Artificial Intelligence in Biotech Therapeutics:

Applications, Benefits, and Future Directions

Christian Grant

Deep AI Mind
Copenhagen, 22 January 2026

Artificial Intelligence in Biotech Therapeutics:

Applications, Benefits, and Future Directions

Core AI Use Cases in Biotech

Artificial intelligence (AI) is being applied across the pharmaceutical R&D pipeline – from early discovery and preclinical research through clinical development, manufacturing, diagnostics, and even synthetic biology. In drug **discovery**, AI-driven platforms help identify novel drug targets and design new therapeutic candidates. For example, machine learning models can mine large genomic and phenotypic datasets to uncover disease-driving biological pathways and targets that were not evident via traditional methods . Deep learning-based virtual screening allows scientists to explore "infinitely" large chemical spaces in silico, far beyond the few million compounds of a physical high-throughput screen . Notably, generative AI models are used to propose new molecular structures (small molecules or protein drugs) with desired properties. Pfizer's partnership with startup PostEra is one example – using generative chemistry, they design small molecules and even antibody–drug conjugate components with the model ensuring the outputs are synthetically tractable . This approach accelerated one antiviral program from a "cold start" to in vivo proof of concept in only 8 months (about **40% faster** than Pfizer's prior expectations) . Similarly, generative models are designing biologics: Absci's platform uses billions of protein interaction data points to train AI that can create novel antibodies, enabling rapid iteration from an AI-designed antibody to a lab-validated candidate in as little as six weeks .

In **preclinical research**, AI improves lead optimization and safety testing. Predictive machine learning models assess drug-like properties (e.g. potency, ADME, toxicity) early on, helping chemists prioritize which analogs to synthesize. For instance, Novartis reported using AI to digitize decades of safety data and detect toxicity signals sooner – allowing researchers to kill unpromising compounds early and focus on safer molecules . Image-based deep learning is also leveraged in preclinical biology: companies like Recursion Pharmaceuticals use high-dimensional **phenomics** (microscopic cell images of disease models) combined with neural networks to identify drug candidates that reverse disease phenotypes . By analyzing subtle patterns across

millions of cell images, AI can predict a compound's effect on cellular pathology and suggest novel therapeutic mechanisms.

AI has significant roles in the **clinical development** stage as well. One major application is optimizing clinical trial design and patient recruitment. AI systems can integrate real-world data (electronic health records, prior trials) to simulate trial outcomes under various inclusion/exclusion criteria. As an example, Novartis uses software to model how adjusting parameters like patient age, concomitant medications, or diversity quotas will impact enrollment timelines . By iteratively testing criteria in silico, trial designers can identify an optimal design that recruits faster while maintaining statistical power and patient diversity. AI is also used for **patient stratification** – analyzing genetic or imaging biomarkers to select patients most likely to respond, which increases trial success rates. Additionally, natural language processing (NLP) models automate the matching of patients to trials by parsing medical records at scale, accelerating recruitment. During trials, AI-driven analysis of medical images and digital biomarkers can provide earlier readouts of drug efficacy (for instance, detecting tumor response on scans sooner than human review).

In **biopharmaceutical manufacturing**, AI contributes to process optimization, quality control, and supply chain efficiency. Biotech manufacturing is data-rich – sensors monitor bioreactor conditions, and high-throughput analytics generate vast data on product quality. Machine learning models are employed to tune process parameters in real time for higher yields and fewer batch failures. For example, AI can model the complex relationships in cell-culture manufacturing of biologics, allowing proactive adjustments to maintain optimal growth conditions. AI-based computer vision is used to inspect products on production lines (for fill–finish of vials, packaging integrity, etc.), detecting defects more reliably than manual inspection. These technologies lead to more consistent product quality and lower production costs. AI also improves supply chain logistics by predicting demand and equipment maintenance needs. According to DSM-Firmenich, implementing AI in production and supply chain management has increased efficiency and reduced time-to-market, partly by identifying process bottlenecks and maintenance issues before they cause delays .

In **diagnostics** and **precision medicine**, AI helps identify patients for targeted therapies. Pharmaceutical companies are increasingly developing AI-powered companion diagnostics – for example, algorithms that analyze pathology slides or genetic data to find biomarkers indicating a patient will respond to a given drug. Roche, which has large pharma and diagnostics divisions, notes that AI image analysis can reveal which tumors express certain protein targets, informing both drug discovery

and patient selection . AI-based diagnostic models also enable earlier disease detection from medical images, wearable sensors, or multi-omics data, which can expand the eligible population for new therapies by finding patients sooner.

Finally, AI plays a growing role in **synthetic biology and biotherapeutics design**. Companies like Ginkgo Bioworks integrate AI into the engineering of cells and biological products. Their platform uses machine learning to design genetic modifications and predict the performance of engineered microbes that produce therapeutic proteins or other biomolecules. A notable trend is training large "foundation models" on protein sequences and structures – for instance, Ginkgo and Google Cloud launched a large language model for proteins to assist in enzyme design . These models, akin to GPT but for biology, can suggest genetic edits or predict which enzyme mutations will improve activity. In one early example outside human therapeutics, Firmenich (now DSM-Firmenich) created a novel flavor molecule using an AI design algorithm . In the therapeutic arena, generative AI is being used for **de novo enzyme design** (e.g. generating enzymes that break down disease-related metabolites or environmental toxins) and for creating improved gene therapy vectors. Overall, across discovery and development, AI's core use cases span **predictive modeling** (e.g. will a molecule hit a target or be toxic?), **deep learning on complex data** (images, sequences, medical records), **generative design** of new candidates, **multi-modal data integration** (combining chemistry, biology, clinical data), and emerging **AI-driven experimental automation**.

Main Benefits of AI

Biotech and pharma companies are pursuing AI because it promises to make R&D faster, cheaper, and more productive. A primary benefit demonstrated to date is **speed**. AI can dramatically shorten certain steps in drug development. Insilico Medicine, for example, reported that it advanced 22 AI-discovered drug candidates from initial idea to IND-enabling studies in an average of just 12–18 months – a timeline previously unheard of . Its lead program (a novel anti-fibrotic drug) went from AI hypothesis to a Phase I trial in about 30 months . Similarly, Recursion achieved discovery to IND in 18 months for a new compound . These cases suggest AI-driven workflows can **reduce preclinical development times by 40–50%** in favorable scenarios . Even for more novel biology projects, independent analyses predict AI can make early development ~35% faster and ~25% cheaper than traditional approaches . Faster development not only gets novel medicines to patients sooner, but also extends

the effective patent life and revenue window for successful drugs – a major business advantage.

Another key benefit is **improved success rates and decision-making**. While AI-designed drugs are still too few to fully assess clinical success improvement, early indications show AI can help avoid failures. AI excels at analyzing high-dimensional biological data, which helps researchers make more informed decisions on targets and lead compounds. By learning from vast datasets (e.g. omics profiles, phenotypic screens, historic trial data), AI models can flag drug candidates likely to fail due to toxicity or lack of efficacy before costly human trials. For instance, prediction models that incorporate human genetic data and physiological complexity aim to filter out compounds that might only work in animal models but not humans . This **failure mode forecasting** could curb the industry's high attrition rates. There is also evidence that AI-augmented drug design yields compounds with better properties: Absci's AI was able to design an antibody with higher affinity and stability than any human-designed variant in that program . In general, AI's ability to search huge chemical and biological spaces can lead to more **novel solutions** – finding molecules, targets, or biomarker patterns that human experts might overlook. Novartis noted that AI-guided virtual screening led them to "new chemical space" for tough targets, yielding chemical matter very different from their prior attempts . Such novel first-in-class drugs can be especially valuable if they address diseases that lacked viable targets.

On the operational side, AI brings **efficiency and cost savings**. Many AI applications automate labor-intensive processes, freeing scientists from repetitive tasks. At Pfizer, internal assessments showed that automating complex analytical workflows with AI (including using large language models as "copilots" for researchers) cut the time required by up to 80% in some cases . AI-powered automation increases human productivity – enabling teams to manage larger project portfolios or focus on high-level creative work instead of data wrangling. In manufacturing, AI-driven process control has reduced waste and improved yields, translating to lower cost of goods. DSM-Firmenich estimates that data analytics and AI together have contributed to **lower costs and faster production ramp-ups**, in part by optimizing supply chain and production processes in real time .

From a scientific perspective, one of AI's biggest benefits is the **integration of diverse data** to generate insights. Pharma R&D involves many data types (genomic data, imaging, chemistry, clinical data). AI models, especially modern multi-modal models, can combine these sources to uncover patterns no single dataset would reveal. This helps researchers understand disease mechanisms more holistically. Roche scientists

note that AI allows them to ask “bigger questions” – instead of studying one target at a time, they can analyze networks of genes and proteins and their interplay in disease . This systems-level view can identify optimal intervention points for complex diseases. AI is also enabling **personalized medicine** benefits: by analyzing patient data at scale, AI can identify subpopulations likely to benefit from a therapy (or those at risk of side effects), allowing more targeted and effective treatment strategies.

Another oft-cited advantage is **quality and consistency**. Machine learning algorithms don’t suffer from human fatigue or bias in execution, so when used for tasks like image analysis or QC inspection, they can be more consistent. For example, an AI diagnostic algorithm will apply the same criteria to every MRI or pathology slide, reducing variability in how results are interpreted. This consistency improves the reliability of diagnosis and patient selection for trials, ultimately benefiting outcomes. AI also facilitates **continuous learning**: models can be updated with new data to improve over time, creating a virtuous cycle of learning in an organization. Companies that deploy AI across R&D report that it acts as a “force multiplier” for human expertise – the combination of AI algorithms with human scientific intuition yields better results than either alone. Overall, when implemented well, AI leads to faster innovation cycles: more experiments or simulations can be run in a given time, more hypotheses tested, and more shots on goal in the search for effective therapeutics .

Finally, there are **business and strategic benefits**. AI capabilities are increasingly seen as a competitive differentiator in the biopharma industry. Companies with strong AI pipelines may attract more partnerships and investment. For example, by 2025 it’s estimated that **30% of newly approved drugs may involve AI in their discovery** (a stark increase from a decade prior) . Leaders like AstraZeneca, Pfizer, and Lilly have all highlighted AI as central to their R&D strategy, reporting concrete achievements such as AI-identified drug candidates entering clinical trials . Embracing AI also signals to investors and partners that a company is forward-looking and efficiency-driven. In practice, major pharmas have formed high-profile alliances (e.g. Pfizer with multiple AI startups, Roche with Recursion, BMS with Exscientia) to gain access to novel AI platforms. These alliances can yield milestone payments and shared IP, benefiting the AI startups commercially as well. In summary, the benefits of AI in biotech include **speeding up discovery, reducing costs, enhancing decision quality, expanding scientific possibilities, and improving operational efficiency**, all of which ultimately can translate to better patient outcomes and stronger business performance.

Key Challenges

Despite the promise, there are significant challenges in deploying AI in biotech R&D. One fundamental hurdle is **data quality and availability**. AI systems are only as good as the data they are trained on – the adage "garbage in, garbage out" fully applies . Biotech data can be noisy, biased, or siloed. For instance, experimental datasets may have batch effects or inconsistent annotations, and negative results often go unreported, skewing learning. Many companies have had to spend years curating and digitizing their research data before AI could yield value . At Pfizer's 2025 AI summit, a core theme was that high-quality, relevant data is the true driver of AI – requiring robust data pipelines, annotation, and constant maintenance . Another limitation is simply data scarcity in certain areas: for rare diseases or novel targets, there may be too little historical data for an AI model to learn meaningful patterns. Unlike internet applications that enjoy billions of data points, biotech often operates with small-n datasets (e.g. a few hundred patients in a trial), making it hard to train reliable models. Ensuring that models are trained on **diverse, representative data** is crucial, otherwise AI tools may not generalize – a lesson seen when early diagnostic AIs trained on homogeneous populations failed in broader use. In drug discovery, lack of negative examples (failed compounds that weren't published) can also mislead models. Companies are addressing this by generating massive proprietary datasets (as Ginkgo Bioworks is doing with its Datapoints program to produce millions of biological data points for model training) , but data bottlenecks remain a challenge for smaller firms.

Another significant challenge is **model interpretability and validation**. Many AI algorithms, especially deep learning models, act as "black boxes" – they may predict an outcome or propose a molecule without providing a clear rationale. In a highly regulated, safety-focused field like therapeutics, this lack of transparency is problematic. Researchers and regulators want to understand *why* a model made a recommendation (e.g. identified a target or predicted a toxicity) before acting on it. If an AI proposes a new drug candidate, scientists must still validate the underlying biology: is the predicted target relevant to disease? Are there off-target effects the AI didn't foresee? Often the first few AI suggestions can be false positives, which means extensive lab work to filter them. This need for **experimental validation** can erode some of AI's efficiency gains. As one commentary noted, accelerated AI timelines have not yet translated into better Phase II success rates – in part because many AI-designed molecules still act on known targets or mechanisms, and the real test (improved efficacy in humans) remains to be proven. There is concern that if AI primarily optimizes within the constraints of training data (e.g. focusing on well-validated targets), it might not solve the translational gap in medicine unless paired with strategies to incorporate human complexity earlier . Regulators like the FDA have voiced that while AI can support drug development decisions, sponsors are responsible for the **reliability of AI-generated evidence**. In mid-2025, FDA released a draft guidance outlining principles for AI used in regulatory submissions, emphasizing

transparency, reproducibility, and the need for rigorous validation of any model that influences safety or efficacy decisions . Ensuring compliance with such guidelines (and the equivalent efforts by EMA in Europe) will be a new challenge for companies – they must document how an AI model was developed, its limitations, and how its output was confirmed with experimental data.

Integration of AI into existing **workflows and culture** is another non-trivial challenge. Pharma R&D is traditionally siloed and sequential; implementing AI often requires breaking down silos and fostering collaboration between data scientists and bench scientists. There can be organizational resistance – some experienced chemists and biologists may be skeptical of AI, preferring established methods. Novartis's R&D head noted that initially there were skeptics, and part of their AI program's success involved convincing teams by demonstrating head-to-head comparisons (e.g. how AI-designed compounds improved properties faster than the classic approach) . Training and upskilling of staff is necessary so that domain experts can use AI tools effectively. Companies have instituted internal programs (Pfizer's AI Festival is one example) to build AI fluency among scientists and to encourage a cultural shift toward data-driven experimentation. Change management is tough: researchers must learn to trust AI where it works and also recognize its failure modes. As Fiona Marshall of Novartis mentioned, a pattern early on was teams running their usual lab process in parallel with the AI-guided process to see which performs better, essentially double-working . This can temporarily reduce efficiency, but it's a natural part of adoption until AI earns trust. Over time, as success stories accumulate, resistance tends to lessen – but maintaining human oversight is still important.

There are also **technical challenges** specific to biotech data. Biological systems are exceedingly complex, and AI models sometimes encounter data far outside the distribution of their training set (for example, a patient with a rare combination of diseases, or a completely novel chemotype). In such cases, models may make unreliable predictions. Ensuring that AI systems can detect when they are uncertain or out-of-scope is an open problem. Moreover, multi-modal integration (combining omics, imaging, text, etc.) is computationally intensive and algorithmically challenging – aligning data of different types and scales is not straightforward. The computational requirements for state-of-the-art AI can be immense, raising infrastructure challenges. Training a modern deep learning model on chemical and genomic data may require powerful GPUs or TPUs and significant memory. Not every company has access to supercomputing resources – though cloud computing is alleviating this, it adds cost and potential data security concerns (when uploading sensitive data to cloud servers).

Speaking of security, **data privacy and IP** present challenges. Training AI on patient data (like medical records or genetic data) triggers privacy laws – companies must anonymize and secure data to comply with regulations such as HIPAA and GDPR. If AI models learn from proprietary datasets (e.g. a pharma's collection of drug assays), protecting the intellectual property in those data becomes an issue, especially if external partners or cloud providers are involved. There is also the question of IP rights for AI-generated inventions (e.g. patents on molecules suggested by AI) – currently the human inventors directing the AI are listed, but legal frameworks are evolving.

In the realm of **regulatory and ethical challenges**, beyond validation, there is caution about how AI models might inadvertently reflect biases. For instance, an AI that learns from historical trial data might propagate any demographic bias in those trials, potentially suggesting skewed trial recruitment that unintentionally under-represents certain groups. Regulators will expect evidence that AI tools do not introduce unfair bias or jeopardize patient safety. The FDA's increased scrutiny (they noted 500+ AI-involving submissions from 2016–2023 and rising) means companies must be prepared for questions around algorithmic accountability.

Specific to **AI in robotics** and lab automation, challenges include the reliability and validation of robotic systems in a GMP (Good Manufacturing Practice) or GLP environment. Robots and AI algorithms in labs must be rigorously tested to ensure they handle edge cases (e.g. what if a pipette tip is missing or a sensor malfunctions?). Any error in automated lab work could invalidate an experiment or, worse, a batch of a therapy. Additionally, a point raised by Insilico's CEO is that most lab equipment was designed for humans and often needs retrofitting for robots . Even "automated" equipment often requires human intervention for certain tasks (refilling reagents, calibrations). Humanoid robots are proposed to mitigate this, but they are still early-stage and need to prove they can perform with the dexterity and consistency required. **Regulatory compliance** for robotic processes is another area: companies must demonstrate that an AI-driven robot process yields the same (or better) quality as a manual process. This requires extensive documentation and sometimes developing new standards for validating AI behavior over time.

In summary, while AI offers powerful tools, biotech companies face a confluence of challenges: ensuring robust data foundations, interpreting and validating AI outputs, integrating AI into existing scientific workflows, and satisfying regulatory and ethical requirements. These challenges demand a careful, often phased approach to AI adoption – starting with low-risk use cases to build confidence (like using AI for internal decision support rather than pivotal regulatory decisions initially) and gradually

expanding as trust and experience grow . Companies that navigate these challenges successfully will likely create a strong competitive moat, whereas those that stumble may face setbacks or regulatory pushback. The consensus in the industry is that AI is here to stay, but it must be approached with scientific rigor, proper oversight, and a willingness to address its current limitations head-on .

High-Value Biotech Functions for AI Application

Not all functions in biotech derive equal value from AI – some areas have proven especially ripe for AI-driven transformation. On the **scientific research** side, **drug discovery** is arguably the highest-value function for AI today. This is the stage where key decisions are made that determine a program's fate (choice of target, lead molecule, etc.), and historically it's been slow and failure-prone. AI's ability to analyze complex biological data and generate novel hypotheses can dramatically improve target identification and lead generation. An estimated 90% of drug candidates fail to reach approval, often due to choosing the wrong target or molecule . By applying AI to better select targets and design more optimized leads, companies can avoid pouring resources into doomed projects. For example, identifying a valid drug target that impacts disease biology is of enormous value – AI can sift human genomic data, literature, and phenotypic screens to suggest targets with a higher likelihood of clinical success (Roche/Genentech's AI group emphasizes using ML to find *combinations* of targets or network nodes that drive disease, aiming for more effective intervention points). Even a single high-quality drug target can be worth billions in potential value if it yields a first-in-class therapy.

Within discovery, **lead optimization** and **medicinal chemistry** are also high-value areas for AI. Medicinal chemists traditionally cycle through many design-make-test iterations to refine a drug's properties. AI can shorten this DMTA (design–make–test–analyze) cycle by guiding chemists to the best modifications sooner . This not only saves time but also can yield a superior drug candidate (with optimal potency, selectivity, pharmacokinetics, etc.), increasing the chance of clinical success. Companies report that AI-enabled compound design has led to "better starting points" – for instance, Schrödinger's physics-AI platform helped develop a TYK2 inhibitor (now in Phase III trials) that might not have been found through brute-force chemistry alone . The value of getting a drug to clinic with ideal properties (and avoiding late-stage failures due to suboptimal chemistry) is enormous.

In **preclinical development**, safety/toxicity assessment is a high-value function for AI because of its impact on downstream success and cost. Predicting toxicity or identifying safety biomarkers using AI can save a company from advancing a toxic compound into costly animal studies or human trials. The earlier a flag is raised, the more resources are saved and can be redirected. For example, an AI that can predict off-target interactions or liver toxicity risk based on a molecule's structure and known toxicophores can prompt chemists to tweak the structure early. As noted earlier, Novartis's use of AI to analyze historical safety data aims to catch safety issues sooner – terminating or redesigning a program before it becomes an expensive failure is a high-value intervention.

Moving to **clinical development**, AI's highest-value contributions are in **clinical trial design and optimization**. This encompasses selecting the right patient population, endpoints, and sites – decisions that can make or break a trial. A trial failing due to poor design or slow enrollment can waste hundreds of millions of dollars and years of time. AI models that predict enrollment rates at various sites or that identify criteria to broaden eligibility without compromising safety are extremely valuable. For instance, by using AI to simulate trial recruitment under different scenarios (as described by Novartis) , companies can design trials that recruit in months instead of years. AI can also help ensure trials are inclusive (by modeling strategies to reach underrepresented patient groups) and thereby produce more generalizable results . Additionally, AI is being used to **augment control arms** via external data (e.g., using real-world patient data to serve as a synthetic control arm), which can reduce the number of patients needed in a trial – a direct cost and time saving, and ethically beneficial by requiring fewer placebo recipients.

Another high-impact area is **biomarker discovery and companion diagnostics**, which often overlaps between late discovery and clinical development. AI can identify biomarkers that predict response or stratify patients, enabling **precision medicine** approaches. A successful biomarker can transform a drug's profile by identifying exactly who benefits from it, which can lead to smaller, faster trials and better outcomes. For example, an AI that analyzes tumor images and genomics might find that a particular protein expression pattern predicts response to a new cancer drug – the company can then use that AI-derived biomarker to select patients in Phase II, drastically increasing the measured efficacy. The value is seen in higher success rates and potentially expedited approval if a clear responsive subgroup is identified.

Beyond R&D, **manufacturing and supply chain** functions in biotech also reap high value from AI, though these are sometimes underappreciated. For a biologic drug

that's approved, even small efficiency gains in manufacturing can save tens of millions of dollars annually. AI-driven process control can maximize yield of biologics (like antibodies, which are grown in cell culture). Given that manufacturing costs contribute significantly to the price of therapies (particularly advanced therapies like cell therapies), AI optimizations in bioprocessing have a direct financial impact. Supply chain optimization – predicting demand for drug product, scheduling production, managing cold chain logistics – is another area where AI adds value by preventing stock-outs or overproduction. DSM-Firmenich explicitly notes that AI helps optimize their production processes and supply chain, supporting customers more effectively while lowering costs . In an era of complex global supply chains (highlighted by recent disruptions), having AI to dynamically adapt production plans is a competitive edge.

Quality and compliance functions are also seeing AI contributions. High-value examples include AI for **pharmacovigilance**: using NLP to scan medical reports and social media for adverse event signals faster than traditional methods. Early detection of a safety signal can be worth a lot (preventing patient harm and regulatory penalties). Likewise, AI can streamline regulatory submission processes by checking documents for consistency and completeness, reducing the risk of compliance issues that could delay approvals.

On the business side of biotech, functions like **portfolio management and forecasting** gain value from AI. Companies like Pfizer and Novartis are using predictive analytics to forecast clinical trial outcomes or market uptake, helping them allocate R&D budget more wisely. While not as publicized as AI in the lab, these applications can yield high financial value by focusing resources on the most promising programs (some biotechs use AI from firms like Intelligencia to predict probability of success for each project and prioritize accordingly). Additionally, **commercial strategy** (though outside the therapeutic R&D focus) can benefit – e.g. AI models analyzing real-world data to identify which patient segments to target for a launched drug, or optimizing marketing efforts. Those are high ROI activities, albeit once a drug is on the market.

In summary, the highest-value functions for AI in biotech are those that **impact critical decision points, major cost drivers, or key success factors** in the therapeutic pipeline. Target discovery, lead design, safety assessment, trial design, and manufacturing are top among these. Success in any one of these areas can save years of work or hundreds of millions in cost, or conversely create a breakthrough product that generates substantial revenue. Biotech and pharma executives are therefore especially focused on applying AI in these domains first, where the value proposition is clearest.

Notable AI Technologies Used

Several categories of AI technologies are powering the above applications, each with distinct roles:

- **Machine Learning (ML)** – This broad category includes traditional algorithms (regression, decision trees, random forests, gradient boosting machines, etc.) that learn patterns from data. ML has been a workhorse in pharma for years in the form of quantitative structure–activity relationship (QSAR) models, statistical models for clinical data, etc. These models are interpretable and work well on smaller datasets. For example, a random forest model might be trained on known drug molecules to predict whether a new molecule will likely be orally bioavailable. Pharma companies apply ML to tasks like predicting chemical properties, clustering patients by disease subtype, or forecasting manufacturing yield. Genentech scientists note that they aim to develop models that are **predictive, generative, and interpretable** . A predictive ML model might tell you if a molecule can bind a target (yes/no), or if a patient falls into a responder group. These models often use features engineered by human experts (e.g. molecular descriptors, or lab values from patients). While simpler than deep learning, ML models are faster to train and easier to validate, so they remain widely used especially when data are limited. An example is Genentech's use of an ML model to predict which candidate antibiotics can penetrate Gram-negative bacterial membranes – by training on chemical structures and known outcomes, the model filters an enormous virtual library down to a manageable set of compounds to synthesize . ML models also drive knowledge graphs and repurposing tools that ingest literature and find non-obvious connections (e.g., linking an old drug to a new disease through shared gene networks).

- **Deep Learning** – This refers to neural network models with many layers, which have revolutionized AI in the past decade. Deep learning is adept at handling unstructured and high-dimensional data, making it ideal for genomics, imaging, and speech/text in biotech. Convolutional Neural Networks (CNNs) are used for image analysis – for instance, analyzing histopathology slides or high-content cellular images. Recursion built one of the world's largest biological image datasets (from cell-based screens) and uses CNNs to embed these images into "phenotypic fingerprints," allowing clustering of drugs by their effects . This approach uncovered new relationships (e.g. a drug treating one disease might reverse the cell phenotype of a completely different disease, suggesting a repurposing opportunity). Recurrent neural networks and transformers (sequence-based deep models) are widely used on genetic and protein

sequence data. Models like these can predict what a DNA or amino acid sequence does, or propose new sequences (for example, Moderna has used deep learning on mRNA sequence data to optimize the design of mRNA therapeutics for higher protein yield and stability – deep models capture subtle sequence motifs that human designers might miss). **Transformers**, in particular, have enabled large language models in biomedicine: models trained on protein sequences (like Meta's ESM-2) can "understand" biochemical properties of proteins, and models trained on scientific text can answer complex research questions. DeepMind's **AlphaFold** is a prime example of deep learning in biotech – it uses a specialized neural network to predict 3D protein structure from sequence, a breakthrough that provided the structures of ~200 million proteins including many human ones . The availability of AlphaFold structures has been a boon to drug discovery, allowing researchers to perform structure-based design even when experimental crystallography is too slow or failed. Many pharma companies have integrated AlphaFold into their workflows for target modeling and ligand docking. Overall, deep learning's ability to automatically extract features makes it extremely powerful; however, it demands large datasets and compute, and can be a black box. Nonetheless, where big data is available (e.g. millions of chemical structures, thousands of patient omics profiles), deep nets are often the state of the art.

- **Generative AI** – Generative models are those that learn the underlying distribution of data and can create new synthetic data with similar characteristics. In biotech, generative AI is mainly used for **molecule generation** (both small molecules and macromolecules like proteins). Techniques include variational autoencoders (VAEs), generative adversarial networks (GANs), and modern approaches like diffusion models. These models can propose novel chemical structures that optimize multiple objectives (binding to a target, avoiding toxicity, etc.). A notable example is Insilico Medicine's generative pipeline: it uses generative adversarial networks and reinforcement learning to ideate new molecules and then score them . Insilico's AI platform, as reported, has produced a pipeline of 30 novel drug candidates, with 10 already approved to enter clinical trials – a strong proof-of-concept for generative chemistry. Absci's generative AI for biologics designs antibody sequences that meet very specific criteria (affinity, manufacturability), which a human might not arrive at easily . These generative models often work in tandem with predictive models: for instance, a reinforcement learning loop will generate a molecule, test it with a predictive model for desired properties, then adjust generation accordingly (this is how generative**RL** frameworks work in drug design). **Large language models (LLMs)** can also be considered generative AI – they generate text, which companies are now leveraging for tasks like generating study reports, writing first drafts of regulatory submissions, or even suggesting hypotheses by synthesizing literature. Pfizer's use of an internal generative model named "Charlie" was noted in early 2024, aiming to assist scientists by generating insights from Pfizer's vast data silos . While generative AI opens

exciting possibilities, it also introduces new challenges – such models can hallucinate or produce infeasible suggestions (like a molecule that is impossible to synthesize, as early generative tools often did). Hence, integration of synthetic feasibility checks (like PostEra's emphasis on synthesizability in its generative models) is critical. As hardware and algorithms improve, generative AI is expected to take a larger role, for example designing therapeutic *modalities* beyond conventional structures – we're already seeing generative design of cell therapies (e.g. designing T-cell receptors) and AAV capsids for gene therapy.

- **AI Agents and Automated Decision Systems** – An emerging area is the use of AI agents, which are systems that can perceive an environment, make decisions, and act towards a goal, often using reinforcement learning or planning algorithms. In the context of biotech, AI agents coordinate complex sequences of tasks or even control lab robots autonomously. A practical example is the concept of a "self-driving lab." Companies like Insilico and Eli Lilly (in partnership with NVIDIA) are connecting AI agents to both **wet lab** and **dry lab** components. Lilly's new AI-driven lab aims to tightly link automated experiments with model updates in a continuous loop – an AI agent decides which experiment to run next based on prior results and constantly retrains models on the new data, essentially an autonomous scientific cycle . These systems are sometimes called "lab-in-the-loop" or "robot scientist" frameworks. Genentech's collaboration with NVIDIA in 2023 was geared towards such an AI-orchestrated lab setup for Roche's research, where AI agents manage experimental workflows and data analysis in unison . Another kind of AI agent in biotech is one that navigates knowledge spaces – for example, an AI agent that reads papers (an NLP-based agent) to keep scientists up to date, or an agent that helps design a clinical trial by iterating through protocol parameters (like the scenario Novartis described for trial simulation, one can imagine an AI agent adjusting inclusion criteria until an optimal set is found). Agents can also appear in manufacturing, adjusting operations on the fly for efficiency. This technology is still maturing, but it holds promise to **automate complex decision-making** that currently requires human oversight. Insilico's humanoid lab robot "Supervisor" is paired with what they call *agentic AI* – algorithms that allow the robot to operate laboratory equipment and perform tasks with minimal human direction . In essence, these AI agents act like autonomous lab assistants or process controllers. Over the next few years, we expect to see more AI agents in software form (for data analysis and design tasks) and embodied form (robots in labs and plants).

- **Multimodal AI Models** – Multimodal models can process and link multiple types of data (text, images, numerical data, molecular structures, etc.). Many problems in biotech are inherently multimodal: for example, understanding a disease may require imaging data, genomic sequences, and clinical notes together. Traditional models handled one modality at a time, but new architectures (often based on transformers) can ingest concatenated or parallel inputs of different

types. A use case is in **drug discovery for complex diseases**: one might combine high-content imaging readouts with gene expression profiles and chemical structures to build a holistic model of how different compounds affect a disease phenotype. Recursion, for instance, builds **maps of biology** that integrate genetic perturbation effects (from CRISPR knockout experiments) with phenotypic image data – creating a unified embedding where distance reflects biological similarity. Such an approach required multimodal learning to connect genotype and phenotype. Another example is using clinical trial datasets where you have images (MRIs), clinical diagnostics, and patient demographics; a multimodal AI could learn patterns that span these (e.g. an image finding plus a lab test together predict a certain outcome). **Foundation models** in biology are often multimodal or at least adaptable to multimodal tasks. NVIDIA and Novo Nordisk's collaboration explicitly mentions plans to build customized AI models and agents for research and clinical development, including using single-cell omics data alongside structural data . They also aim to mine vast scientific literature with large language models – combining textual information with experimental data to generate new hypotheses. In manufacturing, **digital twins** are a form of multimodal AI application: a digital twin of a bioreactor combines physics-based models with real-time sensor data (numbers, maybe images from in-process microscopy) to simulate the process. Lilly, for example, is using NVIDIA's Omniverse platform to create digital twins of its production lines, which integrate visual data, equipment data, and process metrics, allowing AI to simulate and optimize manufacturing changes virtually . The trend is clearly toward AI models that can bridge domains – chemical, biological, clinical – which mirrors how human experts work when making decisions. These multimodal models are complex but are expected to yield richer insights than siloed analyses.

In summary, the AI toolkit in biotech spans from relatively simple ML models to very sophisticated deep networks and autonomous agents. Companies often employ a mix: a pipeline may use a simple ML model for quick predictions, a deep learning model for complex pattern recognition, a generative model for design, and an AI agent to coordinate it all. It's the **convergence** of these technologies that is unlocking truly novel capabilities. For instance, a "robot scientist" setup might use a deep learning model to analyze results, a generative model to propose the next experiment, and an AI agent to physically execute that experiment with lab robots – a combination that was science fiction not long ago. While these technologies are distinct, they complement each other, and leading organizations are building integrated platforms to take advantage of the full AI spectrum .

AI-Enabled Robotics and Humanoid Robots in Biotech

Automation has long been part of biotech (with robotic liquid handlers and automated assays), but AI is enabling a new generation of robotics – including humanoid robots – to work in labs and manufacturing facilities. **AI-enabled robotics** refers to robotic systems guided by advanced AI for perception and decision-making, allowing them to perform complex tasks with minimal human input. In cutting-edge labs, companies are deploying robots not just for repetitive liquid-handling, but for more adaptive tasks like experiment setup, instrument operation, and even data-driven decision loops. A headline example is Insilico Medicine's recent introduction of "Supervisor," a bipedal **humanoid robot** integrated into its fully automated drug discovery lab . As a humanoid, Supervisor can navigate a human-oriented lab and manipulate standard equipment – tasks traditional lab robots (which might be stationary or require custom hardware) struggle with. According to Insilico, these humanoid scientist robots are initially assisting with mundane lab chores (pipetting, moving reagents) and acting as an interface for remote researchers (telepresence for lab tours and monitoring) . The broader vision is that such robots, guided by *agentic AI*, will eventually perform entire experimental protocols end-to-end, enabling "lights-out" labs that operate 24/7 without human intervention .

The driver for humanoid robots in labs is largely practical – as Alex Zhavoronkov (Insilico's CEO) noted, most lab instruments are built for human use (with buttons, lids, pipettes sized for human hands) . Retrofitting all equipment for robots is expensive; instead, making robots more human-like in form factor allows them to use existing tools. The trade-off is complexity: humanoid robots are mechanically and computationally complex, and today can be slower than a human. However, AI is improving their capabilities rapidly. Insilico's humanoid is specifically designed to learn from human scientists by observation and to eventually execute delicate lab techniques with precision . This is being achieved by using the robot to gather data (via cameras, sensors) on how human lab techs perform tasks, then training the robot's AI to replicate those skills – essentially an AI apprenticeship.

Beyond Insilico, other companies and consortia are exploring lab robotics. Some pharma labs use mobile robots (like autonomous guided vehicles) to ferry plates and samples between instruments. With AI, these vehicles can navigate dynamic lab environments safely. There are also **robotic arms** integrated with AI vision that can replace some manual operations under sterile conditions. For example, a robotic arm equipped with a camera and deep learning can pick colonies from a culture plate or

load samples into an analyzer, adjusting on the fly if something is misaligned – tasks that earlier robots could not easily do without hard coding. AI also optimizes scheduling in automated labs: when you have many robotic stations, an AI scheduling agent decides which experiment to run in which sequence to maximize throughput (this is akin to an AI air traffic controller for the lab). Lilly's new AI co-innovation lab with NVIDIA emphasizes creating a **continuous learning system connecting "agentic" wet labs with computational dry labs** , effectively describing a setup where AI agents coordinate multiple robotic experiments and analysis in a loop.

In **manufacturing**, robotics and AI are coming together in the push toward Pharma 4.0 (the industry term for smart, autonomous manufacturing). **GMP cleanroom tasks** – such as weighing ingredients, sterile filtration, or filling vials – can be performed by robotic systems to reduce human presence (and thus contamination risk). Some modern biologics facilities use robot arms in isolators for aseptic filling. AI vision systems ensure the robots align correctly and perform quality checks in real time. **Humanoid robots in manufacturing** are still rare, but conceptually they could take on routine inspections or maintenance tasks in facilities that were not originally designed for automation. A recent Barclays analysis predicted that humanoid robots will start appearing in various industries, including pharmaceutical manufacturing, within the next decade as they become more reliable and cost-effective . Tesla's general-purpose humanoid and others are being watched for potential in laboratory logistics or warehouse-style pharma operations. Even if not humanoid, **collaborative robots (cobots)** are increasingly used in pharma plants – these are robot arms with AI safety features that allow them to work alongside humans. They can do tasks like machine tending (loading/unloading machines), packaging, or transport of materials. AI comes into play by giving these robots the ability to detect anomalies (via sensors), adapt to new tasks by learning from demonstration, and optimize their operations over time (through feedback loops).

A particularly interesting application is **AI-coordinated robotics swarms** in large scale bioprocessing. Companies like Sanofi have talked about "digital twin" factories where AI simulations model the entire production line and control a fleet of robotic systems. NVIDIA's Omniverse and similar platforms enable creating a virtual replica of a manufacturing line; AI can test scenarios in the simulation and then guide real robots accordingly . For instance, if a certain buffer solution is running low in a bioreactor, an AI agent could dispatch a mobile robot to replenish it, all while maintaining compliance with sterility – and this could be practiced in VR first to ensure reliability.

In terms of **specific roles**: we see AI-enabled robots acting as **research assistants** (setting up and running experiments). For example, cloud lab companies (like Emerald Cloud Lab or Strateos) use automated systems that scientists can remotely control – AI is used to schedule experiments and manage the robots efficiently so multiple users' experiments run in parallel. Robots handle reagents, liquid transfers, plate reading, etc., guided by software; adding AI means the system can adjust if an experiment deviates (say a culture grows slower than expected, the AI might reschedule subsequent steps or adjust incubation times). The **Supervisor humanoid** at Insilico is a high-profile case, effectively a research assistant that also aids in generating training data for AI models .

For **safety applications**, AI robots can take over dangerous tasks, such as handling biohazardous materials or potent compounds. In a chemical synthesis lab, an AI-controlled robot can run reactions that involve toxic fumes or explosion risk, keeping humans out of harm's way. Also, in manufacturing, robots can perform high-temperature or high-pressure operations in place of humans, with AI monitoring to ensure everything stays within safe parameters.

One key challenge being actively addressed is **validation and compliance** of robotic systems. Any process done by a robot under AI control must be thoroughly validated under regulatory guidelines. This includes ensuring the AI doesn't drift from its intended performance over time (prompting the concept of continuous verification). Companies need to log every action a robot takes for auditability. AI algorithms need to be locked down or change-controlled in regulated environments; a learning robot in a GMP facility can't just update its behavior spontaneously without documentation. This is leading to new regulatory thinking on "Good Machine Learning Practice" for such systems . The FDA's emerging principles on AI emphasize robustness and the ability to explain and justify actions – which extends to robotics (e.g., if a robotic system adjusts a process parameter based on an AI recommendation, there should be a rationale documented).

Humanoid robots in biotech are in their infancy, but the introduction of Supervisor suggests a trend to watch. Early results from Insilico's robotic lab are promising – they reported that their fully automated lab **DORA** and LifeStar systems have been running AI-driven experiments in disease research and even identified an anti-aging compound using a robotic workflow . The addition of the humanoid is expected to further reduce the remaining human touches. The potential is that a combination of humanoid dexterity and AI brainpower could generalize lab automation beyond repetitive tasks to truly flexible, **autonomous research**.

In manufacturing, while we haven't seen a humanoid per se, **AI-driven automation** is advancing. For example, Moderna built a highly automated production line for mRNA vaccines, using AI for quality control decisions – this came to light during the COVID-19 vaccine rollout where robotics helped scale up manufacturing quickly (with AI ensuring each batch met specs through real-time analytics). Companies like Johnson & Johnson have employed mobile robots in warehouses managed by AI systems that learn the facility layout and optimize pick-and-place for pharmaceutical distribution.

Finally, **AI-agent coordination** refers to multiple robots or systems working together under AI orchestration. In a future scenario, one can imagine an AI "lab manager" overseeing a team of specialized robots: one humanoid handling instruments, a robotic arm doing liquid handling, an automated microscope scanning samples, etc., all synchronized by an AI that knows the experiment's protocol and goals. This is analogous to how a human PI coordinates students and technicians in a lab, but here an AI could allocate tasks dynamically (e.g., if one robot finishes early, assign it a new task to avoid idle time). Such coordination is already being tested in some automated labs and is an area of active development.

In summary, AI-enabled robotics in biotech are expanding the possibilities of automation. They promise to increase throughput (experiments and production runs can happen around the clock), improve reproducibility (robots perform the same way every time), and keep humans safer (by taking on hazardous or tedious tasks). Humanoid robots like Insilico's Supervisor symbolize the ambition for **fully autonomous research facilities** . Over the next few years, we expect more labs to implement AI-controlled robotics in incremental steps – perhaps starting with AI scheduling of existing automation, then adding advanced computer vision for QA, and eventually adopting more autonomous robot assistants. The long-term vision is a seamless integration of **physical robots and AI "brains"**, leading to drug discovery and manufacturing that can progress with unprecedented speed and minimal human error. Achieving this will require overcoming technical and regulatory hurdles as discussed, but the industry is clearly moving in that direction.

Image: Insilico Medicine's humanoid lab robot "Supervisor" working alongside human scientists. Such AI-driven humanoids can operate standard lab equipment and perform

tasks like pipetting and reagent handling, aiming to accelerate research in fully automated laboratories .

Recent Examples and Major Trends

The past two years have seen rapid progress in AI applications in biotech, as well as some reality checks. A number of **concrete examples** illustrate both the achievements and the direction of the field:

- **AI-designed drugs entering clinical trials:** Several AI-discovered or AI-designed drug candidates have progressed to human trials, marking important firsts. Insilico Medicine's ISM001-055, targeting fibrosis, became one of the first AI-originated molecules to reach Phase II trials. In early 2024, Insilico published the full case study in *Nature Biotechnology*, documenting how AI identified a novel target and designed a novel small-molecule inhibitor, which then demonstrated efficacy in preclinical models and entered clinical testing . By late 2024, preliminary Phase IIa data showed the drug had favorable safety and some signs of efficacy (improved lung function in IPF patients) . This example is often cited as proof that AI can deliver not just theoretical ideas, but actual drug candidates meeting regulatory standards. Another example is from Exscientia, which had an AI-designed DDR1 kinase inhibitor reach Phase I and an AI-designed immunotherapy (on a known target) reach Phase I as well. These successes, albeit early-stage, have encouraged more investment into AI-driven pipelines.

- **Industry consolidation and partnerships:** A major trend has been the consolidation of AI biotech expertise through partnerships or mergers. In 2025, Recursion Pharmaceuticals (USA) and Exscientia (UK), two leading AI drug discovery companies, announced a merger to combine Recursion's phenomics and data scale with Exscientia's chemistry and clinical design AI . This $700M+ deal (one of the largest in the AI-biotech space) reflects how companies are seeking breadth and end-to-end capability – a single platform that can go from target discovery to molecule design to clinical analytics. Large pharmaceutical companies are also partnering heavily: in late 2023, Roche's Genentech expanded an alliance with NVIDIA to build out a "lab-in-the-loop" concept using NVIDIA's AI platforms . In 2022–2024, almost every big pharma forged some AI partnership: e.g. Pfizer with XtalPi (for AI-enhanced drug solid-form screening) , Pfizer with PostEra (medicinal chemistry) , BMS with Insitro (for fibrosis targets), Merck with Absci (generative biologics), and Bayer with Exscientia (AI research in cardiovascular disease). Tech giants are part of this story too – **Google's**

DeepMind spun out Isomorphic Labs in late 2021 to apply AI (like AlphaFold-derived tech) directly to drug design, and by 2023 Isomorphic Labs had multi-target collaborations including one with Novartis . While Isomorphic has yet to reveal public drug candidates, its partnership implies that pharma trusts AI-born companies with significant projects. Microsoft launched an AI for Health initiative and partnered with Novartis to build AI solutions in a dedicated lab, focusing on uses from discovery to commercialization . Meanwhile, **NVIDIA** has emerged not just as a hardware provider but a key AI software player via its BioNeMo generative AI platform. In 2025, NVIDIA partnered with Novo Nordisk to utilize BioNeMo and AI supercomputing for drug discovery – Novo is using NVIDIA's generative and "agentic" AI tools to model things like single-cell experiment outcomes and to generate novel molecule designs . They are even building large biomedical language models to connect gene, protein, and disease knowledge . This highlights a trend of *AI infrastructure* investment: pharma companies are setting up AI "factories" – dedicated clusters or cloud resources and teams – to embed AI into their core operations. Eli Lilly's January 2026 announcement (co-investing up to $1B with NVIDIA in an AI/robotics lab) exemplifies the scale of commitment now being made .

- **Advances in AI models and tech:** The past two years have also brought technical leaps that the biotech industry is quickly adopting. AlphaFold's impact continued to grow – by 2024, essentially every structural biology or protein engineering group in pharma was using AlphaFold or analogous models (like Meta's ESM) to guide projects. In late 2023, DeepMind further demonstrated **AlphaMissense**, an AI to predict pathogenic mutations, indicating how AI can assist in understanding genetics for therapeutic targeting. Generative AI saw the rise of diffusion models applied to molecule generation (e.g., StablMolecule, etc.), offering more quality and diversity in generated compounds. **Multimodal models** became more accessible – companies like Hugging Face and BioAura released pre-trained models that combine protein sequence and structure data, which smaller biotechs are leveraging to predict protein functions or design antigens for vaccines. On the hardware side, **AI-optimized chips** and cloud services became available specifically for biotech (NVIDIA's DGX systems for pharma, Google Cloud's AI Cloud for Life Sciences, etc.). This democratizes access to powerful AI – a small startup can rent compute to train a model that previously only big pharma could afford. Another trend is the use of **large language models (LLMs)** for knowledge management and even bench work assistance. In 2024, a few biotechs started experimenting with GPT-4 and similar LLMs to help design experiments or summarize findings. For instance, scientists can ask an LLM to suggest possible pathways affecting a disease, using it as an advanced research assistant (with caution due to occasional inaccuracies). Pfizer highlighted integrating generative AI into workflows to automate documentation (like writing clinical study reports drafts) and cited that as a major efficiency gain . We expect that in the next 3–5 years, domain-specific LLMs fine-tuned on biomedical knowledge will become routine tools for

researchers – acting like “co-pilots” during drug design and development documentation.

- **Regulatory developments:** Recognizing the growing role of AI, regulators have started crafting guidelines. In 2024, the FDA (CDER) held workshops on AI in drug development and by 2025 released draft guidance for AI in regulatory submissions . This guidance – and its finalization expected soon – provides clarity that the FDA is open to AI-derived evidence but expects rigorous validation. The FDA is essentially saying: *if you want to use AI to support a drug approval (say to justify a dose selection or as part of a manufacturing change), you need to prove the AI tool is reliable and unbiased.* The European Medicines Agency (EMA) likewise has been discussing AI in its Innovation Task Force. This regulatory attention is a positive trend, because it means by 2026–2027 we likely will have formal pathways to get AI-based methods (like using a machine learning model as a surrogate assay or as a companion diagnostic) approved. It will also set industry-wide standards for GxP-compliant AI (often called Good Machine Learning Practice – GMP for algorithms). In parallel, standards bodies and consortia (like the Alliance for Artificial Intelligence in Healthcare) have been developing best practices on data integrity, algorithm bias testing, etc., which companies are adopting to future-proof their AI processes.

- **Emerging applications and forward-looking trends:** In the next 3–5 years, we anticipate **multi-modal “foundation” models** to play a bigger role. Similar to how in NLP, one big model can be adapted to many tasks, in biotech we might see large models pretrained on enormous biological datasets (sequences, structures, cellular images, medical records) that can then be fine-tuned for specific problems. For example, a foundation model might encode general knowledge of cellular biology, and a company could fine-tune it to predict efficacy of a drug in a specific disease. This could greatly reduce the need for task-specific training data. Another trend is increased use of **digital twins** in both research and clinical settings. In research, that means simulated experiments (a digital cell or tissue model that AI can experiment on before real lab work). In clinical, that means patient digital twins – using AI to model patient outcomes under different treatment options, which could personalize therapy choice. Some hospitals and pharma are piloting AI that can predict individual patient responses by comparing them to large databases (this could guide inclusion criteria or dosing in trials).

We also expect **closer integration of AI across the pipeline**. Up to now, many AI projects were point solutions (one model for target discovery, another for trial recruitment). Companies are now aiming to connect these: data from early research feeding into clinical development models and vice versa (creating a continuous learning across phases). The Lilly-NVIDIA partnership explicitly mentions applying AI beyond

discovery into clinical, manufacturing and commercial realms, integrating robotics and digital twins throughout . In manufacturing, a trend will be more fully autonomous facilities. By 2028 or so, it wouldn't be surprising if a major pharma announces a "dark" (fully automated) manufacturing line where AI monitors and adjusts the process end-to-end with minimal human oversight.

Another forward-looking area is **humanoid and general-purpose robots** becoming more common in labs, as their reliability improves. Right now, Insilico's Supervisor is a unique case, but others are on the way (companies like Sanctuary AI and others are developing humanoid robots targeting lab work and light industry). In a 3–5 year horizon, we might see a handful of pharma or large biotech companies trialing humanoid robots in their high-throughput screening facilities or sample warehouses. Even if not humanoid, the use of **AI for lab automation scheduling and error recovery** will increase – labs will have AI that not only runs experiments but also diagnoses when something goes wrong (e.g., a reagent went bad causing anomalous data) and takes corrective actions or alerts humans appropriately.

Finally, the **economic and workforce implications** are a trend: companies are retraining employees to work alongside AI. The role of a scientist is gradually shifting more towards designing experiments and interpreting AI outputs, with the grunt work delegated to AI/automation . This requires new skill sets, and we see the trend of hiring data scientists in pharma continuing to grow. At the executive level, many pharmas now have Chief Digital or AI Officers (Pfizer, for instance, has a Chief AI Officer spearheading AI integration). This institutionalizes the commitment to AI and ensures it remains a strategic focus.

In conclusion, the biotech industry is **transitioning from the hype phase to the implementation phase** of AI. Major trends like successful AI-designed molecules, big pharma-tech collaborations, more powerful models, regulatory clarity, and advanced automation indicate that AI is steadily becoming embedded in the fabric of drug development. The next few years will likely see the first AI-designed drugs reaching late-stage trials or approval, broader adoption of AI-driven clinical trial design (potentially even an FDA-approved trial designed substantially by AI), and possibly the first instances of **regulatory-approved AI tools** (for example, an AI as a companion diagnostic for a drug, or an AI model as part of a control strategy in manufacturing). Biotech companies that leverage these trends stand to dramatically improve their R&D productivity. Those that don't may find themselves at a competitive disadvantage as the gap widens between AI-enabled "techbio" organizations and traditional

approaches . The convergence of AI with biology is accelerating human therapeutics development toward a more data-driven, efficient, and innovative future.

References

1. Pfizer (2025). *Accelerating Innovation: Pfizer's 2025 AI Festival.* Pfizer News Announcement. – Describes Pfizer's company-wide AI initiative and emphasizes the importance of high-quality data and cultural adoption of AI .
2. Walrath, R. (2025a). *Pfizer doubles down on AI partnership with antibody-drug conjugate deal.* Chemical & Engineering News 103(1), Jan 7, 2025. – Reports on Pfizer's expanded collaboration with PostEra, including results like 40% faster preclinical milestones and an AI-designed antiviral lead achieved in 8 months .
3. Walrath, R. (2025b). *Q&A: What Novartis's biomedical research head sees in AI.* Chemical & Engineering News 103(1), Jan 15, 2025. – Interview with Novartis's Fiona Marshall detailing how Novartis applies AI across drug discovery, including partnerships (Isomorphic Labs, Microsoft, Schrödinger, etc.), and uses in clinical trial design and safety analysis .
4. Roche (2023). *Harnessing the power of AI.* Roche Stories. – Describes Roche/Genentech's perspective on AI in drug discovery, noting the use of ML for predictive, generative, and interpretable models and partnerships like the Recursion collaboration to build cell biology maps .
5. Insilico Medicine (2025). *Insilico Medicine deploys the first bipedal humanoid AI scientist in the fully-robotic drug discovery laboratory.* Press release, Mar 5, 2025 . – Announces Insilico's "Supervisor" humanoid robot and summarizes Insilico's AI platform achievements (30 pipeline assets, 10 INDs, AI-designed drug ISM001-055 reaching Phase II) .
6. Morgan, F. (2025). *The AI drug revolution needs a revolution. npj Drug Discovery*, 4:13 (2025) . – Commentary discussing the need for better translation of AI successes, citing Insilico's 30-month timeline and Recursion's 18-month IND achievement, as well as the observation that AI has sped up R&D but not yet improved clinical success rates .
7. FDA/CDER (2025). *Considerations for the Use of Artificial Intelligence to Support Drug Development.* (Draft Guidance, Oct 2025) . – FDA guidance outlining principles (data quality, model transparency, etc.) for industry using AI in drug development, highlighting the increase in AI-informed regulatory submissions and the FDA's risk-based approach to AI.
8. NVIDIA (2025). *NVIDIA Partners With Novo Nordisk and DCAI to Advance Drug Discovery.* Press release, June 11, 2025 . – Details a collaboration where Novo Nordisk will use NVIDIA's BioNeMo generative AI platform, AI workflow engines, and Omniverse simulations (Gefion supercomputer) to accelerate early research, including generative models for molecules and single-cell data analysis.

9. NVIDIA (2026). *NVIDIA and Lilly Announce Co-Innovation AI Lab to Reinvent Drug Discovery.* Press release, Jan 12, 2026 . – Announces a $1B five-year partnership between Lilly and NVIDIA to build an AI-driven lab with continuous learning systems linking "agentic" wet labs, foundation models for chemistry/biology, robotics and digital twins for manufacturing, and application of AI across discovery, clinical, and operations.
10. Taylor, P. (2025). *AMD shadows NVIDIA with a move into AI drug discovery.* Pharmaphorum, Jan 9, 2025 . – News piece on AMD's $20M investment in Absci and partnership to provide computing for Absci's generative biologic design platform, noting Absci's ability to go from AI design to validated antibody in ~6 weeks and partnerships with multiple pharma (AZ, Merck) leveraging its AI for antibody engineering.
11. Ginkgo Bioworks (2024). *Ginkgo Bioworks Launches Ginkgo Datapoints to Transform AI Model Training in Biology.* Press release, Sep 18, 2024 . – Introduces Ginkgo's data generation service to produce large high-quality biological datasets (e.g. functional genomics, protein data) at scale for training AI models, addressing data bottlenecks by delivering millions of data points in weeks via automated labs.
12. DSM-Firmenich (2024). *Medical Innovations that Save Lives: The Role of AI at DSM-Firmenich Biomedical.* Press release interview, Nov 4, 2024 . – Discusses how DSM-Firmenich's Biomedical division is applying automation, data analysis, and AI to accelerate product development. Emphasizes heavy investment in data management as a foundation for AI, and notes AI's role in analyzing large data volumes to identify promising molecules, optimize supply chains, and improve efficiency (reducing costs and time to market).
13. Future Medicine AI (2025). *Insilico Medicine Introduces Humanoid Scientist to Fully-Robotic Drug Discovery Laboratory.* News article, Mar 5, 2025 . – Coverage of Insilico's humanoid robot deployment in their AI-driven lab, including quotes on bridging human and robotic research (humanoids operating human-designed equipment) and the expectation that combining humanoid robots with AI systems will streamline workflows, reduce human error, and accelerate discovery.
14. C&EN (2026). *Year in Pharma 2025: AI's Influence Grows* (in "C&EN's Year in Pharma 2025" feature). – Summary of how AI shaped pharma in 2025, mentioning milestones like the Recursion–Exscientia merger, increasing number of AI-designed molecules in trials, big tech collaborations, and regulatory developments (cite info consolidated from sources above, e.g., Recursion-Exscientia merger and Insilico Phase IIa result) . (Note: compiled from multiple 2025 events; representative of trend summaries).