AIDDD SUMMIT

INSIGHTS

THE INFRASTRUCTURE REALITY OF AI IN DRUG DISCOVERY

FROM PRODUCTION TO PROMISE

AI in drug discovery and development has reached a decisive phase. After years of experimentation, pilot programs, and ambitious claims, the industry is moving beyond whether AI can create value, toward how it can be deployed reliably and at scale.

Across the AI Drug Discovery & Development Summit 2026, one message surfaced repeatedly: the limiting factor for AI impact is no longer model capability, but infrastructure readiness.

While advances in machine learning, generative chemistry, and multimodal modelling continue, many organizations remain constrained by fragmented data estates, manual workflows, and platforms not designed for AI-native research.

Pharma and biotech organizations generate vast quantities of experimental data, but turning that data into impact remains unclear. AI does not fail because it lacks intelligence; it fails because the foundations beneath it are weak.

Expectations are rising sharply. Boards, investors, and R&D leaders expect AI to compress timelines, improve decision quality, and reduce attrition across discovery and development.

FREDDIE METHERELL

PORTFOLIO DIRECTOR

INFRASTRUCTURE IS NOW THE RATE-LIMITING STEP

A consistent theme across the AI Infra track was that AI capability itself is no longer the primary constraint in drug discovery—infrastructure is.

Experts on-stage described R&D environments where model performance continues to improve, yet real-world impact remains uneven due to fragmented data estates, lost experimental context, inconsistent metadata, and manual data movement.

There was little debate about whether better models were coming. The frustration in the room centred on the gap between what AI can do in theory and what existing systems can support in practice. In several discussions, infrastructure limitations were described not as technical debt, but as the primary reason promising work stalled.

Successful AI adoption was framed as a sequencing problem: core infrastructure first, followed by automation, and only then advanced analytics and AI models. Teams that reversed this order described paying the price later, often through rework, stalled deployment, or loss of confidence.

Subha Madhavan, Vice President & Head of Clinical AI / ML, Pfizer

Where these foundations were in place, organizations reported meaningful reductions in scientific analysis time and improved confidence in AI-supported decisions.

Infrastructure choices were repeatedly described as direct determinants of scientific productivity and trust in AI outputs. The underlying sentiment was clear: infrastructure is no longer a background concern — it is now a defining factor in whether AI delivers value at scale.

FROM PILOT TO PRODUCTION: WHAT WORKS IN PRACTICE

One large pharma organization described plans to generate imaging data at a scale of roughly 200 terabytes per day. By standardizing pipelines, automating image processing, and embedding AI into workflows designed for lab scientists, the organisation reported reductions of approximately 30% in scientific analysis time.

Another organization highlighted that in 2024, more than 85% of its small-molecule projects were AI-assisted following integration of AI tools into the Design–Make–Test–Analyze cycle.

There was an emphasis on operational readiness. Speakers were explicit that these gains did not come from isolated innovation, but from sustained investment in data foundations, automation, and integration.

AI became impactful once it was treated as part of the scientific workflow, rather than as an external capability layered on top.
There was also a clear contrast with earlier-stage efforts discussed elsewhere in the session.

In these cases, AI moved from pilot to production when infrastructure and workflow design were treated as scientific priorities, not deferred technical concerns. The sentiment in the room was that these examples were notable precisely because they were repeatable, rather than exceptional.

CASE STUDIES: FROM PILOT TO PRODUCTION

NOVO NORDISK
Facing plans to generate imaging data at a scale of approximately 200 terabytes per day, Novo Nordisk standardised data pipelines, automated image processing, and embedded AI into workflows designed for lab scientists. This resulted in approximately 30% reductions in scientific analysis time.
Takeaway: Infrastructure-led workflow design unlocked AI impact at scale.

ASTRAZENECA
AstraZeneca reported that more than 85% of small-molecule projects were AI-assisted in 2024, following integration of AI tools directly into the Design–Make–Test–Analyse cycle, supported by production-grade platforms and intuitive interfaces.
Takeaway: AI adoption accelerated once tools became part of everyday scientific workflows.

WHY AI PILOTS CAN FAIL AT SCALE

One of the strongest points of alignment across the session was not about model performance, but about frustration. Speaker after speaker returned to the same observation: AI pilots often work—until organizations try to scale them.

There was little doubt about model capability. Early experiments routinely deliver value in controlled settings. But there was broad recognition that these successes are fragile. What works for a single team or dataset often breaks when exposed to production realities.

Several speakers described how scaling quickly surfaces issues that are easy to ignore during pilots. Pipelines that rely on manual steps become brittle. Data quality problems that were acceptable in exploratory work become blockers once outputs influence real decisions. Latency, reliability, and traceability suddenly matter — and many systems are not built for them.

Governance emerged as a clear point of tension. More than one speaker acknowledged that guardrails are often added too late, after tools are already in use. This creates friction precisely when confidence is needed most, particularly in regulated R&D environments where auditability and reproducibility are essential.

Ownership also becomes blurred at scale. AI initiatives often sit between scientific teams, data groups, and IT. When pilots move beyond experimentation, no single group is clearly accountable, and progress slows as a result.

Workflow fit was another recurring theme. Several examples highlighted pilots that failed to gain traction because they sat outside core scientific processes. Scientists were asked to change how they work to accommodate new tools, rather than having AI embedded into existing workflows. Even strong models struggle to survive this friction.

Lifecycle management drew quieter but serious concern. As models proliferate, tracking versions, monitoring performance, and managing drift becomes harder. Without robust MLOps, trust degrades over time, even when early results are strong.

The sentiment was pragmatic rather than pessimistic. The issue is not that organizations are moving too fast, but that they are scaling without foundations. Pilots show what is possible. Production reveals what is sustainable.

ADOPTION IS A WORKFLOW DESIGN PROBLEM

Discussions around adoption were notably pragmatic, there was little ideological resistance to AI in the room. Instead, conversations repeatedly returned to a simpler point: scientists adopt AI when it fits how they already work — and abandon it when it does not.

Several speakers emphasised that model performance alone is rarely the deciding factor. Early enthusiasm often fades when tools add steps, require manual configuration, or interrupt established workflows. Even high-quality outputs struggle to gain traction if they arrive too late, or in a form that is hard to act on.

Where adoption was working, the pattern was consistent. AI was embedded directly into existing scientific processes. Data flowed automatically from instruments into analysis and downstream tools. Defaults were sensible. Context was preserved. Scientists did not need to make an explicit decision to “use AI” — it was already there, supporting routine decisions.

In contrast, slower adoption was associated with tools that sat outside core workflows. Several examples highlighted systems that required specialist support, bespoke configuration, or separate interfaces. In these cases, AI remained the domain of a small group of power users rather than becoming a shared capability.

This distinction became especially clear when discussion turned to user experience. Experts were clear that usability is not a cosmetic concern. Poorly designed interfaces concentrate capability in technical teams, while well-designed platforms distribute it across the organization. This distinction became more pronounced as AI usage expanded.

Trust also featured prominently in discussions of adoption. Scientists were more willing to rely on AI outputs when results were reproducible, traceable, and clearly situated within the context of their work. Workflow integration played an important role here, preserving provenance and making decision paths visible.

The underlying sentiment was clear. Adoption is not primarily a training problem, nor a cultural one. It is a design problem. Organizations that design AI into workflows see sustained use and reap the benefits.

WHAT LEADING TEAMS ARE DOING DIFFERENTLY

When the discussion turned to organizations making real progress, no one claimed access to better models. What distinguished leading teams was how deliberately they had reshaped the environment around AI.

Several speakers described a move away from treating AI as a collection of experiments. In more mature organizations, AI was discussed as an operating capability. Ownership sat at the platform or portfolio level rather than with individual project teams. This allowed infrastructure, data pipelines, and governance to be built once and reused, rather than reinvented for each use case.

There was clear alignment around sequencing. Teams that were scaling successfully had resisted the urge to lead with advanced models. Instead, they invested first in data foundations, automation, and integration. AI was layered onto stable systems, not used to compensate for gaps in them.

Design choices came up repeatedly. Leading teams assumed most scientists would never configure models or write code. Platforms were built to hide complexity, with sensible defaults and tight workflow integration. This was less about democratizing AI in principle, and more about removing friction in practice.

Trust was another differentiator. Rather than relying on assurances, mature teams had embedded reproducibility, traceability, and monitoring into their systems from the outset. Governance was described as operational rather than performative. This appeared to be a prerequisite for sustained adoption in regulated R&D.

Planning for scale also looked different. Hybrid compute was treated as the default. Model lifecycle management was assumed to be complex and long-lived. Several implied that teams struggling to scale had underestimated this operational burden.

Perhaps most striking was how selective leading teams were. Use cases were narrow and clearly owned. Success was framed in terms of improved scientific decisions and workflow impact, not benchmark scores.

The sentiment was not that these teams were moving faster, but that they were moving with more intent. Cautious teams focused on what AI could do. Leading teams focused on what had to change for AI to work—and acted accordingly.

GENERATIVE AI: PROMISE, WITH CONSTRAINTS

Generative AI drew strong interest, but the tone in the room was noticeably more measured than in previous years. Discussion moved quickly away from novelty toward reliability.

Speakers were clear that the most credible near-term role for generative and agentic AI is as an interface layer, helping scientists interact more easily with complex tools, models, and data. Fully autonomous decision-making was widely viewed as out of scope for regulated R&D.

Non-deterministic behaviour was repeatedly cited as a limiting factor. Small changes in prompts can produce different outcomes, making reproducibility difficult to guarantee. This lack of consistency remains a major barrier to trust in discovery workflows that depend on traceability and auditability.

Many described early experiments with agentic systems that select and orchestrate tools automatically. While promising in theory, these approaches were characterised as fragile in practice. Ensuring the right tools are invoked, in the right order, with consistent results is not yet reliable enough for broad deployment.

The prevailing sentiment was pragmatic rather than dismissive. Generative AI is expected to accelerate discovery but not replace scientific judgement. Its value lies in augmentation, not autonomy.

GENERATIVE AI IN R&D: DO / DON'T

DO	DON'T
• Lower access barriers to tools and data • Constrain GenAI to bounded tasks • Embed GenAI within governed workflows • Prioritise traceability and auditability • Use GenAI as a productivity accelerator	• Rely on autonomous decision-making • Assume deterministic behaviour • Deploy without ownership or rollback plans • Introduce parallel, disconnected interfaces • Confuse impressive demos with readiness

Takeaway: GenAI delivers value when constrained, governed, and embedded into existing systems.

WHAT CHANGES IN 2026

Looking ahead, the focus was less about prediction and more about transition. Speakers did not frame 2026 as a year of breakthrough models, but as a point where expectations around AI deployment become harder to avoid.

There was broad agreement that AI maturity will increasingly be judged by execution rather than ambition. Claims of AI readiness will be tested against infrastructure, integration, and reliability, not the number of pilots or proofs of concept underway.

Several speakers implied that informal approaches will begin to fail. As models proliferate across discovery workflows, MLOps, monitoring, and lifecycle management move from optional to unavoidable. Teams that have delayed these investments are likely to feel increasing friction.

Ownership was another area where change felt imminent. AI is expected to move further into core R&D operations, away from isolated innovation groups. Responsibility for outcomes shifts closer to scientific leadership, rather than remaining with specialist or experimental teams.

Expectations around generative AI also appeared to be changing. The focus is moving away from autonomy and toward controlled, reliable use cases. In 2026, GenAI is more likely to be judged on where it is constrained than where it is deployed.

Finally, several speakers hinted at a compounding effect. Organizations that have invested early in data foundations, workflow integration, and platforms are accelerating faster as AI usage grows. Those that have not may find the gap increasingly difficult to close.

The sentiment was not one of urgency for its own sake, but of inevitability. By 2026, AI in drug discovery is no longer judged by promise. It is judged by how well it works in practice.