The State of Data, Analytics, and AI in Commercial Biopharma

Accelerating commercial AI and analytics adoption at scale

The AI paradox: ambition vs reality

A top pharmaceutical company saw its new AI tool deliver promising results in a key market. Confident in the pilot's success, the organization moved quickly to scale the solution to other markets. But the rollout stalled. Developing the algorithms in new markets took longer and the algorithms that worked in one market failed in others. Adoption by the field teams was low because they did not trust the recommendations the tool generated. Success proved difficult to replicate and the company quickly identified the reason as the lack of a common, reliable data foundation across markets.
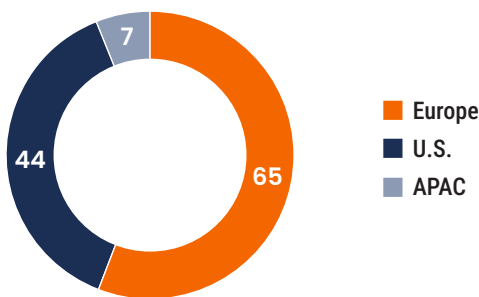
This isn't an isolated case. It reflects a central paradox facing the life sciences industry today: the enormous value potential of AI is being held back by a fundamental data problem.

To explore this disconnect, this report draws on insights from a survey of 116 senior life sciences leaders, from large, global enterprises in Europe, US, and Asia Pacific. A full 85% of these respondents directly oversee their organization's commercial analytics and AI initiatives, making them uniquely qualified to comment on the state of innovation. These leaders reveal a clear narrative: to innovate, business leaders must prioritize building a robust data foundation.
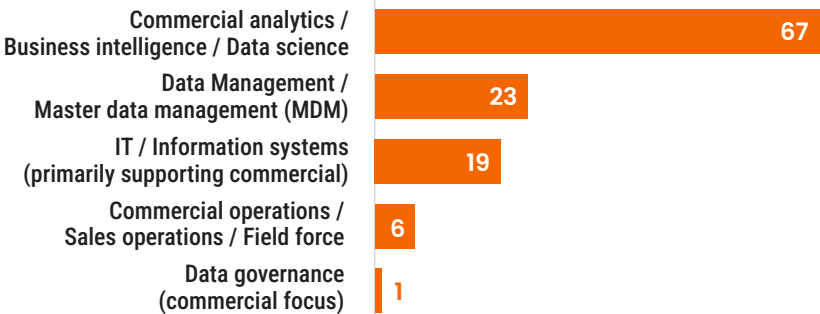
This report will diagnose the three core data issues driving this failure: ***trust, speed, and consistency***, and outlines a clear path to bridge the gap between AI ambition and reality.

Insights from 116 Life Sciences Commercial Data and Analytics Leaders
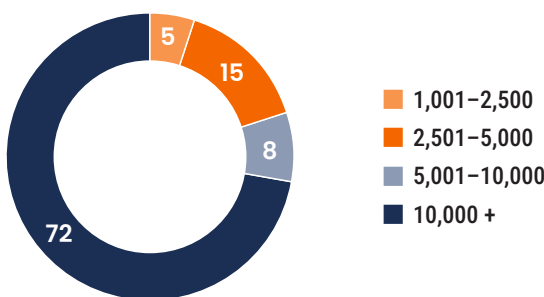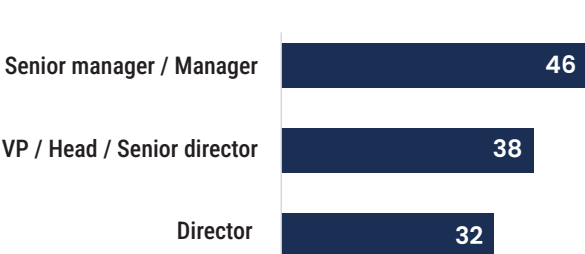
Region

- Europe — 65
- U.S. — 44
- APAC — 7

Job function

Function	Count
Commercial analytics / Business intelligence / Data science	67
Data Management / Master data management (MDM)	23
IT / Information systems (primarily supporting commercial)	19
Commercial operations / Sales operations / Field force	6
Data governance (commercial focus)	1

Company size

- 1,001–2,500 — 5
- 2,501–5,000 — 15
- 5,001–10,000 — 8
- 10,000 + — 72

Seniority

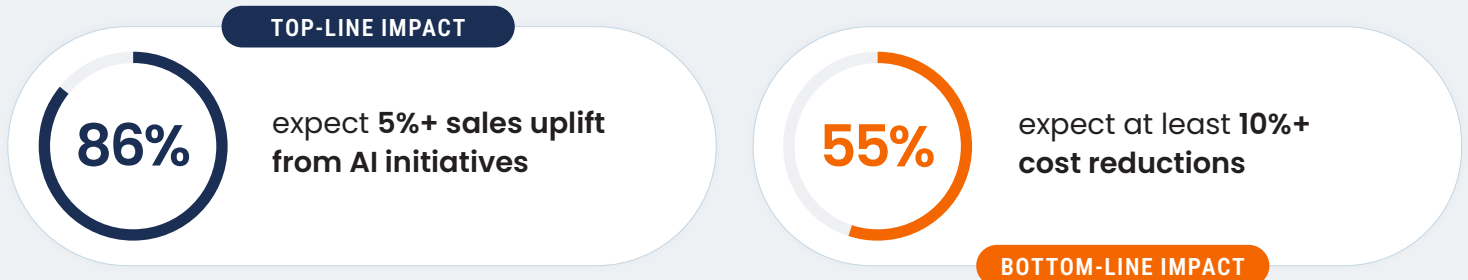| Level	Count
Senior manager / Manager	46
VP / Head / Senior director	38
Director	32

AI's potential drives ambition to scale

Commercial leaders are betting big on AI, anticipating transformative impacts on the top and bottom lines.

The commitment is clear, with organizations actively pursuing a range of sophisticated AI and analytics use cases. Their focus is on high-value usage, including **AI for summarizing healthcare professional (HCP) updates and complete call reports** (72%), **segmentation analytics for targeting** (62%), **and building an HCP 360-degree view in CRM** (60%).
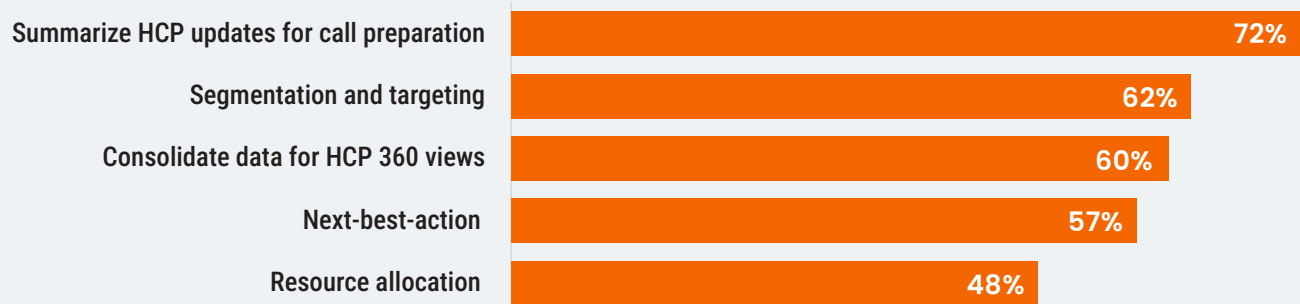
AI's Transformative Potential in Commercial Pharma
High expectations for top-line growth and bottom-line savings

TOP-LINE IMPACT

86% expect **5%+ sales uplift from AI initiatives**

55% expect at least **10%+ cost reductions**

BOTTOM-LINE IMPACT

What analytics or AI initiatives are you currently pursuing in your Commercial function?
Select all that apply.

Summarize HCP updates for call preparation	72%
Segmentation and targeting	62%
Consolidate data for HCP 360 views	60%
Next-best-action	57%
Resource allocation	48%

AI pilots show promise, yet companies struggle to scale

Despite the clear ambition and investment, the value potential of AI at scale remains largely uncaptured. The survey reveals a critical struggle to move initiatives from pilot to enterprise-wide scale, undermining the potential ROI and leaving strategic goals unmet.

Scaling AI Proves Challenging
Limited readiness and success in AI initiative deployment

96%
believe their **data is not well-structured** and not AI-ready for scaling

89%
were not able to **scale more than half** of their initiatives

67%
reported having to **abandon an AI initiative** due to bad data

PROBLEM

Fragmented HCP data leading to:
1. Local vs. global discrepancies
2. Manual cleanup required

CONSEQUENCE

Slower time to market
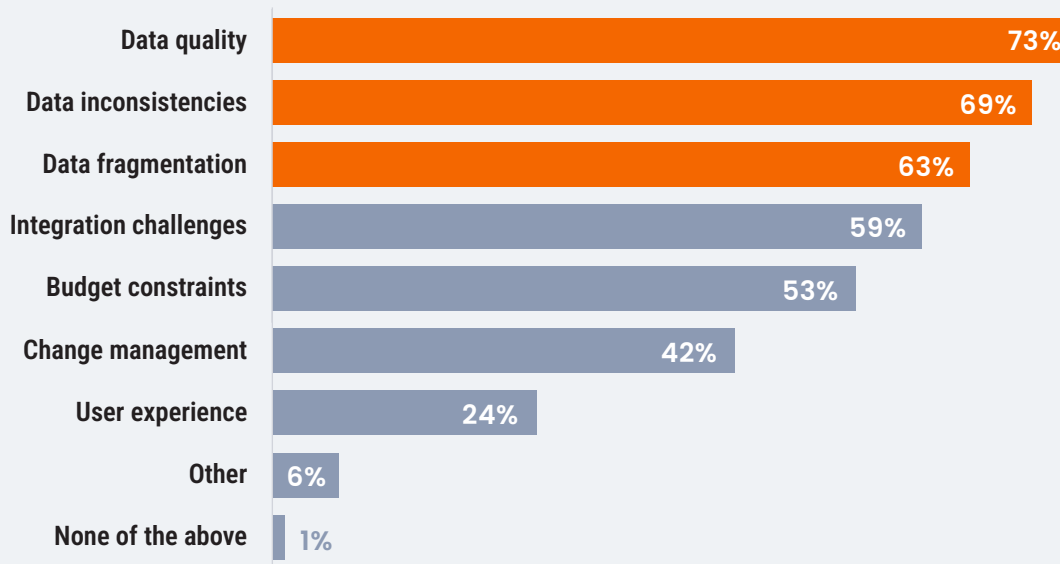1. Two-month delay
2. 15% lower early script volume

This inability to scale is not trivial. One case highlighted that fragmented data led to a **two-month delay** in time-to-market and **15% lower early script volume**. Instead of making progress, companies are caught in a costly loop of isolated pilot projects that fail for the same unaddressed data reasons.

Three main issues stalling progress

Diving deeper into the root causes of why these pilots are not scaling, leaders identify three core issues in their foundational data — data quality, data inconsistencies, and data fragmentation — that directly manifest as enterprise-wide challenges to **trust**, **speed**, and **consistency**.

Which of the following challenges are you currently experiencing in your organization?
Select all that apply.

Challenge	Percentage
Data quality	73%
Data inconsistencies	69%
Data fragmentation	63%
Integration challenges	59%
Budget constraints	53%
Change management	42%
User experience	24%
Other	6%
None of the above	1%

TRUST
The impact of data quality on stakeholder confidence

Unreliable data erodes the foundation of any analytics program by creating a ripple effect of mistrust impacting the entire organization, from the executive level to the customer.

BUSINESS LEADER TRUST: Losing confidence in data-driven initiatives

Mistrust begins when business leaders see skilled teams trapped in a reactive cycle of data cleanup when there is no single source of truth. One leader says **it takes weeks to get the data into a working format**, which diminishes strategic value and erodes confidence, directly impacting strategic planning. For example, without reliable and consistent data, leaders cannot make data-driven decisions for consistent resource allocation and target-setting across markets. As a result, they cannot track performance on a comparable basis and must rely on subjective judgment.

The scale of this foundational challenge is significant:

73% of leaders experience significant **data quality issues**

96% believe **their data is not well-structured** or AI-ready for scaling

81% said **data quality is their top data issue** to address in the next 12 months

This crisis of confidence extends beyond human decision-making. As the industry moves toward using agentic AI, the problem of "hallucination", where AI generates non-factual information, is severely amplified by unreliable input data, making trust more critical than ever.

FIELD TEAM TRUST: Rejection of unreliable tools

For field teams, trust erodes when granular data they receive is obviously incorrect. When tools such as next best action (NBA) use flawed data, it generates poor recommendations that conflict with field teams' deep, on-the-ground knowledge of customers and the system loses all credibility.

As a result, the field force rejects the tools and reverts to individual and varied methods and intuition. One leader says, "A lack of high-quality data made go-to-market strategy a guesswork process. Users trusted their intuition more than insights, which is hard to overcome." The result is a disconnect between the data-driven strategy from headquarters and the reality of execution in the field.

> "
> *A lack of high-quality data made go-to-market strategy a guesswork process. Users trusted their intuition more than insights, which is hard to overcome."*
>
> **Leader**

HCP TRUST: On the receiving end of uncoordinated and poor engagement

Ultimately, internal chaos spills outward and can damage customer relationships and company reputation. For example, many companies pursue sophisticated account-based strategies, yet don't have a clear view of which HCPs belong to which healthcare organization (HCO). The result is a frustrating experience for the customer. An HCP might be contacted by multiple account teams from the same company about the same topic, receive information irrelevant to their specialty, or interact with a new team member lacking historical knowledge of the relationship. A disjointed approach signals a disorganized organization, eroding the HCP's trust and willingness to engage.

SPEED
Analyzing bottlenecks in the data-to-insight value chain

Data is often inaccessible, disconnected, and misaligned. The effort required to access, clean, and assemble data creates delays, making insights irrelevant before they are generated. Friction points at each step, from raw data to actionable intelligence, perpetuate the delays and eventually prevent AI and analytics from being adopted at scale.

This friction impacts companies significantly:

72% of companies spend up to 100 days matching HCP data across sources

62% of companies spend up to 200 days navigating third-party access agreements (TPAs) to use the data they have

Accessibility

The process stalls at the very start because data is often not easily accessible. Leaders report that customer data is fragmented across different regional CRMs in varied formats, making it difficult to find a single source of truth. Even when found, it is often not easily accessible, locked behind internal firewalls or blocked by lengthy contractual hurdles. Navigating TPAs can take months, stalling initiatives before they can begin.

Interoperability

Once companies access data, the lack of a common data standard makes it difficult to use. A lack of interoperability imposes a heavy 'operational tax,' forcing teams to manually align records, reconcile conflicting formats, and coordinate with local teams. For example, a common case is that analysts need to develop, test and run different code to conduct one and the same analysis in different countries since data models (i.e., fields, names, definitions,, structure) do not align. This massive undertaking is a primary cause of delays in the data preparation process.

Freshness

Finally, data's value is compromised if it is not kept up to date. The time it takes to resolve a data change request (DCR) is a critical friction point. Leaders report that even simple DCRs for a name or address can take two weeks to resolve, while more significant inconsistencies can trigger investigations. A leader shares, "We had more than a one-month delay in key functions as we had to run a full investigation to ensure the validity of the entire dataset." When data is not kept current, it cannot be confidently reused.

> "
> *We had more than a one-month delay in key functions as we had to run a full investigation to ensure the validity of the entire dataset.*
>
> **Leader**

CONSISTENCY
The challenge of achieving a single source of truth

Data inconsistency is a pervasive challenge directly preventing organizations from achieving a single, reliable view of their customers and products. The problem is a natural consequence of a complex data ecosystem, where information is spread across a mix of in-house and external sources, different software systems, and various business units, each with its own standards. As a result, the same customer can look completely different depending on the data source, leading to widespread duplication and conflicting records.

The manual effort required to fix the issues is substantial:

95% of companies must remap global-to-local specialties, with 55% doing this annually or more frequently

84% of companies spend up to 100 days on HCP mapping processes alone

Pervasiveness of inconsistencies across the enterprise

The lack of a common data language or standards means that a single, unified view of the customer is nonexistent. Key attributes and standards vary widely across:

Systems	Inconsistencies are rampant across the tech stack. A single customer's data can be spread across different CRMs with different formats, creating technical roadblocks. One leader reports, *"A lack of mapping between SAP and CRM for many markets prevents us from deploying our AI Customer 360 tool."*
Countries and regions	Global standards often break down at the local level. Leaders consistently note: *"Inconsistent data sets from one provider across countries make it difficult to integrate globally."* This is especially true in Europe, where healthcare systems are so different from country to country that it's not that easy to scale initiatives.
Clinical and commercial functions	A significant source of inconsistency is the deep-rooted disconnect between clinical and commercial data. One leader reports that *"There are silos between business units such as commercial, supply chain, and R&D."* The data standards and identifiers used to track a principal investigator in a clinical trial are often completely different from those used for the same person in the commercial CRM, making it nearly impossible to get a true 360-degree view of a customer's journey.

Data coverage and completeness	A major challenge is the lack of comprehensive and complete data for different attributes across the enterprise. For example, a company may have excellent HCP-level data in North America but finds that the same fields "do not exist or are not commercially available" in other regions, making it impossible to apply a consistent analytical
Key master data attributes (e.g. customers and products)	The most fundamental attributes are often available but inconsistent. For customer master data, one leader states, *"I haven't seen a reliable solution across the globe to maintain specialties of HCPs and HCOs."* This inconsistency extends to core product master data as well. The same product may be listed with a local trade name in one country and a global brand name in another. Another leader shares, *"This lack of global taxonomy causes challenges with granular level analytics tracking,"* making it impossible to get a clear, cross-market view of performance without significant manual mapping.

The enterprise-wide impact of inconsistency

This widespread inconsistency creates a massive operational burden and stalls high-value projects. It forces teams into a painstaking effort to manually align and reconcile data with local teams, significantly increasing the resources required and reducing project agility. Ultimately, the inability to connect data across the enterprise makes a unified global strategy impossible, forcing companies to scale on a market-by-market basis rather than globally.

A call for a foundational fix

The survey reveals a clear consensus: deep-seated data issues cannot be solved with surface-level fixes. Leaders are rightly skeptical of a simple technology fix, with 71% stating they don't believe Generative AI alone can solve fundamental data quality and consistency issues. The statistic indicates understanding that AI is only as good as the data it is built upon.

71%
of leaders state they don't believe Generative AI alone can solve fundamental data quality and consistency issues.

Instead, leaders are pointing toward a foundational strategy that directly addresses the core problems: **a unified approach to data harmonization and governance**. Harmonization and governance are the two sides of a coin. While harmonization creates the common data language that solves for speed and consistency, ongoing governance ensures the quality and reliability needed to build trust.

The critical strategic question every organization must face is whether to **create this standard in-house or adopt an industry-wide model**.

Two paths to solve the harmonization-governance challenge

PATH #1: In-house harmonization

The traditional approach is to build and maintain a proprietary, in-house standard. Organizations that choose this path often do so because they aim to accommodate their unique business needs, which often vary across the company. While the in-house path prioritizes full control, it implicitly means choosing not to make the tough strategic decisions required for standardization. As a result, companies are left paying the 'operational tax' of constant data cleansing and harmonization. The effort often yields diminishing returns. While significant progress can be made, achieving the final, most complex level of data quality can be prohibitively expensive.

PATH #2: Industry standardization

This emerging alternative directly addresses what 76% of life sciences leaders believe is the #1 enabler of AI at scale: global data harmonization. This path involves adopting an external, pre-harmonized data foundation. But a truly effective implementation requires more than just clean and harmonized data; it demands a new model: connected data and software supported by deep industry expertise.
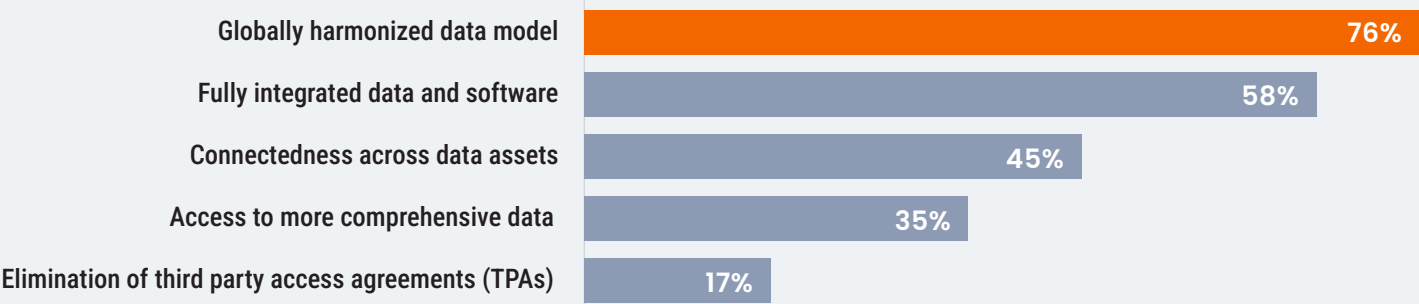
76%

of life sciences leaders believe global data harmonization is the #1 enabler of AI at scale.

Success hinges on three key factors.

01. It must solve the challenge of **balancing global consistency with local autonomy** by providing a centrally governed, harmonized core while still allowing for the flexibility for unique markets needs.

02. It must be delivered with an end-to-end **product mindset**, moving beyond manual file transfers to offer robust APIs and native software integrations that automate the flow of data across systems and lower the integration barrier.

03. Finally, it requires shifting from a simple data provider to a **strategic partner focused on co-creating value**. The goal is a holistic relationship that helps the organization get the most out of its data investment to solve critical business challenges.

What would most help overcome your data challenges?
Select all that apply.

Globally harmonized data model	76%
Fully integrated data and software	58%
Connectedness across data assets	45%
Access to more comprehensive data	35%
Elimination of third party access agreements (TPAs)	17%

Ultimately, the choice between the two paths is not about right or wrong, but about aligning with core strategic priorities. The in-house path prioritizes full control and deep customization, allowing an organization to build a proprietary data asset tailored to its unique business rules. In contrast, the industry standardization path prioritizes speed and scale, enabling an organization to leverage a pre-harmonized foundation to more quickly deploy analytics and AI. Each path requires a different type of investment, capabilities and commitment, and the right choice depends on an organization's specific goals and resources.

What a globally harmonized data foundation means across the organization

The insights from this report offer clear pathways to unlock the full potential of AI and analytics across the organization.

For commercial and business leaders

Championing a globally harmonized data foundation moves the organization from a go-to-market strategy with significant data and insight gaps to one that is truly data-driven. This enables more effective resource allocation and performance management across countries and empowers the precise, targeted AI initiatives required to achieve ambitious goals in sales uplift or cost reductions.

For IT and data leader

Adopting an industry standard for data eliminates the significant 'operational tax' of manual integration and frees teams from repetitive, low-value data cleansing tasks. The approach simplifies your architecture, lowers the total cost of operations, and provides the speed needed to accelerate critical IT projects and focus on high-value innovation like scaling AI.

For analytics teams

With a trusted single source of truth, the AI models built by analytics teams will have the credibility they deserve. This solves the problem of models producing unsatisfying results and unhappy users, allowing teams to move from data wrangling to insight generation.

For field teams

A reliable data foundation means field teams can finally trust the tools they are given. This moves their role beyond 'ticking the box' to driving meaningful, personalized engagement with HCPs. It also means gaining access to the latest AI innovations that were previously stuck in pilots and unable to scale to your market, providing reliable recommendations field teams can actually use.

Connected data and software that supercharges AI efforts

The path to scaling AI is not paved with more technology alone. For the **89% of companies** that have failed to scale their AI initiatives, overcoming the data challenges of trust, speed, and consistency fix is key to finally closing the gap between ambition and reality. By connecting pre-harmonized data and software, commercial teams can solve these data challenges and turn technology and AI into a durable competitive advantage.

Take the next step

Solving the data challenges outlined in this report requires a clear path forward.

Hear from your peers

Hear from leaders at Boehringer Ingelheim, Bayer, and Astellas on their journey toward standardizing on a global data model.

Assess your readiness

Contact us for a complimentary data assessment to learn how to take the first step toward a globally harmonized data foundation.

Explore Data Cloud

Veeva Data Cloud products are pre-harmonized and connected with Vault CRM, providing trusted, AI-ready data that fuels faster insights and smarter commercial execution

Veeva OpenData	Veeva Compass (US Only)	Veeva HCP Access	Veeva Link
Complete HCP and HCO profiles	Patient and prescriber data	HCP access data	Deep data on KOLs

Veeva

ABOUT US

Veeva is the global leader in cloud software for the life sciences industry. Committed to innovation, product excellence, and customer success, Veeva serves more than 1,100 customers, ranging from the world's largest biopharma companies to emerging biotechs. As a Public Benefit Corporation, Veeva is committed to balancing the interests of all stakeholders, including customers, employees, shareholders, and the industries it serves.

For more information, visit **veeva.com/opendata**