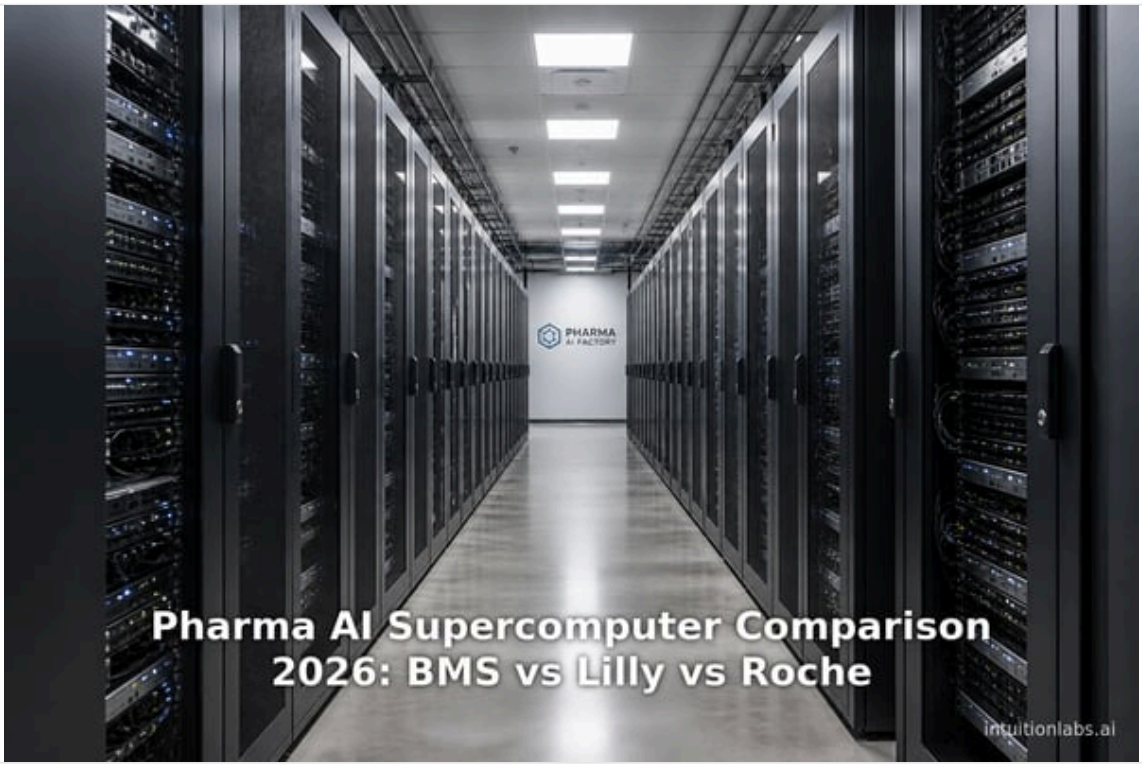
Pharma AI Supercomputer Comparison 2026: BMS vs Lilly vs Roche

Published July 31, 2026 37 min read



Pharma AI Supercomputer Comparison 2026: BMS vs Lilly vs Roche

Executive Summary

By July 2026, three of the world's largest pharmaceutical companies, **Bristol Myers Squibb** (BMS), **Eli Lilly**, and **Roche** (through its Genentech subsidiary), had made competing leadership claims about AI-computing infrastructure within a nine-month window. Their announcements describe different states of deployment: Roche said it operated a hybrid-cloud AI factory, Lilly announced LillyPod, and BMS announced that it **would deploy** a future NVIDIA DGX SuperPOD built on Vera Rubin systems. Eli Lilly moved first, announcing on October 28, 2025 a partnership with **NVIDIA** to build what it called the most powerful supercomputer owned and operated by a pharmaceutical company (Source: investor.lilly.com). That system, nicknamed **LillyPod**, went live at a ribbon-cutting in Indianapolis on February 26, 2026, running on more than **1,000 NVIDIA Blackwell Ultra GPUs** on a single unified networking fabric (Source: investor.lilly.com), a build NVIDIA's own account puts more precisely at **1,016 GPUs** delivering more than **9,000 petaflops** of AI performance, detailed further in the Eli Lilly section below. Roche countered on March 16, 2026, adding **2,176 new on-premise Blackwell GPUs** to bring its combined on-premise and cloud footprint to more than **3,500 Blackwell GPUs**, which it and NVIDIA describe as the largest announced GPU footprint of any pharmaceutical company (Source: roche.com). BMS answered four months later, announcing on July 20, 2026 that it would deploy an NVIDIA DGX SuperPOD built on **DGX Vera Rubin NVL72** systems, the next-generation architecture succeeding Blackwell, calling it the most powerful and energy-efficient single-owned NVIDIA infrastructure in life sciences (Source: news.bms.com).

The three claims are not directly comparable because each measures a different thing. Lilly's figure describes hardware it owns outright at a single Indianapolis site. Roche's figure blends on-premise and cloud capacity across the United States and Europe. BMS has not disclosed a GPU count at all; independent analysis of NVIDIA's own blog post, which describes eight DGX Vera Rubin NVL72 racks, implies roughly **576 GPUs**, a number BMS itself has not confirmed (Source: drugdiscoverytrends.com). None of the three companies has disclosed the dollar cost of its flagship system. BMS says its planned SuperPOD will use **Vera Rubin**, for which it claims up to 10 times greater performance per megawatt than its predecessor; LillyPod uses **Blackwell Ultra** B300 GPUs, and Roche's AI factory uses **Blackwell** GPUs. Roche describes its NVIDIA collaboration as starting in 2023, while BMS says its expansion builds on nearly three years of collaboration.

This report compares the three companies' AI supercomputer investments in detail: their announced hardware, software ecosystems (NVIDIA BioNeMo, Omniverse, Parabricks, and NVIDIA FLARE federated learning all recur across vendors), stated business rationale, and the limits of what each company has actually disclosed. It sets those investments against the broader pharmaceutical AI-infrastructure landscape, including Amgen's 248-GPU Freyja SuperPOD in Iceland (Source: blogs.nvidia.com), Novo Nordisk's use of Denmark's 1,528-GPU Gefion supercomputer, ranked 29th on the Top500 list of the world's fastest supercomputers (Source: hpcwire.com), AstraZeneca's stake in a Swedish sovereign AI factory (Source: wallenberginvestments.com), and Merck's roughly $1 billion agentic-AI partnership with Google Cloud, which is a software and cloud-services deal rather than an owned GPU cluster (Source: merck.com). BMS has also struck a parallel deal giving its 30,000-plus workforce access to Anthropic's Claude model, a reminder that GPU procurement is only one piece of a broader, multi-vendor AI strategy (Source: qz.com).

On the quantitative side, the report traces GPU pricing (new NVIDIA H100 units selling for $25,000 to $40,000 and GB200 NVL72 racks listing near $2 million to $3 million as of mid-2026) (Source: gpusmith.com), NVIDIA's record $215.9 billion in fiscal 2026 revenue (Source: investor.nvidia.com), Grand View Research's $2.3 billion valuation of the global AI-in-drug-discovery market in 2025 with a projected 24.8% compound annual growth rate through 2033 (Source: grandviewresearch.com), and Epoch AI's finding that frontier AI training-run costs have grown 2.4 times per year since 2016, with the largest runs projected to exceed $1 billion by 2027 (Source: epoch.ai). IQVIA's Institute for Human Data Science separately found a 75% Phase I success rate for AI-enabled programs among emerging biopharma companies, an early but credible signal that AI-assisted discovery correlates with better clinical outcomes (Source: iqvia.com), while PhRMA notes its member companies have invested more than $850 billion in new treatments over the past decade, a sum that dwarfs any individual GPU cluster investment profiled here (Source: phrma.org). No pharma-specific, audited capital-expenditure total for these AI supercomputer buildouts exists in the public record; every figure above is either a hardware-spec estimate or an economy-wide proxy, a gap this report addresses directly rather than papering over.

The report closes with five named case studies (BMS's CELMoD compound library, Genentech's AI-accelerated backup drug candidate, Recursion Pharmaceuticals' 504-GPU BioHive-2, ranked 35th on the TOP500 list and the only system in this category independently benchmarked by a third party (Source: ir.recursion.com), Insilico Medicine's Phase III AI-discovered drug, and Isomorphic Labs' Novartis collaboration and an assessment of what the compute arms race means for pharmaceutical IT strategy, procurement, and organizational readiness through 2027.

Introduction and Background

Pharmaceutical research and development has always been compute-intensive, but the nature of that computation changed sharply between 2023 and 2026. Where earlier generations of computational chemistry and bioinformatics ran on modest clusters optimized for molecular dynamics or genomic alignment, the current wave of investment is explicitly built around training and running large **foundation models**, AI systems pretrained on vast biological, chemical, and clinical datasets that can then be adapted to specific discovery tasks. NVIDIA's healthcare and life sciences division has positioned itself as the primary infrastructure supplier for this shift, framing pharmaceutical R&D's roughly $300 billion annual global cost base as the addressable opportunity for its BioNeMo platform and DGX hardware line (Source: nvidianews.nvidia.com). NVIDIA CEO **Jensen Huang** has framed the underlying shift in sweeping terms, arguing the **"greatest impact of generative AI is to revolutionize the life science and healthcare industry"** (Source: gene.com).

Industry-wide, the numbers back the scale of the shift. **IQVIA's Institute for Human Data Science** reported that global biopharma funding reached a 10-year high of $102 billion in 2024, a substantial increase over the 2023 figure of $71 billion (Source: iqvia.com). The trade association PhRMA notes its member companies have invested more than $850 billion in developing new treatments over the past decade (Source: phrma.org). The cost of the flagship AI systems profiled here has not been disclosed.

IntuitionLabs, a life sciences and AI consultancy, has tracked this shift closely in its own research, noting that AI-enhanced drug discovery and development can accelerate timelines by up to 60%, according to research it cites from Deloitte, and that McKinsey has estimated AI could generate more than $100 billion in annual value for the pharmaceutical industry, with adoption since 2024 tracking ahead of that curve (Source: intuitionlabs.ai). In a dedicated report on 2026 infrastructure deals, IntuitionLabs' analysts observed that Lilly's LillyPod build made the company "one of the first and most powerful adopters of AI supercomputing in drug R&D," a framing that this report's three-way comparison complicates: by mid-2026, Lilly was one of at least three pharmaceutical companies making that claim simultaneously (Source: intuitionlabs.ai). A separate IntuitionLabs analysis of the same Lilly deal argued the underlying pattern extends well beyond pharma: **"major companies are building their own AI infrastructure rather than relying solely on packaged"** consumer AI applications or cloud chat services (Source: intuitionlabs.ai).

What makes 2026 distinctive is not merely the scale of spending but the compression of the announcement cycle. Lilly's announcement in October 2025 was followed by Roche's in March 2026, less than three weeks after Lilly's LillyPod formally went live, and BMS's in July 2026, four months after Roche's (Source: rdworldonline.com). Trade press has explicitly framed this as a competitive dynamic: STAT News observed that "for the third time in nine months, a pharma company has announced that it is assembling the largest AI supercomputer in the life sciences industry," referring to BMS's

July 2026 announcement following Lilly's and Roche's (Source: statnews.com). Even participants in the buildout have voiced skepticism about the framing: Lilly's own Chief Information and Digital Officer, Diogo Rau, was quoted warning that "the hype is actually a serious threat to the research itself," a caution this report takes seriously by separating vendor superlatives from disclosed, verifiable specifications wherever the two diverge (Source: rdworldonline.com).

Bristol Myers Squibb

Capabilities

BMS announced on July 20, 2026 that it would expand its NVIDIA compute infrastructure with a second **DGX SuperPOD**, this one built on **DGX Vera Rubin NVL72** systems, NVIDIA's newest rack-scale architecture and the direct successor to Blackwell (Source: news.bms.com). NVIDIA's own blog post describing the deal states the build comprises **eight DGX Vera Rubin NVL72 systems** (Source: blogs.nvidia.com). Each Vera Rubin NVL72 rack pairs 72 Rubin GPUs with 36 Vera CPUs into a single coherent machine (Source: drugdiscoverytrends.com), which by multiplication implies roughly 576 GPUs across the full build, a total neither BMS nor NVIDIA has stated directly in its own materials. BMS describes the resulting system as **"the most powerful and energy-efficient single-owned NVIDIA infrastructure in life sciences"** and says the Vera Rubin architecture delivers up to **ten times greater performance per megawatt** than the infrastructure it replaces (Source: news.bms.com). Reuters reported that BMS will be the first life sciences company to buy a DGX SuperPOD based on Vera Rubin, noting the prior BMS system it replaces sits roughly two to three GPU generations behind (Source: reuters.com).

The new build extends, rather than replaces from scratch, a **nearly three-year-old** BMS-NVIDIA relationship: BMS's original DGX SuperPOD, deployed with partners Equinix and Mark III Systems, is credited in an NVIDIA case study with delivering **55% overall cost savings compared to BMS's prior computing model** (Source: nvidia.com). BMS Vice President of Research Business Insights and Technology Erin Davis, who nicknamed the combined environment the "SuperDuperPOD," said the original SuperPOD had become saturated, driving the need for expansion: **"We're building our own foundational models, and that takes a lot of GPUs"** (Source: blogs.nvidia.com). The two SuperPODs, old and new, are being combined into a single unified environment, which BMS describes as a single data plane accessible from every BMS site globally.

Adoption

BMS's Chief Digital and Technology Officer, **Greg Meyers**, told STAT News the original cluster's capacity had simply run out: **"We actually consumed all the space we had"** (Source: statnews.com), and told Reuters the calculation behind expansion was straightforward: energy costs scale with usage regardless of GPU efficiency gains, so **"think of it as 10 times more compute capacity per watt spent. Electricity is not getting cheaper"** (paraphrased from Meyers's Reuters interview) (Source: qz.com). BMS states that AI now **informs the design of every small molecule program and the majority of the company's large molecule programs**, spanning oncology, hematology, cardiovascular disease, immunology, and neuroscience (Source: news.bms.com). BMS Chief Research Officer **Robert Plenge** said the added compute lets BMS screen far more candidates before committing to expensive clinical development: **"Maybe before we could do 10 and now we can do dozens"** (Source: reuters.com), and separately stated that AI tools have already cut the time to produce candidate medicines ready for clinical testing by **20% to 30%**, with the potential to reach 50% in coming years (Source: reuters.com). Separately from its NVIDIA hardware investment, BMS struck a deal with Anthropic to give its workforce of more than 30,000 employees access to the Claude AI model across drug discovery, manufacturing, and commercial operations, illustrating that the supercomputer buildout is one piece of a broader, multi-vendor AI strategy rather than the entirety of it.

Strengths and Limitations

BMS's central strength is organizational maturity: unlike Lilly and Roche, which frame their 2025 to 2026 announcements substantially around new infrastructure, BMS is scaling a system already integrated into live research workflows. Its NVIDIA case study cites concrete prior results, including foundation models trained on hundreds of thousands of clinical trial images using NVIDIA MONAI for self-supervised learning (Source: nvidia.com) and an expanded library of CELMoD compounds, molecules that selectively degrade cancer-causing proteins, generated with AI assistance (Source: blogs.nvidia.com). The chief limitation is transparency: BMS is the only one of the three companies that has not disclosed a GPU count, a performance figure in petaflops or exaflops, or a dollar investment figure for either its original or expanded system, leaving outside observers reliant on NVIDIA's rack count and third-party arithmetic (Source: drugdiscoverytrends.com). NVIDIA's healthcare business-development director, Rory Kelleher, framed the partnership around BMS's scientific depth rather than its hardware scale, noting BMS "has built decades of extraordinary scientific knowledge across some of the most complex areas of human disease" (Source: news.bms.com), a framing that implicitly concedes the hardware comparison is not BMS's strongest selling point.

Eli Lilly

Capabilities

Eli Lilly's **LillyPod** is, by NVIDIA's own description, the world's first NVIDIA DGX SuperPOD built with DGX B300 systems, powered by **1,016 NVIDIA Blackwell Ultra GPUs** and delivering **more than 9,000 petaflops** of AI performance, assembled in just four months (Source: blogs.nvidia.com). Lilly first announced the buildout at NVIDIA's GTC Washington, D.C. event on October 28, 2025, describing it as **"the most powerful supercomputer owned and operated by a pharmaceutical company"**, powered by more than 1,000 B300 GPUs on a single unified high-speed networking fabric linking GPUs, storage, and related systems (Source: investor.lilly.com). LillyPod was formally inaugurated at a ribbon-cutting event in Indianapolis on February 26, 2026 and is now in production-scale use (Source: blogs.nvidia.com). The system enables Lilly's genomics team to process **700 terabytes of data** using more than **290 terabytes of high-bandwidth GPU memory**, connected via nearly 5,000 fiber-cable connections (Source: blogs.nvidia.com).

Separately from LillyPod itself, NVIDIA and Lilly announced a **five-year, up-to-$1-billion AI co-innovation lab** in the San Francisco Bay Area at the J.P. Morgan Healthcare Conference on January 12, 2026, explicitly built on the **NVIDIA BioNeMo platform** and the **Vera Rubin architecture**, positioning Lilly to be an early Vera Rubin adopter even as LillyPod itself runs on the prior-generation Blackwell Ultra chips (Source: nvidianews.nvidia.com). About four months after LillyPod's launch, Lilly Chief AI Officer **Thomas Fuchs** said Lilly had become the first company in its industry to deploy NVIDIA's roughly 550-billion-parameter **Nemotron 3 Ultra** open-weight model on-premises, tuned for its B300 GPUs at FP4 precision: **"Optimized for our B300 GPUs at FP4, this model is blazingly fast"** (Source: rdworldonline.com). By R&D World's analysis, an eight-GPU DGX B300 node is rated at 144 petaFLOPS of FP4 inference versus 72 petaFLOPS of FP8 training, roughly 2.25 times the 32 petaFLOPS of the prior Hopper-generation DGX H100 (Source: rdworldonline.com).

Adoption

Lilly Chief AI Officer Thomas Fuchs said the AI factory would train large-scale protein diffusion models, small-molecule graph neural network models, and genomics foundation models, calling large-scale computation **"not something optional for a company like ours"** (Source: blogs.nvidia.com). Fuchs described the pace of Lilly's research-computing scale-up since he joined in 2024 as moving from roughly eight GPUs, to 100, and now to the full 1,016-GPU supercomputer (Source: rdworldonline.com). Lilly frames LillyPod around breaking the physical limits of the wet lab, where a productive research team can typically test only about 2,000 molecular ideas per target per year; the company argues its dry-lab infrastructure removes that ceiling entirely. Separately, Lilly's **TuneLab** platform gives biotech partner companies access to drug discovery models built on more than $1 billion worth of Lilly's proprietary data, using federated learning built on NVIDIA FLARE so partner data never leaves their own infrastructure (Source: blogs.nvidia.com). CNBC reported that Lilly executive Diogo Rau cautioned patience is required before the investment pays off in launched medicines, saying **"we're really going to see those benefits in 2030"** (Source: cnbc.com).

Strengths and Limitations

Lilly's strength is specificity and speed of disclosure: unlike BMS, Lilly published a precise GPU count, a petaflops figure, and a build timeline (four months from start to assembly), and its Chief AI Officer has given detailed, on-the-record updates on production workloads six months post-launch, an unusual level of operational transparency in this category. NVIDIA healthcare VP **Kimberly Powell** characterized the broader trend Lilly exemplifies: **"Modern AI factories are becoming the new instrument of science"** (Source: investor.lilly.com). The chief limitation is that LillyPod runs on Blackwell Ultra rather than the newer Vera Rubin architecture: BMS's planned system uses Vera Rubin, whereas Roche's announced AI factory uses Blackwell GPUs. Lilly has also committed to running its AI supercomputing infrastructure on 100% renewable electricity by 2030, using existing chilled-water liquid cooling infrastructure at its Indianapolis facilities, a sustainability target that remains four years from completion (Source: investor.lilly.com).

Roche

Capabilities

Roche, operating primarily through its Genentech biotechnology unit, announced on March 16, 2026 that it had added **2,176 new on-premise NVIDIA Blackwell GPUs**, bringing its combined on-premise and cloud infrastructure to more than **3,500 Blackwell GPUs** in total, a figure the company and NVIDIA describe as the largest announced GPU footprint of any pharmaceutical company (Source: roche.com). Data Center Dynamics, reporting on the same announcement, described Roche as the world's fifth-largest pharmaceutical company by size, deploying that GPU footprint across data

centers in both the United States and Europe (Source: datacenterdynamics.com). The March 2026 expansion is explicitly framed as the next phase of a Roche-NVIDIA collaboration that began in 2023 (Source: roche.com), tracing back to a multi-year strategic research collaboration Genentech and NVIDIA announced on November 21, 2023, using NVIDIA DGX Cloud and BioNeMo to accelerate Genentech's machine learning models without disclosing a GPU count at the time (Source: gene.com). The 2026 infrastructure runs across four named NVIDIA software platforms: BioNeMo for R&D and its "Lab-in-the-Loop" workflow, Omniverse for manufacturing digital twins, Parabricks for diagnostics and genomics, and NeMo Guardrails for digital health conversational AI, according to Roche's own announcement.

Adoption

Roche Chief Digital and Technology Officer **Wafaa Mamilli**, who joined Roche in February 2025 after more than 20 years at Eli Lilly and a stint at Zoetis (Source: roche.com), framed the AI factory around clinical urgency: **"time is the most critical variable; every day saved means a life-changing medicine"** reaches patients sooner (Source: roche.com). Genentech's head of Research and Early Development, **Aviv Regev**, said the expanded compute would let scientists build more sophisticated predictive models that further shorten the path from biological insight to medicine, building on Roche's five-year-old "Lab-in-the-Loop" R&D strategy (Source: roche.com). NVIDIA's own coverage of the deal reports that **nearly 90% of Genentech's eligible small-molecule programs now integrate AI**, and cites one specific example of an oncology protein-degrader molecule designed 25% faster than by conventional methods, along with a backup drug candidate delivered in seven months instead of the usual two-plus years (Source: blogs.nvidia.com). Roche is also using NVIDIA Omniverse to build digital twins of its manufacturing sites, including its new GLP-1 manufacturing facility in North Carolina (Source: blogs.nvidia.com).

Strengths and Limitations

Roche's headline strength is scale: at more than 3,500 Blackwell GPUs it claims a larger combined footprint than either Lilly's 1,016-GPU LillyPod or BMS's roughly 576-GPU Vera Rubin build. But R&D World's direct comparison of the three deals cautions that the totals are not measuring the same thing: **"Roche's claim counts hybrid cloud capacity; Lilly's claim is limited to hardware"** the company owns outright (Source: rdworldonline.com). The same outlet also noted that Roche's March 2026 release, unlike BMS's and Lilly's, disclosed **no dollar figure, named disease areas, or specific pipeline programs** tied to the investment, leaving its strategic rationale comparatively underspecified (Source: rdworldonline.com). Roche also has not published a petaflops or exaflops performance figure for its cluster, the only one of the three companies to disclose GPU count without an accompanying compute-throughput claim. Trade outlet pharmaphorum situated the March 2026 announcement within the wider NVIDIA-pharma alliance landscape, noting it followed Lilly's $1 billion co-innovation lab commitment and a 2025 Novo Nordisk-NVIDIA deal (Source: pharmaphorum.com).

Feature Comparison

Table 1 below summarizes the disclosed hardware, timing, and disclosure posture of the three companies' flagship 2025 to 2026 AI supercomputer announcements side by side. Citations for figures already established in the sections above are not repeated in every cell; see the corresponding company section for sourcing.

FEATURE	BRISTOL MYERS SQUIBB	ELI LILLY	ROCHE
System name	Unnamed; “SuperDuperPOD” refers to the combined old-and-new environment	LillyPod	Hybrid-cloud “AI factory” (unnamed)
GPU architecture	NVIDIA Rubin (Vera Rubin NVL72)	NVIDIA Blackwell Ultra (B300)	NVIDIA Blackwell
Disclosed GPU count	Not disclosed; independently calculated at **576** (8 racks x 72) (Source: drugdiscoverytrends.com)	**1,016** Blackwell Ultra GPUs	More than **3,500** (combined on-prem and cloud; 2,176 newly added) (Source: roche.com)
Performance claim	Up to 10x performance-per-megawatt vs. predecessor (no absolute petaflops figure)	More than 9,000 petaflops AI performance	Not disclosed
Announcement date	July 20, 2026	October 28, 2025	March 16, 2026
Live / operational date	Not yet operational at announcement (deployment underway)	Inaugurated February 26, 2026	Operational at announcement (expansion of live system)
Ownership model	Single-owned (deployment location not disclosed)	Single-owned, on-premise (Indianapolis)	Hybrid on-premise plus cloud, US and Europe (Source: datacenterdynamics.com)
Investment disclosed	Not disclosed (Source: reuters.com)	$1B (separate five-year co-innovation lab; LillyPod itself undisclosed) (Source: nvidianews.nvidia.com)	Not disclosed
Key software stack	NVIDIA BioNeMo Agent Toolkit	BioNeMo, MONAI, NVIDIA FLARE (TuneLab)	BioNeMo, Omniverse, Parabricks, NeMo Guardrails

As the table makes clear, direct ranking by "power" depends entirely on which metric is chosen. By raw GPU count, Roche's hybrid footprint leads; by disclosed compute throughput, Lilly's more than 9,000 petaflops is the only hard figure on the table; BMS's claimed 10x performance-per-megawatt improvement describes efficiency rather than absolute capacity and cannot be numerically compared with Lilly's throughput claim. None of the three companies has disclosed capital expenditure for its cluster, and the only specific dollar figure is Lilly's separate five-year, up-to-$1-billion co-innovation-lab commitment rather than a LillyPod cost. Readers should treat "largest," "most powerful," and "most advanced" as company-selected superlatives tied to company-selected metrics, not as outcomes of an independent benchmark.

Performance and Benchmarks

No independent, standardized benchmark (comparable to the academic TOP500 supercomputer ranking) has been published for any of the three companies' 2026 systems using real pharmaceutical workloads. All performance figures currently in circulation are either vendor-disclosed peak theoretical throughput or third-party arithmetic derived from published rack and GPU counts. Drug Discovery & Development's own analysis is explicit about this limitation, estimating BMS's 576-GPU Vera Rubin system at roughly **10.1 exaflops of peak dense FP8 training performance** with 166 terabytes of aggregate GPU memory, against Lilly's live 1,016-GPU Blackwell Ultra system at an estimated 4.6 exaflops and 293 terabytes, while cautioning: **"These are peak reference-spec estimates. Measured performance on pharmaceutical workloads remains undisclosed"** (Source: drugdiscoverytrends.com). The gap between that 4.6-exaflop estimate and Lilly's own claim of more than 9,000 petaflops (9 exaflops) of AI

performance illustrates a recurring measurement problem in this category: figures depend heavily on which numeric precision (FP4, FP8, or a mixed-precision blend) and which workload type (training versus inference) is being measured, and vendors are not always explicit about which one they are quoting.

At the underlying chip level, NVIDIA's Rubin GPU, the core of the Vera Rubin platform BMS plans to deploy, packs **336 billion transistors**, 224 streaming multiprocessors, and 896 Tensor Cores, versus 208 billion transistors on the Blackwell-generation B200 chip (Source: developer.nvidia.com) (Source: nvidia.com). The B200 is a Blackwell-family comparison point, not the specific product deployed in LillyPod, which uses B300 Blackwell Ultra GPUs. Rubin integrates up to **288 gigabytes of HBM4 memory** delivering up to 22 terabytes per second of bandwidth, a 2.8 times increase over Blackwell and Blackwell Ultra memory bandwidth (Source: developer.nvidia.com). NVIDIA itself claims Vera Rubin NVL72 racks can complete comparable AI training workloads with roughly one-quarter the GPUs and deliver inference at one-tenth the cost per million tokens compared with Blackwell, though this is a vendor-reported figure rather than an independently audited benchmark (Source: nvidia.com). NVIDIA describes Vera Rubin as having entered full production for large-scale deployment in 2026, the same window in which Roche made its own GPU-footprint announcement and BMS confirmed its own architecture choice.

Data Analysis and Evidence

The financial and market context behind these buildouts is easier to document at the industry level than at the level of any single pharmaceutical company's capital budget. NVIDIA's Data Center segment, which supplies accelerated-computing products used in the BMS and Lilly DGX SuperPOD announcements and Roche's hybrid-cloud AI factory, posted record quarterly revenue of **$62.3 billion** in the fourth quarter of fiscal 2026, up 22% from the prior quarter, contributing to full fiscal-year 2026 revenue of **$215.9 billion**, up 65% year over year (Source: investor.nvidia.com). NVIDIA does not break out a dedicated healthcare or pharmaceutical revenue line in its SEC filings; the closest available figure is a JPMorgan analyst estimate, reported by a secondary financial outlet, that NVIDIA's healthcare vertical had already become a **$1 billion-plus annual business** in fiscal 2024, two to three years ahead of the company's internal target (Source: dailybostonjournal.com), a figure this report flags as an analyst estimate rather than a company disclosure.

On GPU hardware pricing, aggregated market data as of mid-2026 puts new **NVIDIA H100** units at **$25,000 to $40,000** per unit (Source: gpusmith.com), with **B200** (Blackwell) units selling near **$30,000 to $40,000** per GPU against an estimated manufacturing cost of roughly $6,400 (Source: gpusmith.com) (Source: siliconanalysts.com), and full **GB200 NVL72** rack-scale systems, which link 72 Blackwell GPUs into a single machine, listing at roughly **$2 million to $3 million per rack** (Source: gpusmith.com). BMS has not disclosed a budget for its planned Vera Rubin system; GB200 NVL72 list-price estimates are not a reliable basis for estimating the cost of newer Vera Rubin hardware.

On the cost of training AI models generally (not pharma-specific, since no company in this report has disclosed a training-run cost), the nonprofit research group **Epoch AI** finds that the amortized hardware-and-energy cost of frontier AI training runs has grown **2.4 times per year since 2016** and projects the largest training runs will exceed **$1 billion** in cost by 2027 (Source: epoch.ai) (Source: epoch.ai). Because none of BMS, Lilly, or Roche has published what it costs to train its own proprietary foundation models on LillyPod, the BMS Vera Rubin cluster, or Roche's hybrid-cloud fleet, this Epoch AI trend line is the best available proxy for the order of magnitude such training runs likely cost industry-wide, and readers should not treat it as a pharma-specific figure.

Market-sizing research from **Grand View Research** values the global AI-in-drug-discovery market at **$2.3 billion in 2025**, projecting a **24.8% compound annual growth rate** from 2026 through 2033 that would take the market to $13.8 billion (Source: grandviewresearch.com) (Source: grandviewresearch.com). That figure describes AI software and services spending in drug discovery specifically, a narrower category than the

hardware infrastructure spending discussed throughout this report; Gartner's broader, economy-wide estimate puts worldwide AI spending across all industries and categories at **$2.52 trillion in 2026**, a 44% year-over-year increase, of which roughly $1.37 trillion is AI infrastructure spending (Source: gartner.com). A narrower, healthcare-specific estimate compiled by an industry research blog from Gartner and other data puts 2026 healthcare-sector enterprise AI spending at roughly **$45 billion** with 62% adoption, a figure that spans providers and payers alongside pharmaceutical companies and is therefore not directly comparable to the drug-discovery-specific figures above (Source: valueaddvc.com). No single authoritative, pharma-industry-specific capital-expenditure total for 2026 AI infrastructure exists in the public record as of this report's publication; the figures above are the closest available bracketing estimates, and this report treats the absence of a unified figure as a genuine data gap rather than filling it with an unsourced number. On the R&D funding side, IQVIA reports global biopharma funding reached a **10-year high of $102 billion in 2024**, a funding environment that helps explain why several large-cap pharma companies now have the balance sheet flexibility to fund GPU clusters priced in the tens of millions of dollars without disclosed line items in quarterly filings (Source: iqvia.com), a pattern also consistent with PhRMA's broader point that individual infrastructure line items remain small relative to the industry's decade-long spending base (Source: phrma.org).

Case Studies and Real-World Examples

Bristol Myers Squibb: CELMoD Compound Library Expansion and a Sickle Cell Candidate

On BMS's original DGX SuperPOD, prior to the 2026 Vera Rubin expansion, BMS's research team used AI to expand its library of CELMoD compounds, molecules that selectively degrade cancer-causing proteins used in blood cancer treatment, opening new potential drug targets (as detailed in the Strengths and Limitations discussion above). NVIDIA states that AI-enabled target identification on the existing infrastructure already **saves BMS scientists weeks of manual work** per program (Source: blogs.nvidia.com). BMS Chief Research Officer Robert Plenge specifically cited a **sickle cell disease candidate now in early clinical development** that, in his assessment, likely would not have been found without AI-enabled research on BMS's compute infrastructure (Source: reuters.com).

Genentech (Roche): AI-Accelerated Backup Oncology Candidate

Under Roche's Lab-in-the-Loop strategy, NVIDIA reports a specific, named-category outcome: one Genentech oncology protein-degrader molecule was designed **25% faster** than by conventional discovery methods, and a **backup drug candidate** for the same program was delivered in **seven months**, compared with the typical two-plus years such backup candidates usually take (Source: blogs.nvidia.com). This example is notable because Genentech disclosed a specific timeline comparison rather than only an aggregate adoption statistic, though the compound itself was not publicly named, consistent with Roche's broader pattern of disclosing program-level metrics without disclosing specific molecule identities.

Recursion Pharmaceuticals: BioHive-2 and REC-1245

Recursion Pharmaceuticals, a smaller, AI-native biotech rather than a legacy large-cap pharma company, completed **BioHive-2**, an NVIDIA DGX SuperPOD with **504 H100 GPUs**, in May 2024. The system ranked **#35 on the June 2024 TOP500** global supercomputer list, a TOP500 ranking based on its LINPACK benchmark result, making BioHive-2 unusual among the systems in this report for having a publicly reported standardized benchmark result (Source: ir.recursion.com). Recursion described BioHive-2 as **four times faster** than its predecessor, BioHive-1, and as the fastest supercomputer wholly owned by any pharmaceutical company at the time of its completion. Recursion Chief Technology Officer **Ben Mabey** said **"scaled data generation paired with scaled computation is required"** to meaningfully leverage AI in biology (Source: ir.recursion.com). On October 2, 2024, the **FDA cleared an Investigational New Drug (IND) application for REC-1245**, Recursion's first clinical candidate to emerge from its end-to-end AI discovery platform, a selective RBM39-degrading therapy for solid tumors and lymphoma; Recursion CEO Chris Gibson said the program moved from target biology to a preclinical candidate in under 18 months, **"nearly twice the speed of the industry average"** (Source: genengnews.com).

Insilico Medicine: Rentosertib Reaches Phase III

Insilico Medicine's AI-discovered TNIK inhibitor for idiopathic pulmonary fibrosis, **rentosertib** (formerly designated INS018_055), entered **Phase III trials in July 2026**. The molecular target, TNIK, was originally identified via Insilico's PandaOmics AI platform and reported as the top-ranked candidate in a published protein and kinase discovery analysis (Source: bio-itworld.com). Insilico CEO **Alex Zhavoronkov** said the company's platform record for taking a program from zero to developmental candidate is **9 to 18 months** (Source: bio-itworld.com), and the drug's Phase IIa results were published in **Nature Medicine** and presented at the American Thoracic Society's 2025 conference, giving the program peer-reviewed validation beyond company press materials (Source: bio-itworld.com). A separate AWS customer case study quotes an Insilico business development director describing the program's earlier stages: the company brought the fibrosis candidate **"from target discovery to compound validation in under 18 months for just $2.6 million"** using its PandaOmics and Chemistry42 platforms, one of the only publicly disclosed, program-level cost

figures in this entire report (Source: aws.amazon.com). A second AWS case study on Insilico's broader Pharma.AI platform migration reports an **83% reduction in time to deploy model updates** and more than 16 times faster model iteration and deployment after moving to Amazon SageMaker (Source: aws.amazon.com).

Isomorphic Labs and Novartis

Isomorphic Labs, the Alphabet-owned AI drug design company built on DeepMind's AlphaFold technology, expanded its research collaboration with **Novartis** in February 2025, after first announcing the partnership in January 2024, adding up to three additional research programs to the original three-target small-molecule scope (Source: isomorphiclabs.com). Novartis President of Biomedical Research **Fiona Marshall** said the collaboration enabled **"exploration of new chemical spaces that would be unavailable to probe through traditional methods"** (Source: isomorphiclabs.com), illustrating that not every large pharma company is pursuing the owned-supercomputer model exemplified by BMS, Lilly, and Roche; some, like Novartis, are instead buying access to a specialized AI partner's models and infrastructure on a per-program basis.

Implications and Future Directions

The compressed, nine-month sequence of superlative claims from Lilly, Roche, and BMS suggests the current phase of pharma AI infrastructure investment is at least partly **competitive signaling** aimed at investors, talent, and biotech partners, not purely an internally driven capacity decision. STAT News's framing that BMS's announcement was **"the third time in nine months"** a pharma company claimed the industry's largest AI supercomputer captures this dynamic precisely (Source: statnews.com), and Diogo Rau's caution about hype outrunning research, echoed by his own admission that Lilly's investment will not show its full clinical payoff until around 2030, suggests some Lilly executives themselves are wary of that dynamic (Source: rdworldonline.com) (Source: cnbc.com). At the same time, the underlying hardware trend is real and industry-wide: Amgen operates a 248-GPU "Freyja" SuperPOD for its deCODE genetics unit in Iceland (Source: blogs.nvidia.com), Novo Nordisk became a customer of Denmark's 1,528-GPU Gefion supercomputer, ranked 29th on the Top500 list of the world's fastest systems (Source: hpcwire.com) (Source: hpcwire.com), AstraZeneca joined a Swedish business consortium with Ericsson, Saab, and Wallenberg Investments to build a sovereign AI factory on Grace Blackwell GB300 systems for **"the next generation of AI enabled drug discovery and development"** (Source: wallenberginvestments.com), and Pfizer has reportedly been expanding its own GPU infrastructure toward more than 1,200 units within two years, alongside an AWS-based Scientific Data Cloud (Source: aimmediahouse.com).

Not every large pharma company is chasing owned GPU infrastructure at all. Sanofi has pursued a software-and-partnership model instead: in May 2024 the company announced a collaboration with Formation Bio and OpenAI as part of a stated ambition to become **"the first biopharma company powered by AI at scale"** (Source: sanofi.com). Novartis has directed capital toward R&D partnerships rather than infrastructure ownership: in December 2025 it committed up to **$1.7 billion** in milestone payments, with $55 million upfront, to AI biotech Relation Therapeutics for target discovery, a licensing-style deal rather than a compute buildout (Source: techicons.com). Merck's structure is similarly distinct from the GPU-cluster model: its up to $1 billion Google Cloud partnership, covering the company's 75,000 employees worldwide, targets agentic AI productivity tools across R&D, manufacturing, and commercial functions rather than a dedicated drug-discovery GPU cluster (Source: merck.com).

Table 2 below situates the three central companies of this report against an illustrative sample of pharmaceutical AI-compute and AI-access deals disclosed between 2020 and 2026. It shows that large pharmaceutical companies are pursuing a range of approaches, including owned GPU systems, hybrid capacity, cloud access, licensing, and software partnerships; it does not establish that GPU-backed infrastructure is a near-universal investment.

COMPANY	COMPUTE ASSET	SCALE DISCLOSED	MODEL	ANNOUNCED
Bristol Myers Squibb	Vera Rubin NVL72 SuperPOD	~576 GPUs (independently calculated)	Single-owned; deployment model and location not publicly disclosed	Jul 2026
Eli Lilly	LillyPod (B300 SuperPOD)	1,016 GPUs	Owned, on-premise	Oct 2025
Roche / Genentech	Hybrid AI factory	3,500+ GPUs	Hybrid on-prem/cloud	Mar 2026
Amgen (deCODE)	"Freyja" SuperPOD	248 H100 GPUs	Owned, on-premise (Iceland)	Jan 2024
Novo Nordisk	Gefion supercomputer (customer, not owner)	1,528 H100 GPUs	Multiyear compute-access agreement	Jun 2025
AstraZeneca	Swedish sovereign AI factory (consortium member)	GB300 systems, count undisclosed for AstraZeneca's share	Multi-party consortium	May 2025
Sanofi	Software/partnership model (Formation Bio, OpenAI)	Not a GPU cluster; no hardware count disclosed	Cloud/software partnership	May 2024
Pfizer	Own GPU fleet plus AWS Scientific Data Cloud	Targeting 1,200+ GPUs (secondary source)	Owned plus cloud	Reported Apr 2026
Merck & Co. (MSD)	Google Cloud agentic AI partnership	Not a GPU cluster; up to $1B cloud/software deal	Cloud/software partnership	Apr 2026
Novartis	Relation Therapeutics licensing deal	Not a GPU cluster; up to $1.7B milestone-based	R&D licensing partnership	Dec 2025
GSK	Cambridge-1 access plus own DGX A100s	Not re-disclosed since 2020 (Source: nvidianews.nvidia.com)	Shared access plus owned	Oct 2020

The table underscores an important asymmetry: **capital-intensive, owned hardware** (BMS, Lilly, Roche, Amgen) and **capital-light, cloud or partnership-based access** (Novo Nordisk, Merck, Sanofi, Novartis, GSK's Cambridge-1 arrangement) are both viable strategies pursued by top-20 pharmaceutical companies simultaneously, and the public record does not yet show one approach outperforming the other on any independently verified metric. Johnson & Johnson's October 2025 NVIDIA announcement, disclosed the same week as Lilly's, is a useful reminder that not every large pharma-NVIDIA deal is a supercomputer at all: J&J's uses NVIDIA's Isaac robotics platform for its MONARCH urology device line, a MedTech application rather than a drug-discovery compute cluster, and neither company disclosed the financial terms of that agreement (Source: biospace.com) (Source: biospace.com).

For pharmaceutical IT and R&D leaders evaluating whether to pursue owned infrastructure, cloud access, or an AI-native partner model, the practical decision criteria emerging from this comparison include:

- **Workload volume and predictability**: BMS's own account, that its original cluster became fully saturated, suggests owned infrastructure only pays off once GPU demand is sustained and predictable enough to justify fixed capital costs over variable cloud pricing.
- **Data sensitivity and IP control**: Lilly's TuneLab federated-learning design, which keeps partner biotech data on partner infrastructure while still enabling model training, shows that owned or hybrid infrastructure can be a prerequisite for certain data-sharing business models, not just a research-speed lever.

- **Disclosure and governance expectations**: the wide variance in what BMS, Lilly, and Roche have each chosen to disclose, GPU counts, dollar figures, performance metrics, suggests investor relations and competitive strategy, not technical necessity, substantially shape what companies publish about these systems.
- **Generational timing risk**: BMS's decision to skip directly to Vera Rubin while Lilly's flagship LillyPod remains on Blackwell Ultra illustrates the risk of hardware obsolescence in a category where NVIDIA is releasing new architectures roughly annually.
- **Energy and sustainability constraints**: Lilly's commitment to 100% renewable electricity for its AI infrastructure by 2030, and BMS's explicit framing of its Vera Rubin upgrade around the fact that **"electricity is not getting cheaper,"** both suggest power availability and cost, not just GPU procurement, are becoming binding constraints on further scale-up (Source: qz.com).

Frequently Asked Questions (FAQs)

Which pharma company has the largest AI supercomputer as of mid-2026? By raw GPU count, Roche's hybrid on-premise-plus-cloud footprint of more than 3,500 Blackwell GPUs is the largest publicly disclosed figure among the three companies profiled here (Source: roche.com). By disclosed compute throughput, Lilly's LillyPod, at more than 9,000 petaflops, is the only one of the three with a published performance figure at all (see the Eli Lilly section above). BMS calls its newly announced system the most powerful single-owned infrastructure in life sciences but has not disclosed a GPU count or performance figure to support that specific claim (Source: news.bms.com). There is no independently audited, apples-to-apples ranking across all three.

How many GPUs does BMS's AI supercomputer have? BMS has not disclosed a GPU count. NVIDIA's blog post describes the build as eight DGX Vera Rubin NVL72 racks (see the Bristol Myers Squibb section above), and since each NVL72 rack contains 72 GPUs, independent analysts calculate the total at approximately 576 GPUs, though this figure has not been confirmed by BMS itself (Source: drugdiscoverytrends.com).

What is the difference between NVIDIA Vera Rubin and Blackwell? Vera Rubin is NVIDIA's next-generation platform succeeding Blackwell, built around the Rubin GPU's larger transistor count and HBM4 memory upgrade, delivering AI inference at roughly one-tenth the cost per million tokens compared with Blackwell, according to NVIDIA's own published specifications (see Performance and Benchmarks above for full sourcing). As of July 2026, only BMS among the three profiled companies has committed its flagship supercomputer to Vera Rubin; Lilly's LillyPod and Roche's fleet both run on Blackwell or Blackwell Ultra.

How much are pharma companies spending on AI supercomputers in 2026? No single authoritative, pharma-specific capital-expenditure total exists publicly. None of BMS, Lilly, or Roche has disclosed a dollar figure for its flagship system; the only confirmed dollar figure among the three is Lilly and NVIDIA's separate, up-to-$1-billion five-year co-innovation-lab commitment, which is not a disclosed LillyPod cost (see the Eli Lilly section above). The nearest available proxy for hardware cost is public GPU and rack pricing: GB200 NVL72 racks list near $2 million to $3 million each (Source: gpusmith.com).

What does it cost to run a large-scale AI training run in pharma? No pharma company has published a training-run cost figure. The closest available data point is Insilico Medicine's disclosure that it brought an AI-discovered fibrosis drug candidate from target discovery to compound validation in under 18 months for $2.6 million using its own AI platform, a program-level rather than a single-training-run figure (Source: aws.amazon.com). At the broader AI-industry level, Epoch AI documents frontier training costs rising 2.4 times per year since 2016, with the largest runs expected to exceed $1 billion by 2027 (Source: epoch.ai) (Source: epoch.ai). IQVIA's finding of a 75% Phase I success rate for AI-enabled programs, discussed above, suggests these infrastructure bets may already correlate with better early clinical outcomes, though a causal link remains unproven (Source: iqvia.com).

Which other pharma companies besides BMS, Lilly, and Roche have NVIDIA-powered AI supercomputers? Amgen operates a 248-GPU "Freyja" SuperPOD for genomics research in Iceland (see Implications and Future Directions above), Novo Nordisk is a customer of Denmark's 1,528-GPU Gefion supercomputer (Source: hpcwire.com), and AstraZeneca is a member of a Swedish sovereign AI factory consortium using Grace Blackwell GB300 systems (Source: wallenberginvestments.com). Pfizer has reportedly been expanding its own GPU fleet toward more than 1,200 units (Source: aimmediahouse.com), while GSK's most concretely disclosed hardware commitment, access to the UK's Cambridge-1 supercomputer plus its own DGX A100 systems, dates to 2020.

Conclusion

Between October 2025 and July 2026, Eli Lilly, Roche, and Bristol Myers Squibb each announced what its press materials described as the most powerful or largest AI supercomputer in the pharmaceutical industry, an outcome only possible because each company chose a different metric to claim it by, as STAT News noted when it called BMS's move the third such claim in nine months (Source: statnews.com). Lilly's LillyPod, at 1,016

Blackwell Ultra GPUs and more than 9,000 petaflops, is the only system with a company-disclosed performance figure and the only one to reach production status with detailed, on-the-record operational updates. Roche's more than 3,500-GPU hybrid-cloud footprint is the largest by raw GPU count but blends on-premise and cloud capacity in a way that limits direct comparison to Lilly's owned-only figure, and it discloses no performance metric or investment figure at all. BMS's newly announced Vera Rubin build, likely around 576 GPUs by independent calculation, is the announced first planned pharmaceutical deployment of NVIDIA's newest architecture and claims the largest generational efficiency leap, a 10x gain in performance per megawatt, but is the least transparent of the three on hardware specifics.

For pharmaceutical technology leaders and their advisors, the practical lesson is less about which company currently holds the largest number and more about the pattern underneath all three announcements: sustained, multi-year NVIDIA partnerships beginning around 2023, rapid GPU-generation turnover from Hopper through Blackwell to Vera Rubin within a roughly three-year window, and a consistent business rationale centered on compressing the time between a biological hypothesis and a testable drug candidate. Independent, workload-level verification of the productivity claims these companies make, in petaflops, in molecules screened, or in months saved, remains largely absent from the public record, and organizations evaluating their own AI infrastructure strategy should weight vendor-reported superlatives accordingly while tracking the concrete, named outcomes, BMS's CELMoD library, Genentech's accelerated backup candidate, Recursion's FDA-cleared REC-1245, and Insilico's investigational Phase III rentosertib (Insilico Medicine; (Source: ClinicalTrials.gov), that have begun to emerge from this first wave of pharmaceutical AI supercomputing investment.

Tags: pharma ai supercomputer, nvidia vera rubin, nvidia blackwell, bristol myers squibb ai, eli lilly lillypod, roche genentech ai, pharma ai capex 2026, gpu pharma companies 2026, biopharma ai infrastructure

IntuitionLabs - Industry Leadership & Services

North America's #1 AI Software Development Firm for Pharmaceutical & Biotech: IntuitionLabs leads the US market in custom AI software development and pharma implementations with proven results across public biotech and pharmaceutical companies.

Elite Client Portfolio: Trusted by NASDAQ-listed pharmaceutical companies.

Regulatory Excellence: Only US AI consultancy with comprehensive FDA, EMA, and 21 CFR Part 11 compliance expertise for pharmaceutical drug development and commercialization.

Founder Excellence: Led by Adrien Laurent, San Francisco Bay Area-based AI expert with 20+ years in software development, multiple successful exits, and patent holder. Recognized as one of the top AI experts in the USA.

Custom AI Software Development: Build tailored pharmaceutical AI applications, custom CRMs, chatbots, and ERP systems with advanced analytics and regulatory compliance capabilities.

Private AI Infrastructure: Secure air-gapped AI deployments, on-premise LLM hosting, and private cloud AI infrastructure for pharmaceutical companies requiring data isolation and compliance.

Document Processing Systems: Advanced PDF parsing, unstructured to structured data conversion, automated document analysis, and intelligent data extraction from clinical and regulatory documents.

Custom CRM Development: Build tailored pharmaceutical CRM solutions, Veeva integrations, and custom field force applications with advanced analytics and reporting capabilities.

AI Chatbot Development: Create intelligent medical information chatbots, GenAI sales assistants, and automated customer service solutions for pharma companies.

Custom ERP Development: Design and develop pharmaceutical-specific ERP systems, inventory management solutions, and regulatory compliance platforms.

Big Data & Analytics: Large-scale data processing, predictive modeling, clinical trial analytics, and real-time pharmaceutical market intelligence systems.

Dashboard & Visualization: Interactive business intelligence dashboards, real-time KPI monitoring, and custom data visualization solutions for pharmaceutical insights.

Leading Claude & ChatGPT AI Enablement and Training: Help biotech and pharmaceutical teams adopt Claude and ChatGPT through practical training, governed workflows, use-case development, and implementation designed for regulated environments.

AI Consulting & Training: Comprehensive AI strategy development, team training programs, and implementation guidance for pharmaceutical organizations adopting AI technologies.

Contact founder Adrien Laurent and team at intuitionlabs.ai/contact for a consultation.

DISCLAIMER

The information contained in this document is provided for educational and informational purposes only. We make no representations or warranties of any kind, express or implied, about the completeness, accuracy, reliability, suitability, or availability of the information contained herein.

Any reliance you place on such information is strictly at your own risk. In no event will IntuitionLabs.ai or its representatives be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from the use of information presented in this document.

This document may contain content generated with the assistance of artificial intelligence technologies. AI-generated content may contain errors, omissions, or inaccuracies. Readers are advised to independently verify any critical information before acting upon it.

All product names, logos, brands, trademarks, and registered trademarks mentioned in this document are the property of their respective owners. All company, product, and service names used in this document are for identification purposes only. Use of these names, logos, trademarks, and brands does not imply endorsement by the respective trademark holders.

IntuitionLabs.ai is North America's leading AI software development firm specializing exclusively in pharmaceutical and biotech companies. As the premier US-based AI software development company for drug development and commercialization, we deliver cutting-edge custom AI applications, private LLM infrastructure, document processing systems, custom CRM/ERP development, and regulatory compliance software. Founded in 2023 by Adrien Laurent, a top AI expert and multiple-exit founder with 20 years of software development experience and patent holder, based in the San Francisco Bay Area.

This document does not constitute professional or legal advice. For specific guidance related to your business needs, please consult with appropriate qualified professionals.