It also clarifies ambiguous terms by asking: "When we say 'recent,' do we mean the last fiscal quarter? And which segments are we focusing on — enterprise, mid-market or SMB?"
- **Data retriever agent:** Runs a series of SQL queries to pull relevant data. One for historical pricing data, another for retention rates and another for segment definitions. It might use a predefined SQL function to return the results for one of these queries.
- **Analysis agent:** Applies correlation and regression techniques, potentially calling an AI prediction model to forecast how changes in retention patterns will influence next quarter's revenue. This could involve a fine-tuned model trained on historical pricing and retention data, now integrated as just another tool in the agent's toolkit.
- **Result validator agent:** Checks whether the model's forecasts align with known seasonal trends and previously observed patterns. If something doesn't match business expectations, it might ask the data retriever agent for a deeper historical view or consult documentation on price changes to ensure the reasoning is sound.
- **Visualization expert agent:** Summarizes the findings. Perhaps it generates a chart showing how price changes correlate with retention dips in the SMB segment while enterprise accounts remain stable. It then translates the statistical findings into a straightforward narrative for the business analyst.

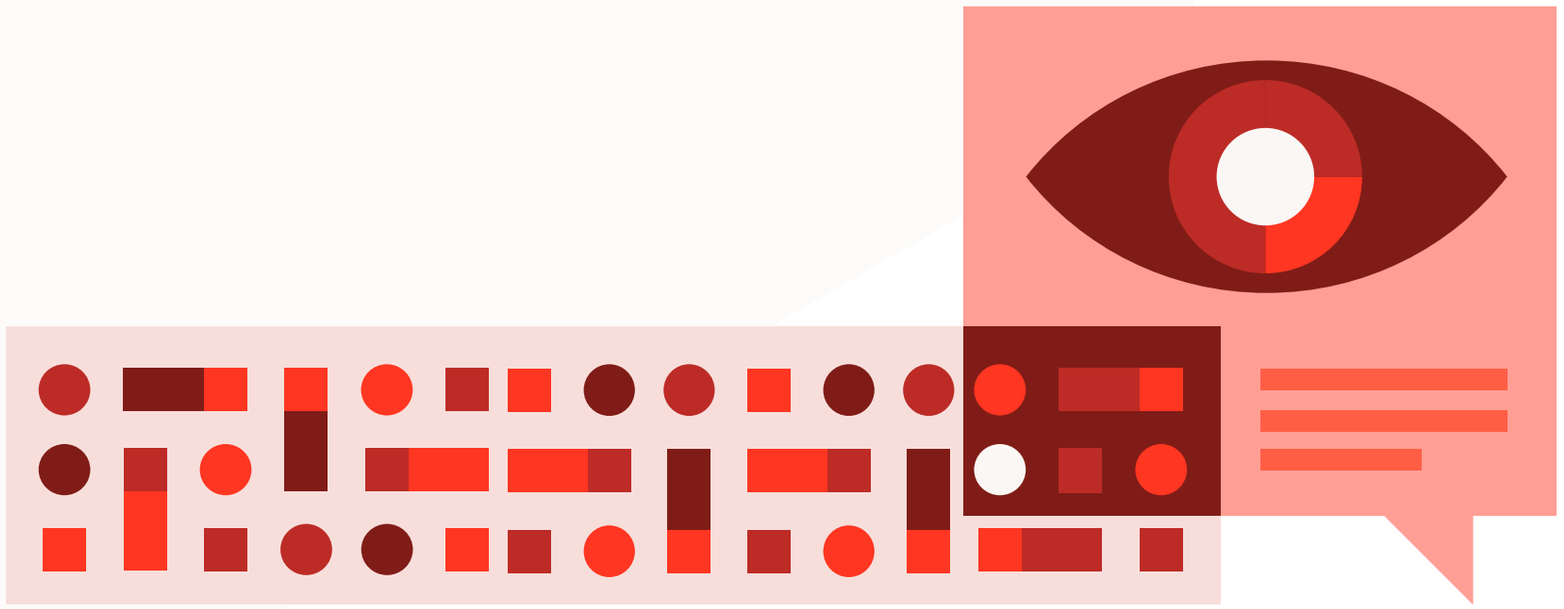
The final answer emerges as a refined, context-aware explanation accompanied by a forecast chart delivered in near real-time. Instead of writing SQL code, digging through spreadsheets or relying on guesswork, the analyst receives an informed, validated and easy-to-digest result.

From data complexity to business clarity

Compound AI transforms how organizations interact with data. Rather than forcing business leaders to think like analysts, compound AI thinks like a business leader. It automatically clarifies vague requests, retrieves and verifies relevant data through SQL queries, taps into predictive AI models, and presents findings in business terms.

In traditional workflows, getting an answer to a complex question could take days. With compound AI, it takes minutes — or even seconds — and the reasoning behind every insight is transparent and auditable. The system's ability to incorporate feedback means each interaction refines not only the data and models but also how agents coordinate and solve problems together.

By harnessing compound AI, organizations unlock a richer, more intuitive way of making decisions. The result is a significant leap forward in agility, accuracy and trust, setting the stage for a new era of data-driven strategy and execution.



CHAPTER 7

Next-Generation Self-Service Analytics in Practice

Let's imagine that the weekend is over and it's time to prep for a Monday morning meeting. Instead of surfing through dozens of spreadsheets, trying to navigate a few dashboards, or trying to find your favorite SQL query, you simply ask, "What's our revenue by store for last week?" Within seconds, you see results organized in a clear, accessible format. You follow up with: "Which stores are running low on inventory?" and instantly get a ranked list.

This back-and-forth feels like having an expert data analyst on call who is ready to clarify trends, highlight issues and guide decision-making.

This conversational approach to analytics is here. New compound AI–driven systems like those powering Databricks AI/BI empower users to explore data with everyday language. Rather than spending hours interpreting dashboards or waiting on specialists, simple iterative questions can be used to dive deeper into the data. The technology works behind the scenes to clarify questions, generate accurate SQL queries, retrieve relevant data and present insights in a format that's easy to understand. We will explore Databricks AI/BI in detail in the next chapter, but first, let's take a closer look at some real-world examples to showcase the potential benefits of these new AI systems.

The shift in self-service analytics

The evolution of self-service analytics marks a fundamental change in how businesses interact with data:

- **From static dashboards to dynamic conversations:** Instead of staring at a fixed report, users have fluid discussions with the system.
- **From technical queries to natural language:** Forget about being a SQL expert. Just ask questions in plain language.
- **From guesswork to guidance:** The system can prompt follow-up questions or offer suggestions, helping users refine their analysis.

These shifts mark a new era of analytics adapting to how business leaders think, rather than forcing leaders to think like analysts.

Real-world examples across departments

To illustrate these capabilities, we'll introduce AI-Bot, our example of an AI-powered assistant, and show how different teams interact with it using natural language. We'll start with simple queries and gradually increase the complexity, demonstrating how AI-Bot adapts to each scenario. This progression will show the breadth of what's possible and how the system learns and responds more intelligently with each interaction.

SALES LEADERSHIP: FAST, TARGETED PIPELINE ANALYSIS

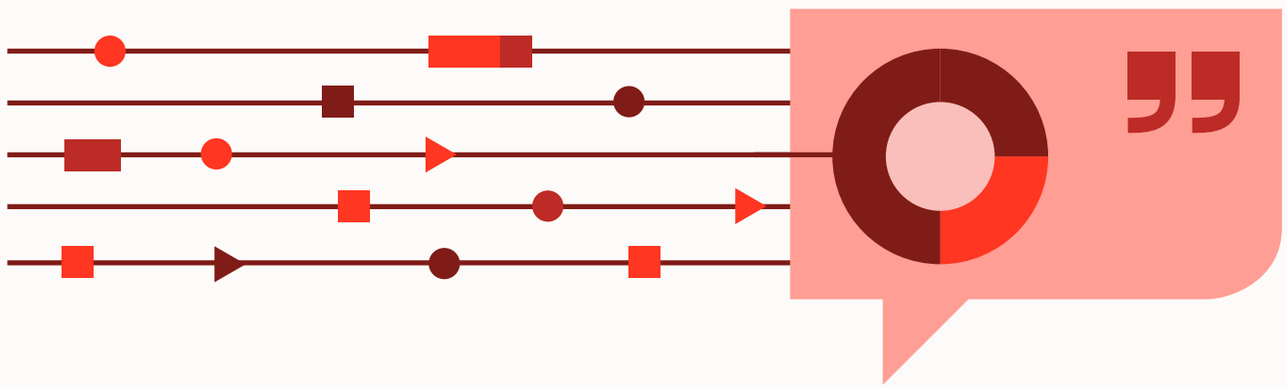
Key benefit: Rapid drill-down on performance metrics

James, a sales director, uses conversational analytics to monitor regional performance. Instead of writing queries or sorting CSV files, he can simply ask: "How did revenue in the Western region change last quarter?"

- **AI-Bot:** "Revenue in the Western region increased by 8% last quarter."

Following this prompt, James responds: "Thanks. Show me revenue by store, and highlight any that fell below their targets."

In seconds, he sees underperforming stores and digs deeper — "Check inventory levels in Store #45" — to identify if stockouts contributed to missed targets. Instead of spending hours preparing data, James uses conversational analytics to swiftly diagnose performance issues and take immediate action



MARKETING CAMPAIGN OPTIMIZATION

Key benefit: Discovering root causes with natural follow-ups

Rachel, a marketing director, wants to understand why Q3 campaigns lagged behind Q2 and asks, "How did our Q3 campaigns perform versus Q2?"

- **AI-Bot:** "Q3 conversions were down 10% compared to Q2. Would you like to see which channels were most affected?"

She follows up with, "Yes, break it down by channel."

AI-Bot identifies email engagement as the weak spot. Another nudge leads Rachel to examine send times, revealing that early morning sends outperform others. With a few conversational turns, she pinpoints an actionable fix — adjusting campaign timing — without juggling multiple tools or manually correlating data.

FINANCE PLANNING AND ANALYSIS

Key benefit: Immediate insight into complex financial drivers

For the CFO's team, understanding margin shifts traditionally involved painstaking manual analysis. Now, they can ask, "What caused the 3% margin drop in the enterprise segment?"

- **AI-Bot:** "Margins declined due to three factors: (1) Raw material costs increased by 12%, (2) three major clients renegotiated contracts at lower rates, and (3) a shift toward lower-margin products reduced overall profitability. Would you like to see the potential impact of a 5% increase in contract renewal rates?"
- **FP&A analyst:** "Yes, show me the impact of a 5% renewal rate increase."
- **AI-Bot:** "A 5% increase in contract renewal rates could improve margins by an estimated $1.2M next quarter."

One follow-up reveals that adjusting renewal strategies could improve margins by $1.2M next quarter. Complex financial inquiries that once took days now resolve in minutes, allowing leaders to confidently refine plans and negotiate with suppliers or clients on the fly.

INDUSTRY EXAMPLE: LARGE RETAIL ENTERPRISE

Challenge

Grupo Casas Bahia struggled with delays in data analysis, taking up to 5–6 hours. This hindered the company's ability to perform real-time forecasting and make proactive business decisions.

Solution

The company deployed a data intelligence system that:

- Unified and centralized data across business functions
- Standardized analysis methods to ensure reliable, consistent data insights
- Integrated Databricks AutoML models with AI/BI Genie to enhance demand forecasting
- Enabled inventory analysis to predict and prevent stockouts, ensuring product availability
- Integrated with Microsoft Teams, allowing users to query data through natural language and access real-time insights within their collaboration environment

Results

Analysis and decision-making processes became significantly faster, with data processing times reduced to minutes. These improvements enhanced demand forecasting, inventory management and business responsiveness to market needs.

https://www.databricks.com/customers/grupo-casas-bahia/ai-bi-genie?itm_data=bimeetsaiebook

Building trust through transparency

As more business decisions become based on these insights, trust becomes paramount. Modern compound AI solutions ensure that every data point is traceable to its source and that each metric's logic is transparent. For example, a CFO reviewing margin analysis can see:

- **Data lineage:** Exactly which databases and tables the figures came from.
- **Calculation transparency:** A breakdown of how margins were calculated, including which costs and revenues were considered.
- **Confidence scores:** Trust indicators showing visual cues and labels to validate the reliability of the insights, helping teams prioritize actions.

This transparency reassures users that behind every quick answer is well-governed, validated data. Over time, trust fuels more frequent and impactful use, embedding data-driven thinking into daily operations.

Learning and improving over time

As users interact with these tools, the system evolves. Each query and clarification teaches it more about an organization's data, terminology and analytical needs. Initially, only regional sales might be explored, but over weeks or months, the system learns key metrics, anticipates follow-up questions and provides more relevant guidance.

- **Adapting terminology:** Familiarity with product lines, regional codes and industry jargon leads to more precise responses.
- **Tailoring responses:** By recognizing patterns in user inquiries, the system can offer helpful suggestions, highlight recurring issues and streamline complex investigations.
- **Continuous improvement:** Feedback loops — where users can confirm accuracy or request revisions — help refine calculations and surface the most pertinent data faster.

In short, as the system learns from each interaction, it grows more aligned with your strategic goals and more adept at steering you toward valuable insights.

Expanding into advanced insights

After mastering basic queries, organizations can go deeper. Consider James from the sales team, now comfortable exploring store performance:

- **James:** "If I offer a 5% discount next month, can we estimate the potential revenue change in the Western region?"
- **AI-Bot (advanced mode):** "Based on historical promotions, a 5% discount may boost monthly revenue by about 3%, though margins might decrease by 1%. Would you like to compare this scenario to other discount levels?"

This forward-looking scenario shows how conversational analytics can evolve from a retrospective reporting tool into a strategic advisor, helping leaders model outcomes and make data-backed business decisions.

Delivering impact

As conversational analytics tools mature and weave into everyday workflows, their influence spreads organization-wide. They do more than expedite analysis; they foster a data-driven culture where smart decisions become instinctive.

- **Accelerated decision-making:** Faster answers mean teams can act promptly on fresh insights.
- **Data-driven collaboration:** Shared, transparent insights unite cross-functional groups around consistent facts and figures.
- **Empowered teams:** Business users gain autonomy, reducing bottlenecks and reliance on specialized analysts.
- **Improved data and AI literacy:** As everyone grows comfortable interacting with data and AI, the entire organization becomes more fluent in using insights to guide strategy.

Over time, this shift from reactive reporting to proactive intelligence turns data into a sustainable competitive advantage. The insights gained today lay the groundwork for better decisions tomorrow, ensuring that conversational analytics doesn't just solve problems — it drives transformative growth.

CHAPTER 8

Databricks AI/BI: Advancing to Real-World Intelligent Analytics

For two decades, BI promised to democratize data, yet many organizations still struggle to move from raw information to meaningful insights. Data teams spend months building dashboards that often remain underutilized, while business users wait on specialists to create reports and dashboards. The tension between the need for immediate, actionable intelligence and the complexity of corporate data has defined the next generation of self-service analytics.

A new era of AI-driven BI

Recent advancements in large language models (LLMs) and data intelligence have opened new possibilities in analytics. Instead of forcing users to learn SQL or navigate complicated BI tools, we can let them simply ask questions and receive contextually accurate answers in everyday language. This is the idea behind AI/BI, an approach that uses AI-driven semantics and a deep understanding of organizational data to translate common questions into precise queries and insightful results.

This is why Databricks undertook the development of AI/BI. Their vision was to deliver a solution that allows both business and technical users to move rapidly from high-level questions like “What is our revenue growth in the Western region?” to trustworthy, validated, data-backed answers. Databricks AI/BI merges the governance and reliability of the lakehouse with the intelligence and accessibility of modern AI.

Why Databricks built AI/BI

Databricks created AI/BI to address three core challenges that organizations face:

1. **Complexity vs. accessibility:** Critical insights often remain hidden behind obscure metadata and complex data models. Databricks AI/BI uses LLMs trained on a company's specific data and semantic context, so it can interpret plain-language questions and translate them into accurate SQL queries and verified answers.
2. **Static dashboards vs. dynamic conversations:** Traditional BI often relies on static dashboards that remain unchanged for weeks. AI/BI shifts from a "report-centric" model to a dynamic, conversational approach. Users can iteratively refine their questions and discover insights organically, rather than relying on pre-built views or waiting for new reports.
3. **Delays and bottlenecks:** Instead of waiting days for custom analyses prepared by experts, users can get timely answers by asking simple questions. With AI/BI, data teams can focus on strategic initiatives while nontechnical users confidently explore and create insights from data on their own.

The core principles of AI/BI

Rather than pursuing a bolt-on AI strategy, Databricks built AI/BI from the ground up with these three guiding principles in mind:

- **Data intelligence at the core:** AI/BI is not limited to surface-level text parsing. It understands a company's data model, relationships and semantics, ensuring responses are both contextually relevant and rigorously correct.
- **LLM-powered reasoning:** LLMs enable users to ask natural questions. However, the true advantage comes from combining these LLMs with a sturdy foundation of data intelligence and lineage. This ensures that answers are not only plausible but also auditable and grounded in verified enterprise data.
- **Trust, transparency and governance:** AI/BI is not a black box. Users can see how the answers were derived, trace them back to the original data source and verify the calculations. This openness builds trust and encourages adoption, because users know that every insight meets the organization's standards for accuracy and transparency.

AI/BI key capabilities

With AI/BI, Databricks brings together the foundational elements we've explored throughout this book — data intelligence, compound AI, and dynamic, conversation-driven analytics — into a cohesive solution. Its two core components work hand in hand to deliver insights — when and where they matter most.

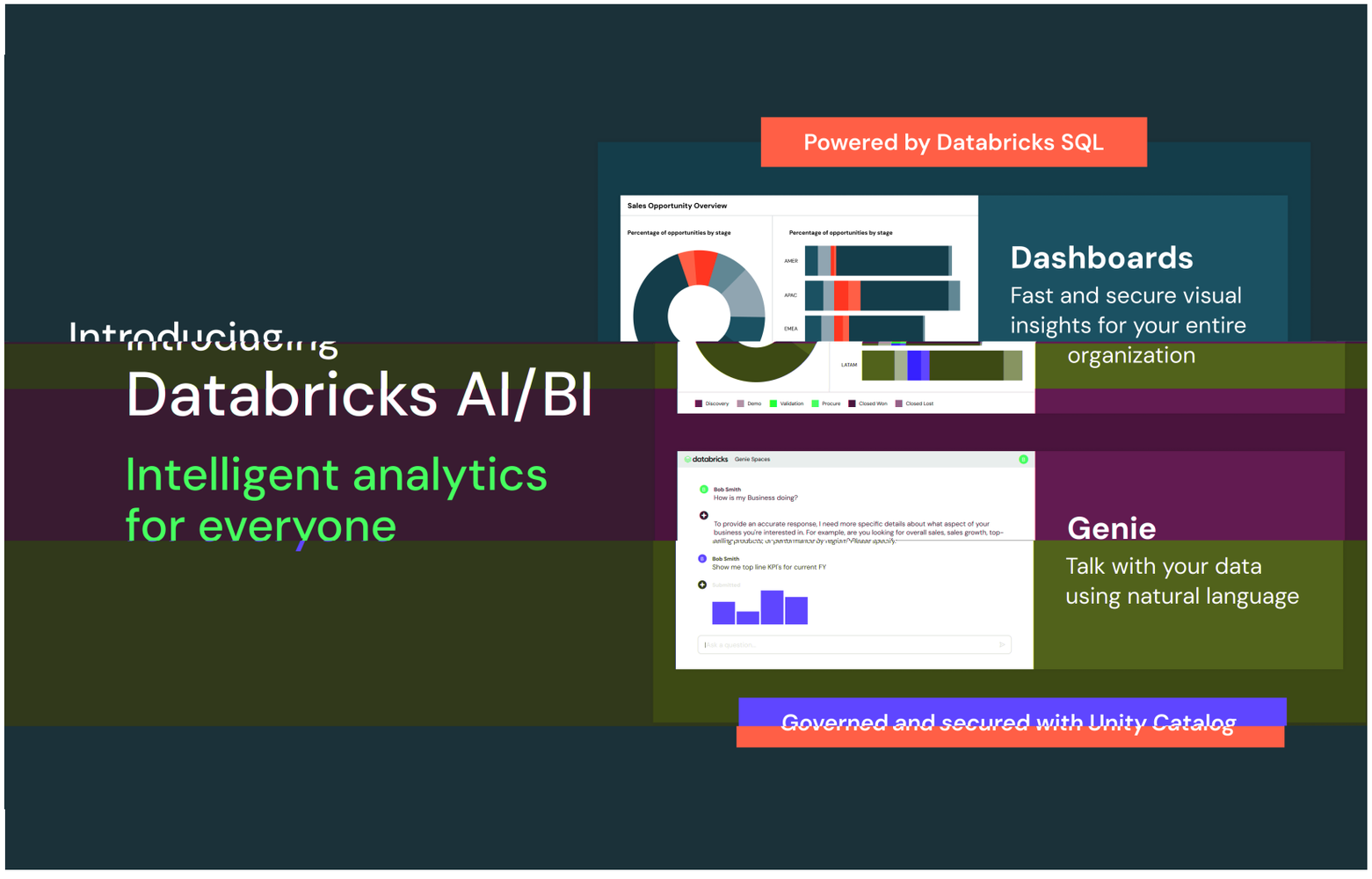


Figure 8.1: Databricks AI/BI — intelligent analytics for everyone

AI/BI Dashboards is an AI-assisted experience for BI practitioners to quickly create analytical datasets, interactive dashboards and data visualizations for their business teams. Native to the Databricks Data Intelligence Platform, AI/BI Dashboards enable Databricks customers to generate visual analytics and insights from the entirety of their data without the need to maintain a separate BI system.

AI/BI Genie is the star of the show, elevating self-service analytics beyond interactive dashboards. Business users can now independently converse with their data in natural language, iterating and refining questions as easily as if they were speaking with an expert analyst. By continuously learning from user behavior and feedback, Genie grows smarter and more contextually aware over time. Rather than making unfounded assumptions with ambiguous requests, it seeks clarification from the user, ensuring that insights remain accurate, actionable and free from "hallucinations." This interactive feedback loop brings the vision outlined in previous chapters to life, delivering on the promise of data intelligence that genuinely understands your business and speaks your language.

AI/BI Dashboards and AI/BI Genie are tightly integrated into the Databricks Data Intelligence Platform, ensuring one copy of data with unified governance and security through Unity Catalog.

From vision to reality: AI/BI Genie in action

Consider Sarah's situation. She once spent hours sifting through dashboards and reports to find meaningful insights about her marketing spend and customer churn. With AI/BI Genie, her morning is different:

- **8:30 AM:** Over coffee, Sarah asks, "How's my business doing?"
- **8:31 AM:** Genie interprets her question and finds an official SQL query function designed to answer this exact question. Genie runs the function and returns the answer labeled as "Trusted" so Sarah can be confident the answer is exactly what she needs.

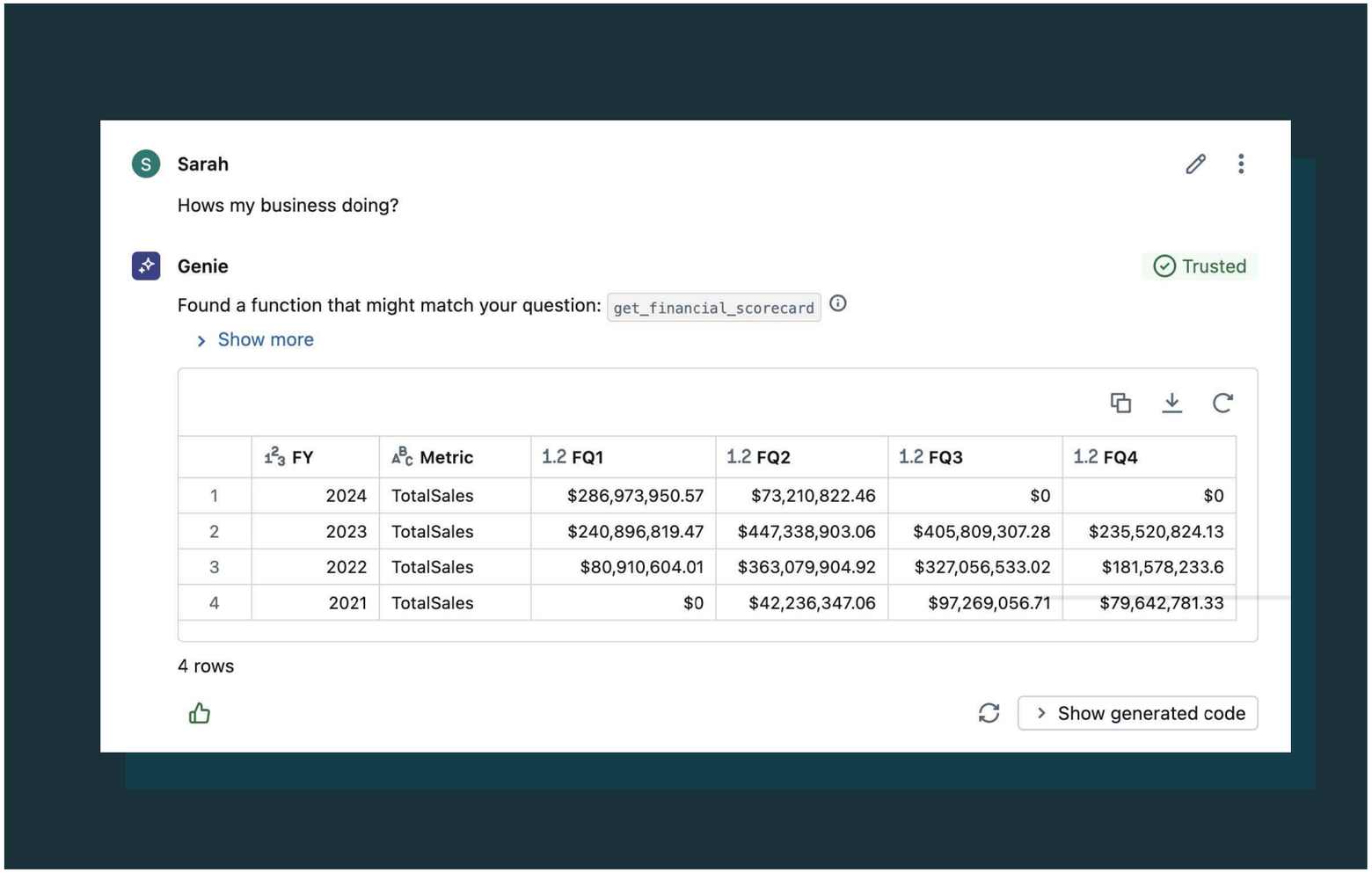


Figure 8.2: Databricks AI/BI Genie delivers trusted answers

- **8:31 AM:** Interested in marketing spend, Sarah seeks further information: "Show total spend by territory in FY24."
- **8:32 AM:** Again, Genie interprets her query and returns a relevant answer, explaining its reasoning as part of the process.

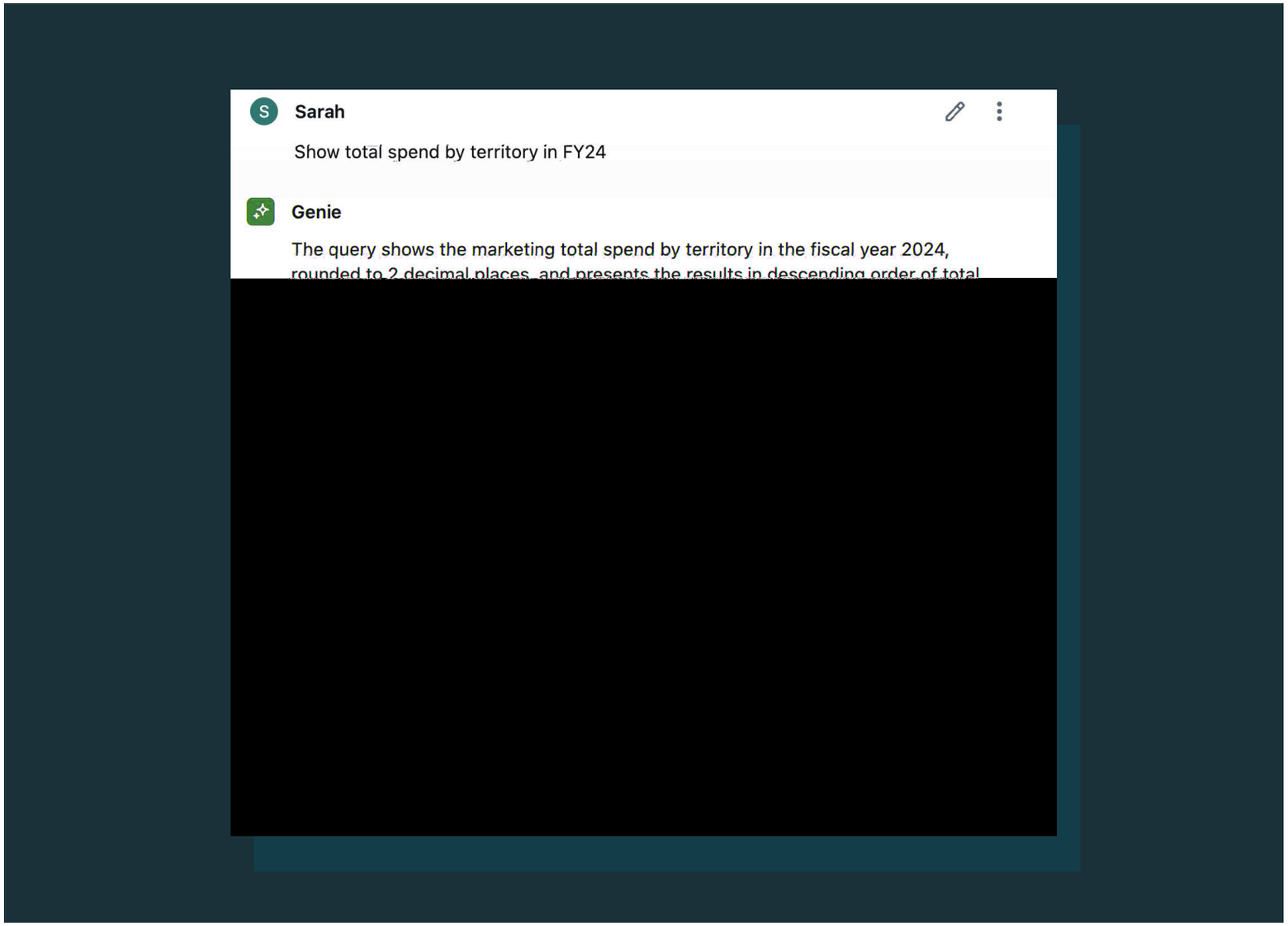


Figure 8.3: AI/BI Genie provides explanations — which fosters trust

- **8:33 AM:** Curious about customer churn, Sarah requests: "Show lost revenue due to customer churn in FY24 Q1."
- **8:33 AM:** Genie hasn't encountered the concept of customer churn before. Instead of making up something and hallucinating, it asks Sarah for clarification.

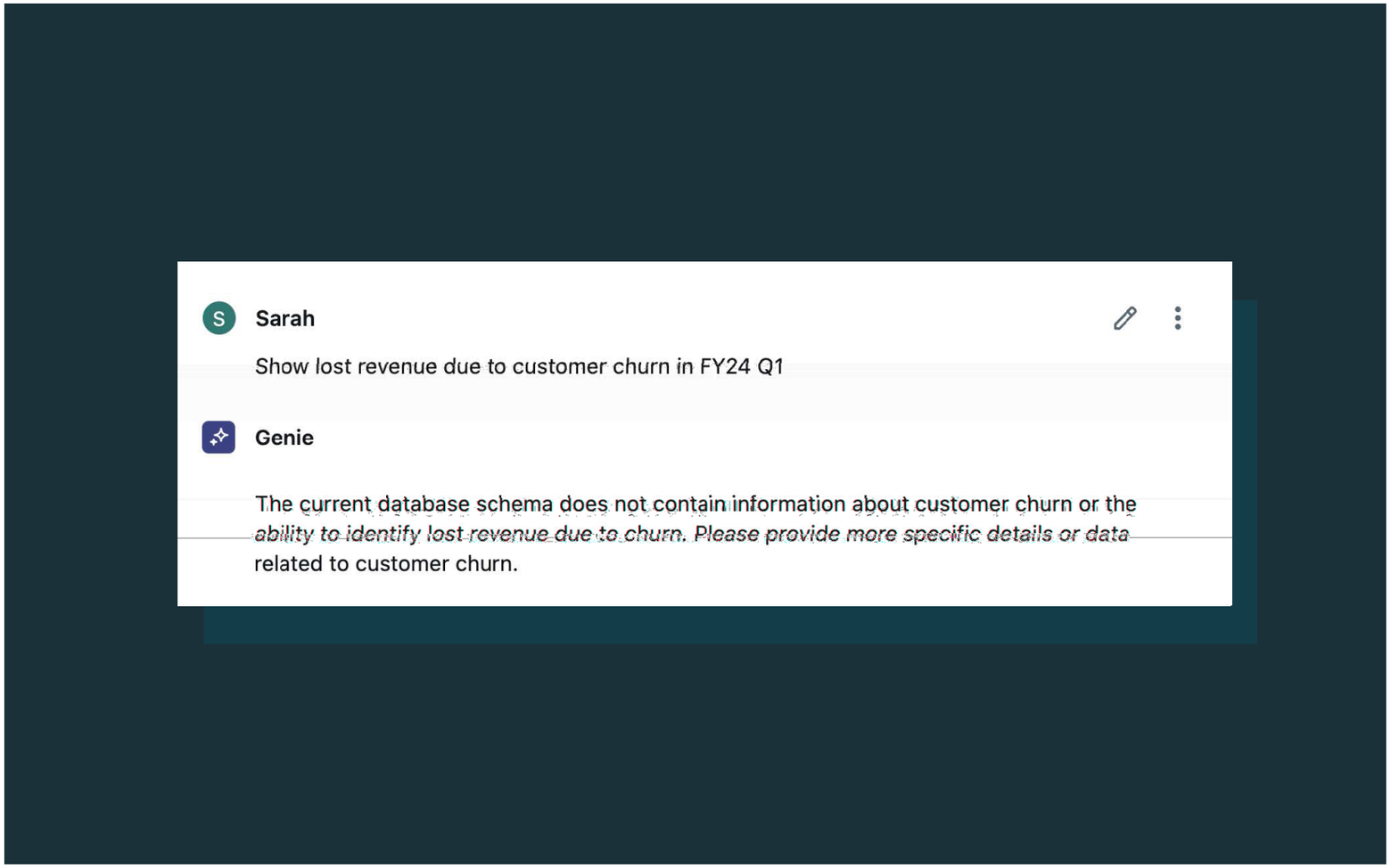


Figure 8.4: Genie asks for clarification when it encounters unknown terms

- **8:34 AM:** Sarah describes the definition of customer churn to Genie in the exact same way she would to a human analyst.
- **8:35 AM:** In real time, Genie learns this new definition and applies it to the data.

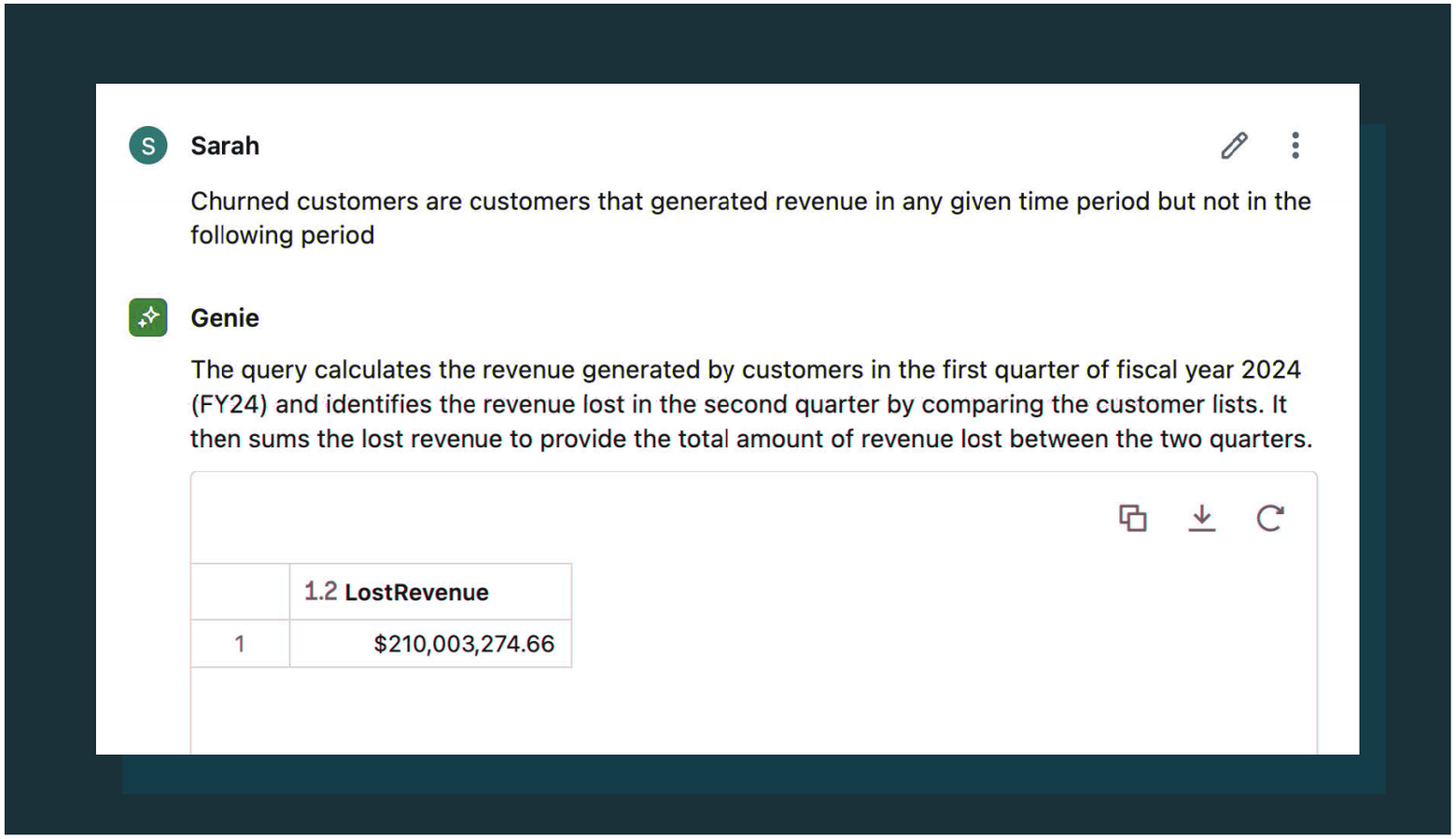


Figure 8.5: Genie uses the trusted definition to deliver trusted results

- **8:35 AM:** For the results to become actionable, Sarah needs to know exactly which customers churned in Q1. She also wants to narrow it down to just the Southwest territory since she knows her marketing spend is highest there from a previous question. She requests: "Show by customer but only for Southwest territory."
- **8:35 AM:** Genie is able to use the definition of churn it learned in the previous step and applies additional logic to the query to return exactly what Sarah needs. Genie also provides a detailed explanation of its reasoning so Sarah knows the answer is correct.

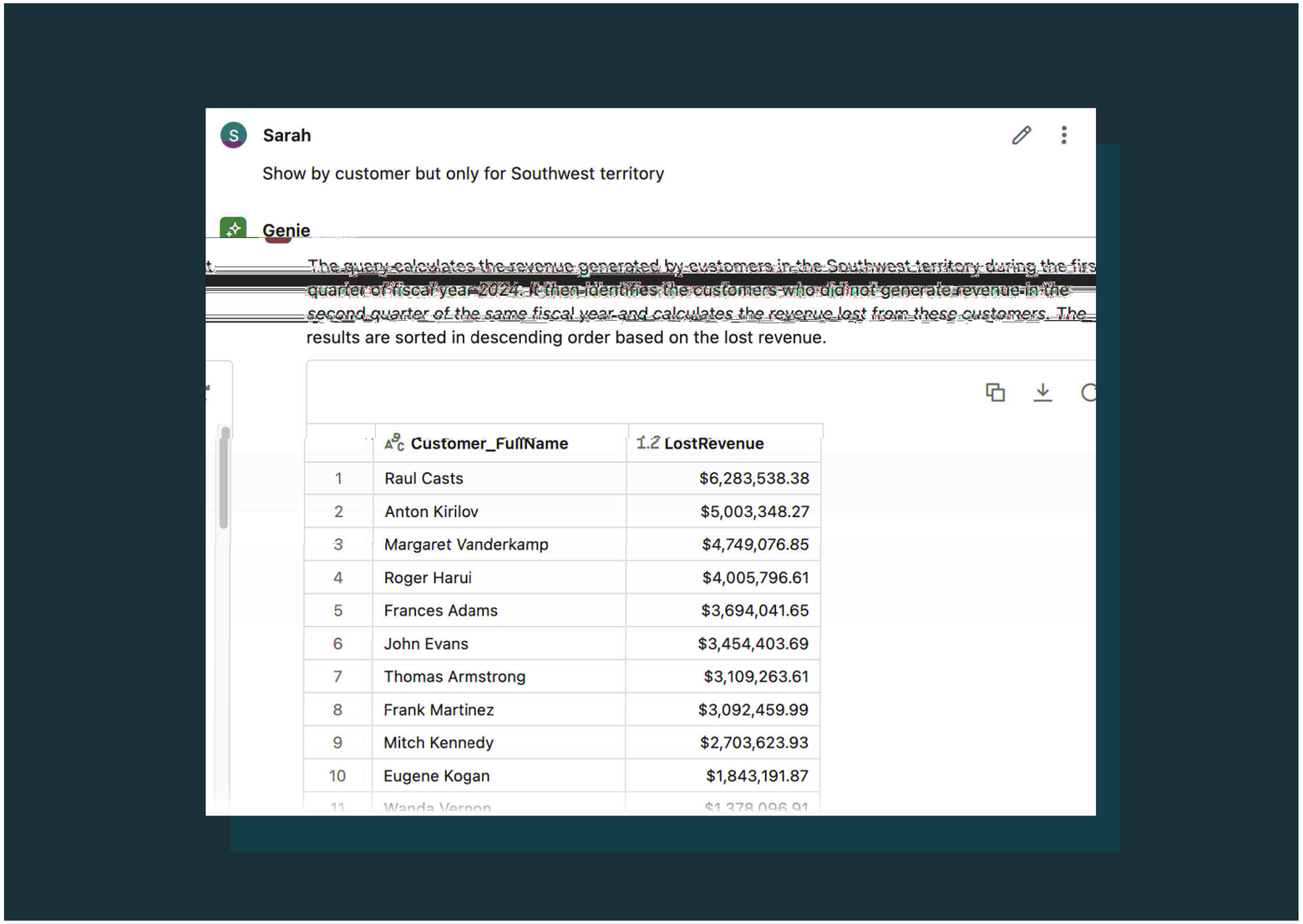


Figure 8.6: Genie drills downs and explains the query logic

This seamless interaction shows how AI/BI helps users think strategically, act quickly and trust the quality of their insights.

Real-world outcomes: driving business value today

The transition to AI-driven BI is already underway. Organizations across industries are helping their employees interact with data as easily as if they were speaking to a colleague. This shift leads to:

- **Accelerated decisions:** Answers arrive in seconds, reducing time to insight and enabling proactive strategy adjustments.
- **Democratized analytics:** Removing technical barriers lets more employees join data-driven conversations, reducing reliance on a handful of specialists.
- **Confidence in data:** Clear lineage, transparent calculations and built-in governance give decision-makers confidence that the insights they rely on are both timely and correct.

Real-world impact

Let's look at some examples of how organizations are realizing this vision:

HEALTHCARE INNOVATION: PREMIER INC.

Premier, a technology-driven healthcare improvement company, serves two-thirds of U.S. healthcare providers, underscoring their monumental scale and impact on the industry. Premier is on a continuous journey to eliminate fragmented data, reliance on manual processes and limited accessibility to advanced analytics, which slows the delivery of insights critical to improving patient care. With their adoption of the Databricks Data Intelligence Platform and AI/BI Genie, Premier is revolutionizing their analytics capabilities, enabling natural language queries, 10x faster SQL creation and seamless integration of data across systems. Now empowering healthcare providers with intuitive, self-service analytics to benchmark care, identify opportunities for improvement and accelerate decision-making, Premier is helping to transform healthcare operations and save lives on a national scale.

https://www.databricks.com/customers/premier/ai-bi-genie?itm_data=bimeetsaiebook

GAMING EVOLUTION: FUNPLUS

FunPlus, a global gaming company with millions of daily users, is leveraging AI/BI Genie and Streamlit to accelerate data-driven decision-making. AI/BI Genie is integrated into their BI visualization platform via an API, providing an interface for business users to easily query operational metrics, props production, and order and sales data using natural language, greatly improving the efficiency of accessing data across the organization.

https://www.databricks.com/customers/funplus?itm_data=bimeetsaiebook

IT FINANCIAL MANAGEMENT IMPROVEMENTS: MAGICORANGE

MagicOrange empowers enterprises to optimize technology spend and derive actionable insights from vast and complex datasets. Since adopting Databricks AI/BI Genie, MagicOrange has transformed their operations, enabling real-time, natural-language querying for both internal teams and clients. This has reduced report generation times to minutes and saved the company hundreds of thousands of dollars annually by eliminating costly legacy tools. With Genie's advanced analytics and Databricks' robust security and scalability, MagicOrange is unlocking new levels of innovation, setting the stage for future advancements in AI-driven insights.

https://www.databricks.com/customers/magic-orange/AI-BI-Genie?itm_data=bimeetsaiebook

> "At T-Mobile, we have leveraged AI/BI to democratize access to critical datasets across our business. It empowers our users, even without SQL expertise, to independently analyze information and trends through natural language queries."
>
> — Arvind Krishnamoorthy, Senior Data Scientist, T-Mobile

> "Users were constantly pinging analysts to say, 'Hey, what did sales look like for the Southwest region compared with the previous year?' The idea of being able to just ask Genie, rather than hunt for the right analyst and hope they get the answer right, has been very exciting for the business."
>
> — Shahmeer Mirza, Senior Director of Data, AI/ML and R&D, 7-Eleven

Getting started with Databricks AI/BI

Databricks includes AI/BI capabilities at no additional cost with Databricks SQL, lowering adoption hurdles and accelerating time to value. Whether you are in the exploratory phase or rolling out across the enterprise, a wide range of resources is available:

Introduction to AI/BI: Better understand foundational concepts and capabilities:

- https://www.databricks.com/product/business-intelligence?itm_data=bimeetsaiebook
- https://www.databricks.com/product/business-intelligence/ai-bi-genie?itm_data=bimeetsaiebook
- https://databricks.com/blog/introducing-aibi-intelligent-analytics-real-world-data?itm_data=bimeetsaiebook

Interactive demos and tutorials: Watch and install demos, practice refining complex inquiries step-by-step and see how everyday questions become actionable insights:

- https://www.databricks.com/resources/demos/library?q=ai%2Fbi

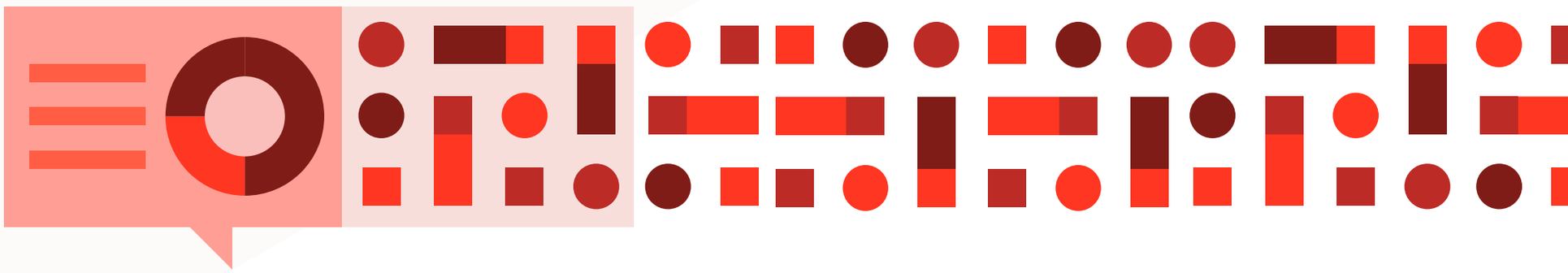
Technical documentation and training: Dive into implementation details, best practices and optimization strategies for accuracy and trust:

- https://docs.databricks.com/en/ai-bi/index.html
- https://www.databricks.com/resources/learn/training/get-started-sql-analytics-and-bi?itm_data=bimeetsaiebook

Databricks is being used by several thousand practitioners every day. The company understands the need to equip users with skills to unlock insights from their data. New courses address this need by equipping learners with the knowledge to utilize Databricks SQL, Delta Lake and Databricks AI/BI Dashboards — key tools for turning raw data into actionable insights.

The introductory Get Started With SQL Analytics and BI Course helps learners to:

- Use Databricks to complete common data analytics tasks
- Use Databricks SQL for data analysis and reporting
- Leverage Delta Lake capabilities to support data exploration and manipulation
- Create Databricks AI/BI Dashboards within Databricks to visualize data
- Create AI/BI Genie spaces that allow business users to converse with their data, use questions to perform queries, and self-serve analytics and visualizations



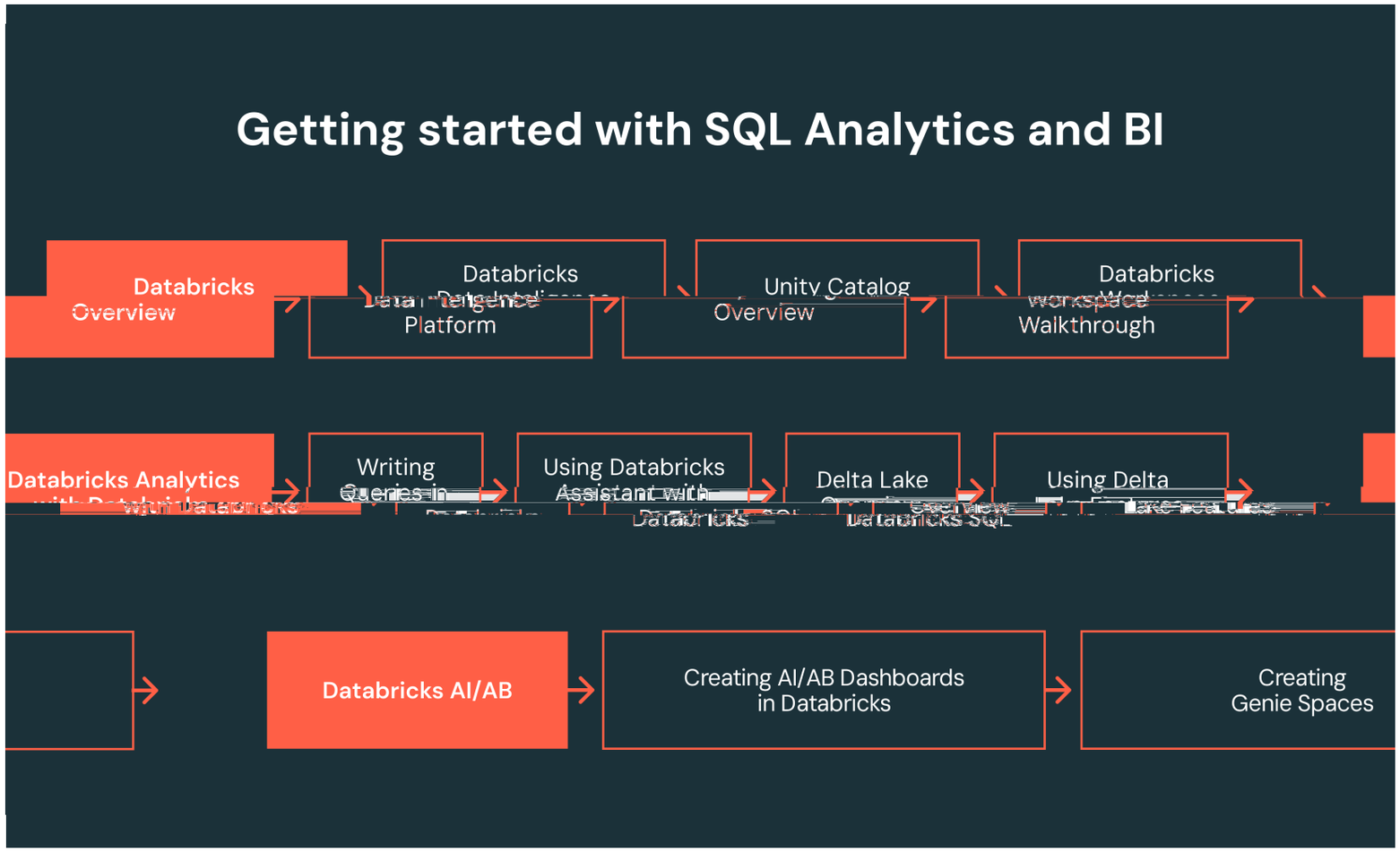


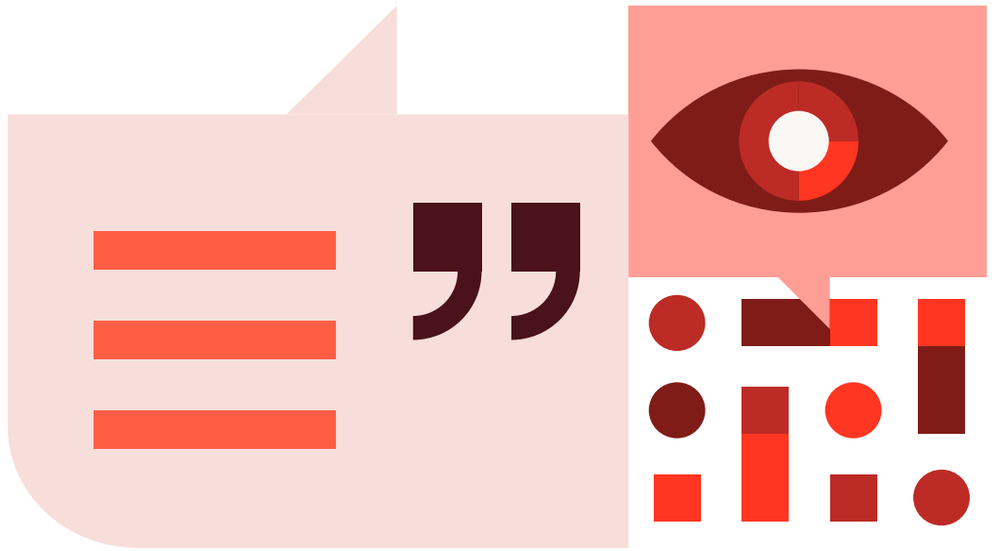
Figure 8.7: Getting Started With SQL Analytics and BI

By blending data intelligence, large language models and conversational interactions, Databricks AI/BI bridges the gap between data potential and data reality. The platform's unified governance and intuitive interface ensure that everyone — not just data specialists — can participate in shaping strategies and driving business outcomes. This is the culmination of the journey we've explored in this book, one of moving from rigid reports and dashboards to an intelligent, dynamic analytics ecosystem that learns with each interaction.

When considering how to implement these tools in your organization, reflect on the core concepts we've covered:

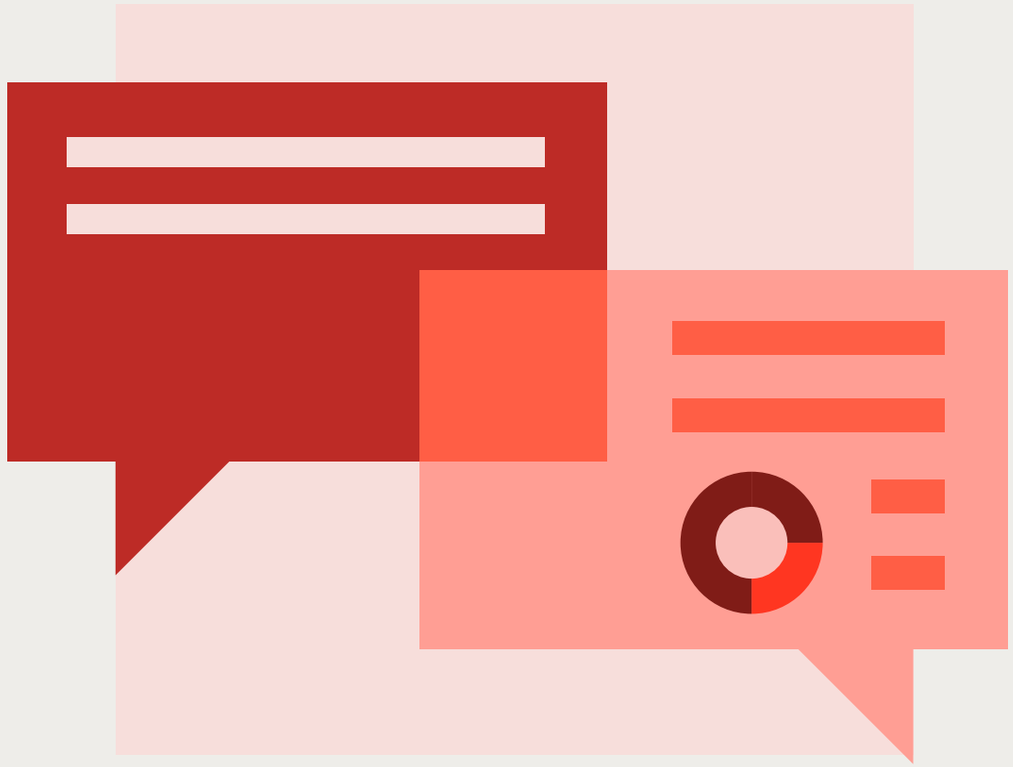
- **Data intelligence** provides the semantic backbone to understand information
- **Compound AI** orchestrates multiple specialized agents to solve complex problems
- **Conversation-driven experiences** transform how leaders ask questions and find answers

Together, these elements create an environment where insights flow naturally, decision-makers gain confidence in their data, and teams collaborate more effectively. In short, AI/BI represents a practical, forward-looking path to making data-driven decision-making a reality — rather than just an aspiration.



About Databricks

Databricks is the data and AI company. More than 10,000 organizations worldwide — including Block, Comcast, Condé Nast, Rivian, Shell and over 60% of the Fortune 500 — rely on the Databricks Data Intelligence Platform to take control of their data and put it to work with AI. Databricks is headquartered in San Francisco, with offices around the globe, and was founded by the original creators of Lakehouse, Apache Spark™, Delta Lake and MLflow. To learn more, follow Databricks on LinkedIn, X and Facebook.



databricks