



Building a Global Data Foundation for Scaling AI

AI use cases are rippling across commercial biopharma, helping companies make faster, more informed decisions. Yet almost 70% of top generative AI (GenAI) users cite poor data quality as their most significant obstacle in unlocking AI's full potential. As the adoption of applications grows, the true competitive edge lies in the quality of the data fuelling them.

To fully harness AI, commercial leaders are establishing a scalable, seamlessly connected data foundation across markets, functions, and disease areas. Without it, companies' AI pilots could amount to isolated experiments. Those who focus on creating standardised and well-integrated data can unlock AI's full potential to gain a competitive advantage and drive long-term success.

Data Consistency and Connectivity: The Foundation of AI

Commercial biopharma teams are uniquely positioned to strategically leverage AI as they collect vast amounts of data, including customer, sales, medical engagement, and social media activity. The next step is to harmonise the data – essentially to “speak the same language” to generate accurate and scalable insights.

Consider a common scenario: One system lists a healthcare professional (HCP) as “John Smith” and another as “J. Smith.” Or perhaps “cardiology” is recorded in one database while “heart medicine” appears in another. AI may fail to connect the variations, leading to errors, duplication, and unreliable insights. These inconsistencies often stem from diverse data sources that don’t speak to each other, creating friction for AI and significantly reducing its ability to provide value.

In another example, a biopharma’s HCP database had over 25,000 specialty classifications, rendering AI-driven insights nearly impossible. The company resolved the issue by implementing global data standards, significantly improving accuracy and scalability.

While AI continues to improve in handling inconsistencies, its success still hinges on the quality of the data it's trained on. This is especially critical in commercial biopharma, where data is often fragmented, sparse and inconsistent, disrupting AI's ability to generate meaningful insights.

Bayer AG's Journey to AI-ready and Globally Standardised Data
Overcoming data consistency challenges requires an organisation-wide approach. Some biopharma leaders are already making strides by prioritising global data standardisation to connect data and run advanced analytics initiatives.

For example, Bayer AG sought to create a 360-degree customer view to provide its field teams with comprehensive

insights before engaging with HCPs. However, data silos across geographies made it challenging to achieve a unified view.

Stefan Schmidt, group product manager at Bayer AG, led the company's data harmonisation efforts. Schmidt understood that AI insights would remain unreliable without a centralised, accurate data foundation. “Our global data landscape was fragmented – different countries relied on different sources. To see the full picture, we needed a unified customer master,” Schmidt explains.

By harmonising data across geographies and functions, Bayer eliminated inconsistencies and improved accessibility. The company consolidated key data sources – CRM, engagement history, and customer profiles – into a single, intuitive platform for its sales teams.

“In just weeks, we developed a solution that our teams genuinely valued,” Schmidt shares. With a single, connected source of truth, Bayer AG is now positioned for scalable, AI-driven insights across the organisation.

How Commercial Leaders Can Scale AI

Bayer AG's experience demonstrates the power of a globally standardised data foundation and the importance of making it a strategic priority to scale the impact of AI.

To avoid the common pitfalls, commercial leaders must address three key data challenges:

1. Business: Moving AI Pilots from Isolation to Execution

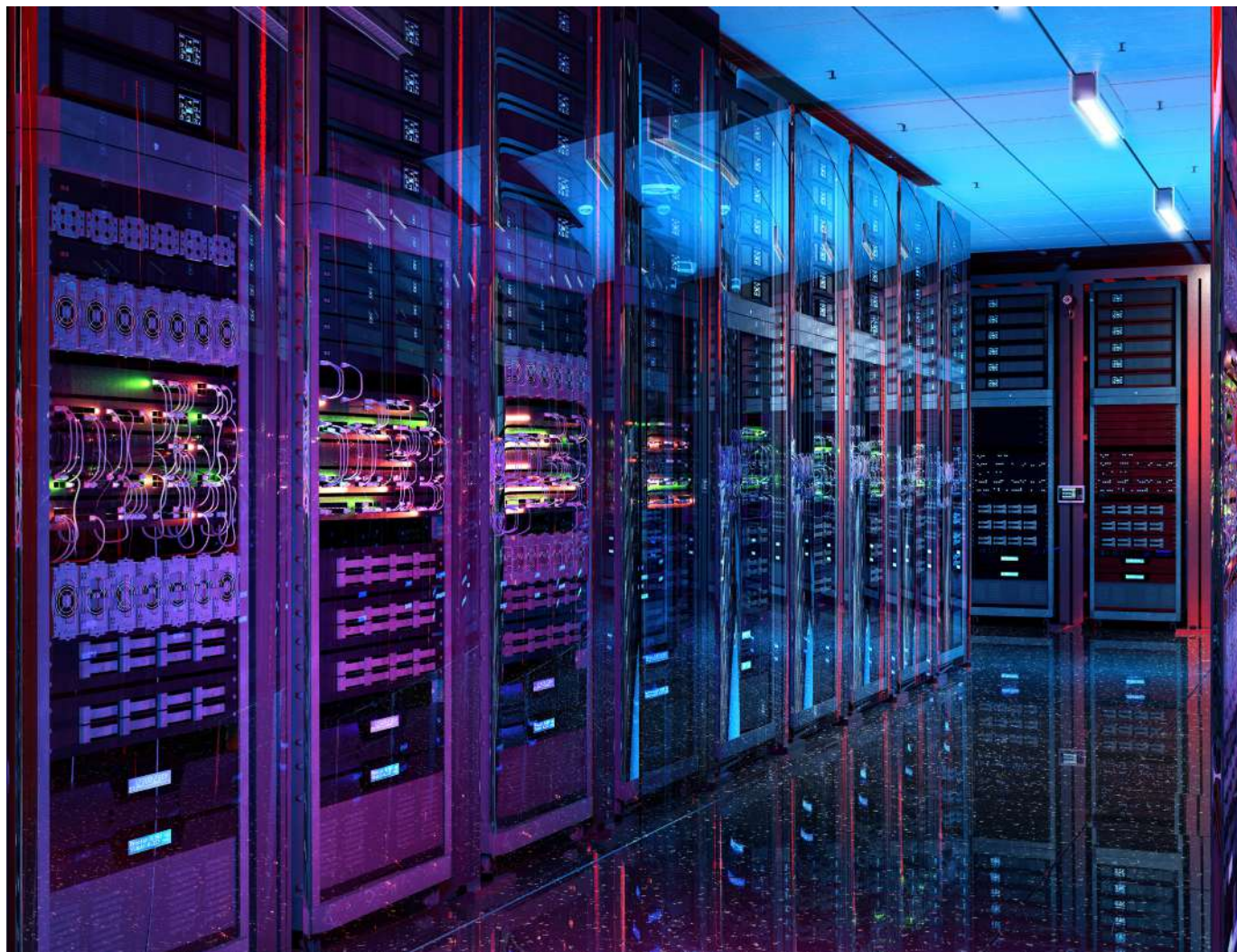
A clear AI strategy, aligned with business priorities, is the strongest predictor of success. Many organisations run local pilots without considering scalability, repeatedly building country-specific solutions based only on country data. This approach prevents data from being connected across countries and limits AI's ability to generate cross-country insights.

To effectively scale AI efforts, commercial leaders should:

- Align AI priorities with long-term business goals to ensure they address high-impact opportunities rather than short-term experimentation.
- Collaborate across functions – data, analytics, digital, and IT – to build a scalable AI roadmap with defined resources, timelines, and investments.
- Establish governance structures that support AI adoption at an enterprise level, ensuring consistency and alignment across regions.

2. Data and Analytics: Establishing Global Data Standards

Once a strategic direction is set, data and analytics teams can ensure access to high-quality, globally standardised, connected data. Piecing together country-specific data will make deploying initiatives across different markets challenging.



To overcome fragmentation, organisations should:

- Standardise data structures globally, ensuring that AI models trained in one region can be applied seamlessly worldwide.
- Invest in connectable data assets that integrate customer, sales, and engagement data across the organisation.
- Continuously refine data quality, ensuring AI models are built on accurate, harmonised data that supports enterprise-wide decision-making.

3. Digital and IT: Reducing Integration Complexity

Technology teams play a pivotal role in making AI scalable by reducing data friction, eliminating costly integrations, and breaking down data silos.

To support AI efforts, technology teams should:

- Align data models across systems to prevent inefficient data mapping and redundant integrations.
- Evaluate process inefficiencies such as third-party access (TPA) agreements that slow down data flow and require unnecessary administrative work.
- Implement scalable data governance frameworks that streamline AI deployment across multiple markets.

Your Data Defines AI's Possibilities

AI adoption in commercial biopharma is accelerating, increasing

the need for high-quality, connected data for more personalised engagement.

Approaching data standardisation with the same urgency as defining AI strategy and infrastructure is critical. After all, the real question isn't, "How can I use AI?" but "How can I make my data work for AI?"

REFERENCES

1. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/a-data-leaders-technical-guide-to-scaling-gen-ai>.
2. <https://www.veeva.com/eu/customer-stories/bayer-driving-insights-and-efficiencies-with-a-global-data-model/>



Karl Goossens

Karl Goossens is Director of OpenData Strategy at Veeva. Since joining the company in 2021, he has played a key role in building Veeva's Commercial Analytics capabilities in Europe, before shifting his focus to advancing the company's OpenData offering. Karl holds a Master's in Process Engineering from ETH Zurich and an MBA from IESE Business School in Barcelona.