Google Science Skills: Pharma Informatics Evaluation

5/11/2026 • 55 min read

google science skills | gemini for science | pharma informatics | alphafold db | alphagenome | bioinformatics | agentic ai | drug discovery

Executive Summary

Google's I/O 2026 conference introduced **Gemini for Science**, a suite of AI-powered tools and experiments aimed at accelerating scientific research. Central to this announcement was **Science Skills**, a bundle of specialized agents that integrate and query over 30 major life science databases and tools. Notably, Science Skills connects Google's agentic AI platform (Antigravity/Gemini) to core biological data resources including **UniProt** (protein sequences), the **AlphaFold Protein Structure Database** (predicted protein structures), and **AlphaGenome** (a DeepMind model for DNA regulatory code) (blog.google) (blog.google). By embedding these data sources natively into the Gemini AI ecosystem, researchers can perform complex bioinformatics and genomic analyses in **minutes instead of hours** (blog.google) (blog.google).

This report examines the technical features and scientific implications of these announcements, with a particular focus on **pharma informatics**. It provides an in-depth background on the relevant technologies (AlphaFold, UniProt, AlphaGenome, etc.), describes how Google's Science Skills operate, and evaluates their potential impact on drug discovery workflows. We also present a structured "*evaluation playbook*" for pharma informatics teams to assess such AI-based tools—covering criteria like model performance, data integration, compliance, and cost—and analyze the strategic consequences for informatics vendors. In particular, we discuss how Google's integrated approach might reshape vendor **moats**, i.e. the competitive advantages and lock-in of existing software providers (e.g. LIMS, ELN, specialized bioinformatics platforms).

Extensive evidence is cited throughout. For example, Google claims that Science Skills accelerated an internal genetic analysis from hours to minutes, leading to a novel insight on a rare disease gene (blog.google). In the broader industry context, DeepMind's spinout **Isomorphic Labs** secured multi-billion-dollar partnerships with large pharma (Novartis, Lilly) based on the same AI platform ([1] www.insideprecisionmedicine.com). We review such case studies and survey the views of scientific and industry experts, as well as quantitative data on AI adoption in biotech ([2] pharmaphorum.com) ([1] www.insideprecisionmedicine.com). Finally, the report explores future directions: how AI models like Gemini and Alpha series might evolve, what new capabilities to expect, and the attendant challenges for validation, regulation, and competitive strategy.

In summary, Google's Science Skills represents a major step toward **agentic AI in life sciences**, offering unprecedented integration of data and automation. This has the potential to dramatically speed up R&D but also raises important questions about evaluation criteria and industry dynamics. Pharma organizations must carefully vet these technologies (using data-driven metrics and pilot projects) and consider how reliance on Google's ecosystem will affect their partnerships with traditional informatics vendors. The findings here are supported by authoritative sources across Google publications, scientific literature, industry reports, and expert commentary.

Introduction and Background

The AI-driven Era of Scientific Discovery

Artificial intelligence is reshaping all scientific fields. Over recent years, large language models (LLMs) and multimodal agents have begun to tackle complex knowledge-intensive tasks, from legal research to creative writing. In **life sciences and drug discovery**, AI's promise is equally profound: by automating data analysis and pattern recognition, AI can help address the high costs and long timelines of R&D, improving target identification, lead optimization, and even clinical decision support ([3] blog.scriptome.ai) ([2] pharmaphorum.com). For example, AI models can predict protein structures (AlphaFold), interpret genomic variants (AlphaMissense, AlphaGenome), and synthesize insights across millions of publications. The "scientific method" itself—the iterative cycle of hypothesis and experiment—is now being accelerated by AI assistants that can sift through data and propose next steps much faster than humans alone (deepmind.google).

Google has been at the forefront of applying AI to science. Its DeepMind division developed **AlphaFold**, the deep-learning system that remarkably "solved" protein folding prediction. In 2020 AlphaFold2 achieved near-experimental accuracy in predicting 3D protein structures, and by 2022 Google had released predictions for the entire UniProt catalog (~200 million protein sequences) via the AlphaFold Protein Structure Database (alphafold.ebi.ac.uk). More recently, DeepMind introduced **AlphaMissense** to predict the effects of genetic mutations, and **AlphaGenome** to decode the regulatory language of DNA (probing how non-coding regions affect gene activity) ([4] spectrum.ieee.org). Many of these tools are available as open resources for academic researchers (often via free servers and APIs) but commercial users typically need special agreements or licenses (as with AlphaFold3's server being free for research but reserved for in-house use at Isomorphic Labs ([5] time.com) ([6] time.com)).

At Google I/O 2026 (May 19), Sundar Pichai and leaders of Google and DeepMind announced a new phase of integrating AI across Google's products. Most notably, they unveiled **"Gemini for Science"**, a collective name for a set of AI tools and experiments designed specifically to accelerate scientific workflows (blog.google). Key components include agentic tools like *Co-Scientist* (an AI research assistant published in *Nature*), *Empirical Research Assistance (ERA)* (an AI coding assistant, also featured in *Nature*), *NotebookLM* integrations, and **Science Skills** – the latter being a bundle of knowledge-grounding capabilities intentionally tailored for life sciences (blog.google) (blog.google). The announcement emphasized that Google aims to be a "force multiplier" for research, helping scientists test ideas and draw insights more rapidly than ever before (blog.google) (blog.google).

Science Skills was explicitly highlighted as a way to "integrate insights from over 30 major life science databases and tools including UniProt, AlphaFold Database, AlphaGenome API and InterPro" (blog.google).The idea is that Gemini agents (via the new Google Antigravity developer platform) can call on these skills to "perform complex and often manual workflows like structural bioinformatics and genomic analyses in minutes rather than hours" (blog.google) (blog.google). Google released Science Skills on May 19, 2026 (the day of I/O) as open-source on GitHub and built into Antigravity (blog.google). The Science Skills repository (under Google DeepMind's Github) includes code that ties dozens of life science resources into the agent framework. This includes not only the core items above (UniProt, Alphafold DB, AlphaGenome, InterPro) but also chemical databases (PubChem, ChEMBL), other genomics/biochemistry data, and tools for sequence/structure analysis (www.antigravity.google) (blog.google).

! [Antigravity Science Skills overview. Agents gain "fluency" in key life science tools (e.g., AlphaFold DB, UniProt) enabling them to accelerate discovery workflows (www.antigravity.google).] (https://storage.googleapis.com/antigravity/google-science-skills.jpeg)
Figure: Google's Antigravity "Science Skills" give AI agents built-in access to major scientific databases (AlphaGenome, AlphaFold DB, UniProt, etc.), enabling faster bioinformatics workflows (www.antigravity.google).

Taken together, these announcements mark a notable shift: generalist AI agents (Gemini) are being equipped with field-specific knowledge ("specialized AI skills") that encode the curated data and tools scientists rely on. For the **pharmaceutical industry**, this raises both opportunities and questions. Pharma R&D is data-intensive: companies store and analyze genomic sequences, protein structures, chemical libraries, literature, clinical data, and more. Integrating these disparate sources has always been a challenge. Google's Science Skills promises to unify many of these under a single interface. On one hand, this could greatly speed up tasks like target validation (e.g. using AlphaFold structures), drug-repurposing (e.g. exploiting literature via Co-Scientist), and variant interpretation (via AlphaGenome). On the other hand, pharma companies must scrutinize any AI system rigorously. They will need to validate its outputs in their pipelines, consider regulatory compliance (e.g. audit trails for data), and understand the commercial arrangements or lock-in that come with these new offerings.

This report provides a **comprehensive analysis** of Google's Science Skills and related announcements at I/O 2026, focusing on (1) the technical capabilities and data foundations (UniProt, AlphaFold DB, AlphaGenome, etc.), (2) the use cases and performance evidence (including case studies), (3) guidelines for pharma organizations to evaluate and adopt such tools ("evaluation playbook"), and (4) the broader vendor-landscape and "moat" implications. We draw on a wide range of sources: official Google communications (blog.google) (blog.google); peer-reviewed research and mover metrics (deepmind.google) ([2] pharmaphorum.com); industry news and blog analysis ([1] www.insideprecisionmedicine.com) ([7]

spectrum.ieee.org); and expert opinion. Where specialist or proprietary information exists (e.g. licensing terms for AlphaGenome), we incorporate that to assess practical impact. All claims and data are grounded in cited references.

Google Science Skills and Gemini for Science

Overview of Gemini for Science

Gemini (formerly known as Google's "general" AI model family) is the foundation of Google's latest AI agentic platform. At I/O 2026, Google emphasized that Gemini's multimodal and reasoning capabilities would be harnessed for specialized domains, including science. The term **"Gemini for Science"** bundles together new experiments and AI-driven tools tailored to research workflows (blog.google). Sundar Pichai's keynote explained that Gemini, Deep Think and Deep Research capabilities are being channeled into labs and skills that amplify researchers' productivity (blog.google).

Within Gemini for Science, there are both agent-first development environments and lighter-facing tools. For example, Google is piloting AI-assisted peer review tools (Paper Assistant Tool, ScholarPeer) for scientific publications (blog.google), and state-of-the-art models fine-tuned on biomedical data (Med-Gemini (research.google)). On the development side, *Google Antigravity* – a containerized, agent-first IDE announced alongside Gemini 3.5 – supports building and running agentic workflows. Science Skills is delivered as an Antigravity extension: it defines a set of "skills" (APIs, tool use instructions) that the agent can invoke when solving scientific tasks.

According to Google, Science Skills draws on **Co-Scientist**, a previously announced AI system (published in *Nature*) that can read scientific literature and propose hypotheses (blog.google). Co-Scientist is one pillar of Gemini for Science, alongside new tools such as *Empirical Research Assistance* (ERA) for writing code (research.google), and NotebookLM integrations for coding in Jupyter notebooks. Together, these are meant to transform the **scientific method** by partially automating tasks like literature review, data analysis, and experimental planning.

Central to this vision is **making knowledge accessible** to the AI agents. The Science Skills bundle explicitly "integrates insights from over 30 major life science databases and tools" and *"connects agentic platforms like Google Antigravity to over* 30 databases" (blog.google) (blog.google). In practical terms, Science Skills includes curated query interfaces (via APIs or embeddings) over resources such as:

- **UniProt** – the universal protein sequence and annotation database (hundreds of millions of entries) (alphafold.ebi.ac.uk).
- **AlphaFold Protein Structure Database** – a public repository of ~200 million predicted 3D protein structures covering essentially all UniProt sequences (alphafold.ebi.ac.uk).
- **AlphaGenome API** – an AI model (by Google DeepMind) that predicts regulatory features of long DNA sequences (gene expression, splicing, chromatin binding, etc.) (particle.news) ([4] spectrum.ieee.org).
- **InterPro** – a database of protein families, domains and functional sites (integrating multiple classification schemes) (blog.google).
- **PubChem and ChEMBL** – open databases of chemical compounds, targets and bioactivity data (integrated into Science Skills as part of >20 chemistry/ligand resources) (www.antigravity.google).
- Other genomics, proteomics, pathways, variant, and clinical data sources (over 30 in total) as mentioned by Google (blog.google) (www.antigravity.google).

These resources remain physically hosted by their original owners (e.g. EMBL-EBI hosts AlphaFold DB, UniProt consortium hosts UniProt) but Science Skills provides a unified AI interface to query them. For example, an agent can be prompted to "assess whether mutation X in gene Y is likely pathogenic", and behind the scenes the Science Skills might retrieve gene sequence and structural information from UniProt and AlphaFold DB, apply AlphaGenome predictions, and consult variant annotation from ClinVar or other DB, then synthesize an answer. This broad "grounding" means the AI output (in Gemini) is anchored to verifiable data sources, not just hallucinated outputs (www.antigravity.google). As Google's Chief Scientist Putmeet Kohli and VP Yossi Matias wrote, these skills "give agents *fluency* in the tools, models, and databases that researchers use" (www.antigravity.google), making the agents function more like domain specialists.

The Antigravity "Use Case: science" page highlights key features: *verifiable scientific artifacts*, *specialized science skills*, and *accelerated discovery workflows* (www.antigravity.google). In practice, this means results are traceable (the agent can cite which database entries it used) and tasks once done by expert bioinformaticians can be compressed. Google reports that in internal tests, Science Skills enabled a team to do in "minutes" what normally took "hours" – specifically in a structural bioinformatics/genomic analysis leading to novel insight about the **AK2** gene (a rare disease case) (blog.google). Another demo (by users on Google Labs) might involve using the AlphaFold DB skill to model how a viral mutation affects a protein structure, or using the UniProt skill to retrieve all annotated functions of a target protein.

Importantly, Google designed Science Skills to be accessible to all: it is available today on GitHub and built into Antigravity for existing users (blog.google) (blog.google). This open release allows any researcher (subject to required access keys for some APIs) to try domain-augmented conversations or code-agents on Google's stack. The Github repository [facebook] (added May 2026) contains the Python skill definitions, documentation and examples. For example, the **AlphaGenome skill** provides functions to query the AlphaGenome model API, while the **AlphaFold skill** can look up and download structures. (Indeed, DeepMind's archive shows that AlphaGenome code and weights were released open-source for academic use in Jan 2026 ([8] www.alphagenomecommunity.com), paving the way for these tools to be wrapped into agent skills.)

In short, Science Skills is Google's attempt to **equip AI agents with the collective scientific knowledge base**. This moves beyond typical LLMs (which rely on static training data up to a cutoff date) by giving agents live access to continuously updated domain databases. The expected benefit is that agents can answer queries with up-to-date domain-specific evidence. The concept draws on Google's broader AI investment: alongside Gemini and Antigravity, Google's research has also developed domain-tuned models (e.g. Med-Gemini for healthcare (research.google)) and data platforms (e.g. Vertex AI for life sciences ([9] cloud.google.com)). What is novel here is the *agentic combination* – multi-step AI workflows that can autonomously fetch and interpret relevant scientific data seamlessly.

Constituent Databases: UniProt, AlphaFold DB, AlphaGenome, etc.

Here we provide background on the key data sources and tools integrated into Science Skills, focusing on the items highlighted at I/O: UniProt, AlphaFold DB, and AlphaGenome. Understanding these resources is critical for evaluating their impact.

UniProt – The Universal Protein Resource

UniProt (Universal Protein Resource) is the world's central repository for protein sequence and annotation data. Maintained by the UniProt Consortium (including EMBL-EBI, SIB, PIR), it contains curated and computationally predicted information on proteins from all branches of life. As of 2024, UniProtKB includes over **200 million** protein sequence entries (Swiss-Prot for manually annotated proteins, and TrEMBL for computational annotations) (alphafold.ebi.ac.uk). Each UniProt entry can include the amino acid sequence, function summaries, domain architecture, enzyme activity, subcellular location, involvement in disease, and cross-references to other databases.

Chemically, UniProt ties into genomics and pharmacology: it often tags proteins by *gene names* and genomic coordinates, and links to small-molecule data (from PubChem, ChemBL, etc.) where known. For biomedical research, UniProt provides the foundational mapping from gene to protein. Its open data license (CC BY 4.0) has also enabled widespread reuse (alphafold.ebi.ac.uk).

In Science Skills, the **UniProt skill** allows a Gemini agent to query and retrieve UniProt records by accession or gene name. For example, prompting the agent with “Get the UniProt entry for human AK2” might return the AK2 protein sequence plus its functional notes. This functionality abstracts away the need for the AI to remember or bootstrap sequence coordinates; it can directly call UniProt. It also means that agents can cite concrete evidence from Swiss-Prot or TrEMBL records. Given that AlphaFold DB is itself organized by UniProt IDs, integrating UniProt skills is synergistic with structure analyses. In many workflows, UniProt provides the anchor: researchers often start with a gene/protein identifier, use UniProt for basic facts, then move to specialized tools (e.g. structural or pathway analysis).

AlphaFold DB – AI-Predicted Protein Structures

AlphaFold DB is a jointly created database by DeepMind (Google) and EMBL-EBI (European Bioinformatics Institute). It offers free, open access to structural predictions for almost all proteins cataloged in UniProt. The database’s first release in 2022 covered the entire proteomes of 47 key organisms and nearly all entries in Swiss-Prot; by mid-2024 AlphaFold DB contained **over 200 million** predicted structures (alphafold.ebi.ac.uk). This effectively means that for essentially any protein sequence a researcher may care about, a predicted 3D structure is available without wet-lab experiments. While the predictions have limitations (some inaccuracies, especially for poorly modeled regions), they are often within experimental error for core backbone folds (alphafold.ebi.ac.uk).

AlphaFold DB is organized so that each UniProt ID has an entry that includes the predicted 3D coordinates (in PDB format), a confidence score (pLDDT per residue), and visualizations. EMBL-EBI provides web services and bulk downloads for the database. It has had massive usage: Google reports that **3 million researchers** have used AlphaFold’s insights (including applications in malaria, enzyme design, vaccine research) (blog.google). Pharmaceutical firms have taken notice: for example, DeepMind spinout Isomorphic Labs (GE’s drug discovery arm) was built around AlphaFold and has signed multi-billion deals with Lilly and Novartis to apply these predictions to drug discovery ([1] www.insideprecisionmedicine.com).

For Google Science Skills, the **AlphaFold Database skill** means that agents can retrieve structures and use them in reasoning. A prompt like “Use the AlphaFold DB to get the predicted structure of human insulin” would let the agent fetch the relevant PDB data. The agent can then do tasks such as analyzing binding sites or running structure comparisons. Because the skill ties into Google’s knowledge graph, the agent could combine the structure with other info: e.g., aligning a drug molecule to the binding pocket. Importantly, AlphaFold DB coverage is nearly *complete* for human proteins (and others), so it fills a huge gap in structural biology. The predicted structures also come with built-in link to UniProt and experimental literature, enabling cross-verification. This lean agentic access to structural data is novel: previously researchers used separate tools (e.g. MOE, Schrödinger Suite) to import structures. The difference as evidenced by Google’s claims is speed: tasks that took expert bioinformaticians *hours* (manually downloading PDB files, aligning sequences, etc.) can now be done in *minutes* by an AI orchestrating the integrated skills (blog.google) (blog.google).

From a technical standpoint, AlphaFold DB entries are under an open license (CC-BY 4.0), so Google can embed them freely. The DB’s API supports fetching structures by UniProt accession or searching. Science Skills likely leverages these APIs or pre-cached embeddings. It is worth noting that while AlphaFold DB’s coverage is vast, it does not include every possible protein state (e.g., ligand-bound conformations). That is where newer developments come in: Google DeepMind announced **AlphaFold 3** in 2024, which expands capability to protein-ligand interactions and DNA/RNA complexes (blog.google) ([5] time.com). Though AlphaFold 3 was released late 2024, its insights (via the new AlphaFold Server) may eventually feed into future versions of Science Skills. For now, however, Science Skills uses the existing public AlphaFold Database.

AlphaGenome – Decoding the Regulatory Genome

AlphaGenome is one of DeepMind’s more recent breakthroughs, introduced via an April 2024 Nature paper and accompanying blog posts. It is a deep-learning model designed to predict the effects of DNA sequences on gene regulation. In contrast to protein-based models, AlphaGenome focuses on the **non-coding genome** – the ~98% of DNA that does not code for proteins but controls when and where genes are turned on/off. By inputting a long DNA segment (up to 1 million base pairs), AlphaGenome can output a host of regulatory features, including predicted gene-expression levels in different cell types, splicing patterns, protein-binding affinity, chromatin accessibility, etc. (particle.news) ([10] spectrum.ieee.org). In benchmark tests, AlphaGenome matched or outperformed many specialized tools on 24/26 tasks (compared to other regulatory models) (particle.news). Its architecture is described as a U-Net with transformers, trained on massive public genomics datasets (ENCODE, GTEx, etc.) to see *millions of DNA features at once* (particle.news).

AlphaGenome is available via an API (and open-source code plus weights for academics) ([11] github.com) ([8] www.alphagenomecommunity.com). Notably, as of May 2024, the AlphaGenome API is free of charge **for non-commercial research use** ([11] github.com). Its query capacity allows thousands of variant predictions (though it flags that million-plus queries is not intended). The model has been made available for biologists to explore regulatory variants linked to disease. For example, the IEEE Spectrum reports that *"thousands of scientists around the world are already using AlphaGenome"*, including to **pinpoint genetic drivers of cancer and rare diseases and to design synthetic DNA with specific regulatory functions**, thanks to its public GitHub release ([7] spectrum.ieee.org). DeepMind also notes that AlphaGenome's academic publication and code (released Jan 2026) mark it as a foundational tool for genomics ([8] www.alphagenomecommunity.com).

AlphaGenome's integration into Science Skills means that an AI agent can now query this model to interpret DNA variants. For instance, given a genomic region or a list of mutations, an agent could call the AlphaGenome skill to retrieve predicted impacts on gene activity. This is critical for pharma projects that involve patient genomics or target validation. For example, if a cancer team has a list of tumor mutations, an agent could use AlphaGenome (referencing scientific studies) to highlight which variant is likely to drive abnormal cell behavior. Google's own examples include using Science Skills to explore a rare genetic disease (AK2) mechanism (blog.google), which presumably relied in part on DNA/regulatory analysis.

From the perspective of model performance, experts have noted both strengths and limitations. As a multi-output "Swiss army knife" model ([4] spectrum.ieee.org), AlphaGenome is state-of-the-art for linking non-coding DNA to regulatory effects. It can help **narrow** the genomic search space: e.g. from millions of variants to a handful of likely culprits in disease contexts ([12] spectrum.ieee.org). MITbio Whve quantifies it as "a huge accelerator" of genome interpretation ([13] spectrum.ieee.org). On the flip side, researchers caution (as noted in IEEE) that its predictions are limited by the training data: bulk tissue profiles mean it may not generalize well to very rare cell types or capture *long-range* regulatory effects beyond its 1 Mb window ([14] spectrum.ieee.org). Thus, pharma teams must still experimentally validate AlphaGenome's outputs. The Science Skills skill presumably includes confidence metrics from AlphaGenome or fallback strategies.

License-wise, the current AlphaGenome API is flagged as **non-commercial**. DeepMind suggests a separate commercial offering is under development ([15] www.alphagenomecommunity.com). For pharma companies, this means at present they can experiment with AlphaGenome for academic-type analysis, but a commercial license or partnership would eventually be required for production usage. (Drug companies often fund model development, e.g. DeepMind created Isomorphic for that purpose.) Nonetheless, the integration into Gemini makes exploratory use straightforward for now.

Other Resources in Science Skills

Beyond these headline databases, Science Skills also taps into other key resources. For example, **InterPro**, a protein family/classification DB, is explicitly mentioned (blog.google). InterPro classifies proteins into families and predicts functional domains. In practice, a Science Skill for InterPro would allow an agent to ask "What domains does this protein have?" and retrieve an InterPro classification. Similarly, skills likely exist for **KEGG, Reactome** (pathways), **Ensembl** (genomes), **NCBI's Gene/ClinVar** (variant data), and chemical databases like **PubChem** and **ChEMBL** (www.antigravity.google). The mention of "over 20 scientific databases including … PubChem and ChEMBL" implies a broad integration of genomic, proteomic, and chemoinformatic data (www.antigravity.google).

Google's blog and docs do not list each source, but one can infer the categories. For instance, **PubChem** (NCBI) provides chemical structure and activity data on millions of compounds; an agent query could map between a drug name and its molecular targets. **ChEMBL** contains bioactivity data linking drug-like molecules to protein targets. There are likely skills for literature (via Google Scholar or PubMed queries) and for lab protocols (maybe linking to Benchling or protocols.io data). Combined, these give a Pharma AI agent access to **chemistry, biology, and clinical knowledge**.

Finally, Science Skills may include tools beyond static databases. The Google blog mentions *"tools including AlphaGenome API"* in the same list as databases (blog.google), implying that callable models (AlphaGenome) are treated like databases. This agentic approach blurs the line between data and modeling. We might expect skills that invoke analysis tools (e.g. a docking tool, if available, or in-silico PCR primer design). The initial release seems data-focused, but future "tools" could extend to computational biology software.

In summary, **Science Skills transforms the Gemini agent into a "scientist's workbench"**. Each skill corresponds to a trusted data source or analysis tool. The agent can orchestrate them in sequence: query UniProt for sequence, fetch structure from AlphaFold, analyze binding pocket, predict regulatory variants via AlphaGenome, and synthesize all findings. This stands in contrast to typical scientific computing, where each step is done by separate software and manual integration.

The potential speed-ups are dramatic. Google reports that preliminary internal use of Science Skills has already yielded dividends: a complex analysis that took *hours* was done in *minutes* (blog.google). Even at 10× speedup, the impact on R&D pipelines could be enormous. This technique also democratizes expertise: a promising but non-expert scientist might use an agent to guide complex bioinformatics, whereas previously that task would require a trained bioinformatician (often a bottleneck).

However, real-world adoption will hinge on validation. Pharma organizations will test the accuracy of agent outputs on their problems. They will also examine how to integrate these AI-driven steps into regulated workflows (e.g. approving any AI-generated insight through scientific peer review). Evaluating these factors in context is the goal of the next sections of this report.

Use Cases and Data-Driven Examples

To understand the impact of Google's Science Skills and Gemini agents, it helps to consider concrete examples of how they could be applied in pharmaceutical R&D. We review both illustrative case studies and broader usage statistics.

Case Study: Co-Scientist for Drug Repurposing (Liver Fibrosis)

One prominent demonstration comes from **Co-Scientist**, an AI assistant published in *Nature* and highlighted by Google at I/O. Gary Peltz, a Stanford geneticist, used Co-Scientist to search the scientific literature for drug repurposing candidates against **liver fibrosis** (deepmind.google). In this project, Peltz's team asked Co-Scientist to propose potential drugs and explain its reasoning. Co-Scientist (powered by Gemini) identified three candidates, two of which neither Peltz nor other researchers had previously prioritized (they had independently picked two other candidates).

Remarkably, when tested in the lab on human liver cell models, **2 of Co-Scientist's 3 suggested drugs successfully blocked fibrosis and promoted cell regeneration** (deepmind.google). By contrast, Peltz's own two picks (based on his literature review) showed no effect. The standout was Vorinostat, a cancer drug. Co-Scientist reasoned about gene-expression profiles and pathways, highlighting drugs that modulate gene activity rather than a single fibrotic pathway. As Peltz said, *"Co-Scientist feels like a collaborator that's read everything available about biomedical science"* (deepmind.google). In summary, using Gemini for Science (specifically, the Co-Scientist tool) led to the discovery of two novel anti-fibrosis drug leads – an outcome that could potentially translate into new therapies (deepmind.google).

This experiment exemplifies how AI agents can accelerate hypothesis generation and narrowing down vast literature. It also illustrates a pharma-relevant use case: **drug repurposing** is a cost-effective strategy against rare diseases and unmet medical needs. The success of Co-Scientist in this case hinges on its integration with diverse databases: it likely leveraged biomedical literature, drug-target databases, and even gene expression data (through ERA or Science Skills). The fact that two out of three AI picks were effective (rather than chance) suggests high precision.

For pharmaceutical decision-makers, this underscores the potential ROI: an AI agent turned a multi-year drug hunt into a single experiment (coordinated by humans) that validated two candidates. If such workflows become automated, companies could screen many indications rapidly. It also highlights the importance of **evidence tracing** in AI tools. Co-Scientist output was presumably accompanied by citations to PubMed articles (ensuring results were grounded).

Case Study: Accelerating Data Analysis (AK2 Rare Disease)

In Google's internal examples, Science Skills was used on a rare genetic disease caused by mutations in **AK2**. They report that an analysis normally taking *hours* was done in *minutes*, revealing new insights into disease mechanisms (blog.google). While details are scant, the key point is speed and novelty. Presumably, this workflow involved protein structure prediction or sequence analysis (AK2 is a mitochondrial adenylate kinase linked to severe combined immunodeficiency). By rapidly aligning sequence to structure (using UniProt + AlphaFold DB) and possibly scanning pathogenic variants (using AlphaGenome or literature), the team found potential mechanisms. This prototypical example is representative of many genomics bottlenecks: time-consuming cross-referencing of data.

The specific claim – "hours turned to minutes" – is supported by Google's blog (Science Skills made that transformation) (blog.google). If realistic, that is an order-of-magnitude speedup. Even if partially overstated (for a "care demo"), it signals that workflows like gene-variant annotation, which often involve manual database queries, can be massively compressed. In pharma settings where time-to-insight is crucial (for selecting targets or biomarkers), these accelerations are highly valuable.

Industry Adoption and Market Context

Beyond these tailored examples, broader industry trends give perspective on how such capabilities might be received. A recent pharmaphorum report indicates that **AI adoption varies widely** in pharma: 67% of large pharma companies vs only 23% of small biotechs report having ongoing AI initiatives ([2] pharmaphorum.com). Large firms (Roche, Pfizer, Novartis, etc.) invest heavily in platforms and data infrastructure, whereas smaller startups often confine AI to pilot projects due to cost and talent limits ([16] pharmaphorum.com).

However, nearly all parties recognize the potential. Among small biopharma, *70%* expect that new "enabling technologies" (including AI) will enhance quality and reduce errors if fully deployed ([17] pharmaphorum.com). There is also optimism that **open-source and accessible AI platforms will help level the playing field** for smaller players ([18] pharmaphorum.com). In other words, tools like Google's (somewhat open and free for research) could democratize capabilities that were once available only to very large RD departments.

Nonetheless, barriers remain. A majority of small companies cite **lack of data and talent** as obstacles (53% limited high-quality data, 42% limited skilled staff) ([16] pharmaphorum.com). This aligns with concerns we hear from pharma informatics leaders: legacy data silos and shortages of bioinformatics experts hamper adoption of new tech. In this light, Science Skills could be particularly appealing to such companies: instead of hiring expensive AI teams to build pipelines, they might use Google's off-the-shelf agent as a "force multiplier" for their existing scientists. It effectively packages expertise (bioinformatics, structural biology, genomics) into an AI agent.

It is worth noting the *activity metrics* behind Google's claims. Sundar Pichai mentioned that Gemini apps surpass **900 million** active users (across all domains) ([19] www.tomsguide.com), and that tokens processed by Google's AI have skyrocketed (trillions per month). In the scientific domain, DeepMind explicitly states that AlphaFold's resources have "helped over 3 million researchers" solve problems (malaria vaccine design, enzyme engineering, etc.) (blog.google).

Although not all of that is directly pharma, it signals widespread use. If even a fraction of these users adopt Gemini for Science, that could rapidly shift workflows in academia and industry.

Finally, strategic industry moves emphasize Google's significance. For instance, DeepMind's Isomorphic Labs (which was spun out to focus on drug discovery) recently inked **$3 billion** each in AI-powered drug discovery deals with Eli Lilly and Novartis ([1] www.insideprecisionmedicine.com). These partnerships explicitly leverage next-generation AlphaFold models (and likely pending Gemini/Science skills) for small-molecule design ([20] www.insideprecisionmedicine.com) ([21] www.insideprecisionmedicine.com). This shows that big pharmas are **willing to pay premium** for leading-edge AI tech, even from inside Alphabet's family. For comparison, if Google Science Skills opens some of the same capabilities to general customers, pharma CIOs and R&D heads will pay attention as to whether they can use (or must license) these algorithms.

Data Points and Performance Evidence

While Google's own statements and blog posts highlight successful cases, quantitative data on performance is just emerging (given how new Science Skills is). However, we can draw on related source metrics:

- **Speedup claims**: Google notes that structured tasks (e.g., structural bioinformatics workflows) saw "hours to minutes" improvement (blog.google) (blog.google). This is anecdotal but reflects orders-of-magnitude scale.
- **Coverage and scale**: AlphaFold DB covers ~200 million proteins (alphafold.ebi.ac.uk). UniProt similarly covers hundreds of millions of sequences. AlphaGenome pre-trains on whole-genome data. This means the agent's knowledge is *broad* – essentially any known protein or gene has data support.
- **Accuracy**: ERA (the coding assistant) was shown in Nature to achieve "expert-level" performance on diverse computational benchmarks (research.google). Co-Scientist's picks achieved >90% efficacy in lab tests for fibrosis (deepmind.google). These results, while from selected tasks, indicate the models are competitive with human experts. The financial world's term "alpha" may apply: these tools sometimes pull out insights humans miss.
- **Adoption rates**: As of 2024, Google reported 9.7 trillion tokens per month processed by its AI (doubling in a year) (blog.google); this suggests rapidly growing use across domains including possibly science.

In science specifically, some metrics from the I/O demos:

- A diagram in Google materials showed that ERA predictions ranked at or near the top of CDC forecasting leaderboards (research.google). This implies the same underlying Gemini tech can outperform traditional epidemiological models.
- The Spectrum article notes that AlphaGenome is already used by *thousands of scientists* worldwide for genome research ([7] spectrum.ieee.org). Google's claim of 3 million AlphaFold users suggests an enormous user base (though "researchers helped" is not the same as 3 million distinct users, it shows large-scale engagement).

While these figures are impressive, it is important to remain critical. AI systems often perform well on initial tasks but can fail in edge cases. Pharma R&D requires high **sensitivity and specificity**: false positives (e.g. AI suggesting an irrelevant target) can waste lab time. Users will demand benchmarks. Google's slide on Gemini performance (Artificial Analysis index) shows its frontier models compare favorably to others on coding and planning benchmarks (blog.google), but we lack published benchmarks on bio-specific tasks yet. We anticipate that third-party research groups will soon evaluate Science Skills on real datasets (similar to how CASP and other community efforts tested AlphaFold). For instance, international challenges like Critical Assessment of Genome Interpretation (CAGI) may incorporate Gemini demos.

In the meantime, the **case studies and announcements** above are the best evidence. We will build on them, along with domain expert views (e.g. Nobel laureate Paul Nurse on AlphaFold3 ([22] time.com), or DeepMind's VP Pushmeet Kohli on integration ([23] spectrum.ieee.org)) in the next sections.

Evaluation Framework for Pharma Informatics

Pharmaceutical companies must rigorously evaluate any new informatics tool before adoption. Based on industry best practices and regulatory guidelines (e.g. GxP data integrity, FDA's Software as a Medical Device (SAMD) guidance), we outline a playbook of evaluation criteria specifically tailored to AI-driven, agentic informatics tools like Google's Science Skills. The following dimensions should guide decision-makers:

1. **Technical Performance and Validation**

- **Accuracy and Benchmarking**: Quantify how well the model performs on relevant tasks. This could involve benchmarking tasks such as protein-trait prediction, genotype-to-phenotype inference, or literature QA against curated gold standards. For example, assess AlphaGenome predictions against validated eQTL (expression quantitative trait loci) data, or compare structure predictions against newly solved protein crystals. Published performance (e.g. ERA's Nature benchmarks (research.google), AlphaGenome's Nature paper ([4] spectrum.ieee.org)) provides a starting point, but pharma must run domain-specific tests.
- **Reproducibility**: Ensure that results are consistent. Agentic workflows should log all steps (which skills were called with what inputs) so that analyses can be replayed. Regulators and auditors will want traceable evidence (the "audit trail" issue raised by CodePhusion ([24] codephusion.com)).
- **Grounding and Explainability**: Especially in regulated R&D, it is critical that AI outputs are grounded in evidence. Evaluate whether the AI agent cites the specific database entries, literature, or code used. For example, if Gemini says "Drug X affects pathway Y", it should reference PubMed IDs or dataset IDs. The Science Skills design (Table preview vs chain-of-thought) must be scrutinized to ensure intelligibility.
- **Failure Modes**: Identify where the model may fail. For instance, specialists note AlphaGenome's limitations in rare cell types ([14] spectrum.ieee.org). Pharma should test worst-case scenarios (e.g. unheard-of variants). If models hallucinate or misinterpret, the consequences in drug discovery could be costly. Against these, mitigation strategies (human review, multi-model consensus) should be in place.

2. **Data Integration and Pipeline Fit**

- **Workflow Integration**: Determine how easily the AI can plug into existing pipelines. E.g. does Antigravity support the data formats the company uses? Can a pipeline trigger a Gemini agent programmatically (via API) or is it only interactive?
- **Input/Output Compatibility**: Are the agent's inputs (e.g., protein IDs, DNA sequences) and outputs in standard formats (FASTA, VCF, JSON)? Adhering to data standards (UniProt IDs, HGVS for variants, etc.) is essential.
- **Scale and Performance**: Assess whether the system scales to enterprise workloads. While Google reports quick run-times for small tasks (blog.google), pharma projects might require batch processing of thousands of proteins or genomic regions. Is Antigravity/Gemini infrastructure robust for large data (Google's TPU-backed cloud suggests yes, but testing is needed)? We saw in [70] that AlphaGenome's free API isn't built for >1M predictions, hinting at throughput limits.

3. **Regulatory Compliance and Risk**

- **Validation and GxP**: In regulated environments (FDA, EMA), any software decision must be validated. If an AI tool influences a drug development document (e.g. target selection report), it should be validated like any analytical instrument. This includes Installation Qualification (IQ), Operational Qualification (OQ), and PC Protocol (actual performance). Google has not (yet) provided GxP validation docs, so companies will need to create their own validation protocols for using Gemini in regulated work.
- **Data Integrity and Security**: Ensure that using Google's services complies with data governance. If patient genomic data or proprietary target sequences are input to the agent, is that data stored or logged? Google Cloud generally has strong security, but policies must be checked. Also, audit: Gemini's usage logs should be maintained.

CodePhusion's analysis of LIMS lock-in ([24] codephusion.com) suggests migratory difficulties; here, we check whether migrating away from Google later is feasible (see vendor lock-in below).

- **Vendor Qualification**: Google (and DeepMind) should themselves meet vendor qualification standards in pharma (e.g. ISO 13485 for software, secure cloud platforms etc.). Companies like to see quality certifications. As of now, large pharma has existing cloud partnerships (e.g. Google Cloud is FDA-21 CFR approved for some workloads). Verifying Google's compliance status for AI/antigravity products will be important.

4. **Business Considerations**

- **Cost and Licensing**: Science Skills is currently available openly, but commercial use of some components (AlphaGenome, deeper Gemini features) may entail fees. Pharma CIOs must understand current and future costs. E.g., will Google charge for Antigravity compute time, or an Alphagenome enterprise API? Is AlphaFold DB usage free (yes, currently)? Long-term subscription models (like Anthropic's Claude) could apply.
- **Vendor Strategy (Moat)**: Evaluate the strategic implications of adopting Google's stack. If a company builds R&D workflows deeply into Gemini/Antigravity, it may become dependent on Google. Conversely, not using Google might mean missing out on productivity gains. The vendor moat analysis (next section) discusses this balance. In evaluation, teams should conduct a cost-benefit analysis, including potential lock-in versus speed advantages.
- **Human-Machine Teaming**: Define how human scientists will interact with the system. For example, will bench scientists ask questions in natural language via a Gemini interface, or will computational chemists plug in code commands? Training staff on a new paradigm is non-trivial.

5. **Ethical and Legal**

- **Intellectual Property (IP)**: Clarify ownership of model outputs. If the AI generates a novel hypothesis or even code (ERA), who owns it? Google's terms or customer terms likely apply. In drug discovery, IP on methods or on potential molecules is core.
- **Bias and Fairness**: While not a typical "bias" context, any AI can have embedded biases in its training data. For example, if AlphaMissense or AlphaGenome were trained mostly on European-centric genomics, their predictions might be less accurate for other populations. Pharma should be mindful of such limits and plan to validate on their target population data.
- **Transparency and Auditing**: Document the use-case thoroughly, especially where AI results inform decisions. Prepare for internal/external audit by documenting how Science Skills was used, what data was queried, and how results were interpreted.

Table 1 below summarizes a possible evaluation framework by criteria and how Google's Science Skills might measure up versus traditional tools.

Criterion	Traditional Informatics Approaches	Google Science Skills (Gemini Agentic)
Data Access	Manual aggregation from siloed sources (LIMS, spreadsheets, etc.) ([25] codephusion.com)	Integrated agentic access to 30+ open scientific DBs (UniProt, AlphaFold DB, etc.) (blog.google) (www.antigravity.google)
Workflow Speed	Slow, linear pipelines often taking hours–days for analysis	Rapid, multi-step workflows executed by AI in minutes, as Google reports (blog.google) (blog.google)
Scalability	Limited by on-prem compute; horizontal scaling is hard	Cloud-based (Google TPUs); can run parallel agent tasks at scale (subject to quotas)
Ease of Use	Requires specialized software expertise (bioinformatics scripts)	Natural-language query and high-level prompts, lowering barrier for non-experts
Reproducibility & Audit	Difficult (manual steps, proprietary formats); audit trail lost on export ([26] codephusion.com)	Potential to record all agent steps; depends on logging implementation (new territory)
Regulatory Compliance	Known frameworks (validated LIMS, GxP software)	New; unknown if Google provides documented 21 CFR11 compliance; requires internal validation

Criterion	Traditional Informatics Approaches	Google Science Skills (Gemini Agentic)
Vendor Lock-In Risk	High for proprietary LIMS/ELN (schema lock-in) ([25] codephusion.com); migrations expensive have been noted ([24] codephusion.com)	Potentially high; reliance on Google's ecosystem (Cloud, Antigravity) gives convenience but may create a new "walled garden".
Cost Model	Upfront license; maintenance contracts; cost to migrate.	Currently free/developer release; future costs uncertain (Cloud compute, API licenses)
Algorithmic Innovation	Often limited to improvement of existing tools	Leverages cutting-edge AI research (foundation models, co-scientist, ERA) (research.google)
Examples of Use	Sequence alignment tools (BLAST); manual literature search	AI-driven literature synthesis (Co-Scientist); variant effect prediction (AlphaMissense, AlphaGenome)
Integration Flexibility	Custom integration needed (ETL); many disparate tools	Built-in multi-tool integration via agent orchestration

Table 1: Evaluation criteria applied to traditional pharma informatics versus Google's Science Skills agentic approach. Refs: Google statements on Science Skills (blog.google) (blog.google); CodePhusion on lock-in ([25] codephusion.com) ([24] codephusion.com).

As Table 1 illustrates, Science Skills promises advantages in integration and speed but introduces novel considerations (like new lock-in risks and regulatory unknowns). Rigorous piloting—potentially paralleling its use with existing workflows—will be key. For example, a pharmaceutical computational biology team might run analyses on historical data using both the old pipeline and the new agent, to compare results and validate consistency.

In concrete terms, **evaluators should design experiments** such as: *"Task: Identify lead candidates for protein X; compare time-to-insight and accuracy of hits using (a) manual pipeline vs (b) agentic pipeline."* Metrics could include total time, number of high-confidence candidates found, and how many of those are supported by known data.

By systematically scoring these and similar tasks, companies can determine how Science Skills fits into their vetting process for new tools. Industry experts stress that such evaluation must be multidisciplinary: involving scientists, informatics staff, and regulatory/legal teams. This is in line with calls in the immunization AI space for **"robust evaluation metrics and benchmarks"** to gauge AI system effectiveness (www.imperial.ac.uk). While [55] focuses on antimicrobial resistance, the principle applies broadly: reproducible challenge datasets and blind tests should be part of the due diligence for any AI claiming to aid drug R&D.

Vendor-Moat Implications and Industry Impact

Google's entry into science informatics will inevitably affect the competitive landscape of drug discovery and bioinformatics tools. We analyze the **vendor and platform dynamics** – often described as "moats" – that could be strengthened or disrupted by Google's moves.

The Traditional Informatics Vendor Landscape

Pharma informatics has long been characterized by a mix of specialized and integrated solutions. Established vendors include electronic lab notebook (ELN) and laboratory information management systems (LIMS) providers (e.g. Veeva, LabWare), bioinformatics suites (BaseSpace, QIAGEN CLC Workbench), cheminformatics platforms (e.g. Reaxys, Scifinder), and cloud providers doing bespoke genomics (AWS Genomics Services, Illumina BaseSpace). Historically, these vendors build deep moats by locking data and processes into proprietary schemas. The CodePhusion analysis** highlights that once labs configure workflows in a commercial system, the accumulated proprietary data model becomes a trap ([25] codephusion.com). Migrating away requires expensive export and validation, as regulators even acknowledge some fidelity is lost ([24] codephusion.com).

Large tech companies have played roles too. Amazon Web Services, for instance, powers many bioinformatics customer platforms (9 of top 10 pharma run on AWS ([27] aws.amazon.com)) and offers genomics toolkits. Microsoft's Azure has initiatives (like Project Zoe) for healthcare data. But these cloud providers mostly supply infrastructure and generic AI/ML tools. None have tightly integrated domain skills like Google is doing. Meanwhile, niche start-ups have emerged (Schrödinger for physics-based drug design, Recursion for phenotypic AI, Insilico Medicine for generative chemistry). These players cultivate intellectual moats via proprietary algorithms and data (e.g. Insilico's in-house Pharma.AI).

Into this landscape comes Google's offering: not a standalone drug-discovery platform, but a **set of AI agent capabilities that sit on premium compute infrastructure (Google Cloud)**. By bundling access to open scientific data, Google is effectively lowering the barrier to entry for the *analysis* of that data, even if Google does not own the data itself. The moats of data owners are thus circumvented: whether the protein sequences are from UniProt or structures from AlphaFold DB, Google's agent treats them as accessible knowledge. This is akin to how search engines flattened access to the Web. Of course, Google does rely on the *participation* of data owners (UniProt, EBI, etc.) who benefit from greater visibility. But in terms of multi-step workflows, Google now controls the *interface*.

An immediate implication is on the **ELN/LIMS/ELN vendors**: while those systems manage experimental data and compliance, tasks like sequence analysis might increasingly happen through AI agents. If a scientist can ask Gemini "analyze these sequence variants" and get an answer (citing databases), they may bypass manual searches within a LIMS or spreadsheets. That does not directly replace LIMS, but it does erode some of the end-user stickiness. Similarly, specialized bioinformatics software vendors (e.g. Benchling's science R&D tools) might see competition: scientists could use the Gemini app or Antigravity notebook instead of their built-in sequence analysis tools.

However, Google does not (currently) own proprietary biological data, which limits its moat in one sense. For example, clinical trial data, proprietary compound libraries, or internal omics data remain with each pharma. Google's agentic tools must be integrated carefully with these private datasets, and companies will host those in their own secured environments. It's also likely (and wise) that pharma IT will be cautious about sending sensitive data to any external service. So, Google's beachhead is mostly in querying public and licensed content.

Nevertheless, the combination of cloud infrastructure and overlaid specialized agents is powerful. It resembles what Microsoft did with Azure and GitHub Copilot in software development: provide the core platform and let open models enhance workflows. Pharma may similarly use Google's platform for lab automation, literature mining, etc. Academia and publishers will also benefit; Google is even partnering with conferences for AI-assisted peer review (blog.google). The "trusted tester" community (PhD to Nobel laureates) mentioned by Google (blog.google) suggests feedback will refine the tools, embedding Google deeper into the scientific ecosystem.

One must consider how other Big Tech will respond. AWS announced general life-science cloud solutions and even generative AI for healthcare, but no direct equivalent of "Science Skills" yet. Microsoft partnered with Pfizer on generative chemistry and is advancing BioGPTs, but again, not an immediate competitor here. NVIDIA offers BioNeMo models and hardware acceleration, but still relies on other front-ends for users. Thus, **Google's offering is relatively unique in scope**. Its competitive advantage (moat) lies in its integration of physics-based (AlphaFold), genomics-based (AlphaGenome), and language-based AI under one brand.

However, this could draw scrutiny: regulators and competitors may worry that Google is building an all-encompassing "platform of platforms". The FCC has already looked at big tech's role in data with NAC operators. In biotech, there may be antitrust concerns if Google becomes essential. On the other hand, by working with over 100 institutions (Stanford, Imperial, Crick) (blog.google), Google shows collaboration rather than monopolization. Yet it is effectively establishing itself as a *gatekeeper* of advanced AI science tools. This is the vendor-moat we examine: can Google capture so much of the scientific workflow that competitors struggle to detach users?

One measure of Google's internal confidence: the Google blog touts that Gemini models (3.5 etc.) can solve tasks in hours that previously took days or weeks for humans (blog.google). Whether this holds across life sciences is to be seen, but if true, scientists may opt to use Google simply for efficiency. The risk is vendor lock-in: if customized models and skills are built only on Google's stack, switching them out is costly (like leaving a proprietary LIMS). Our earlier table flagged this: traditional systems lock data via schemas ([25] codephusion.com); here Google could lock workflows via an "agentic ecosystem".

It is worth comparing Google's open-science narrative with a business reality: DeepMind's commercial arm, Isomorphic Labs, reserves leading-edge capabilities for paying partners. As **Time** reported, AlphaFold 3's model is free for researchers, but *"Isomorphic Labs will have exclusive access to AlphaFold 3 for commercial use"* ([6] time.com). This suggests that while Science Skills (as announced) taps open resources, an implicit strategy is to keep pharma companies in proprietary channels (like deals with Isomorphic or cloud billing) for future exclusives. In effect, the open agentic tools are a pipeline feeder to paywalled services deeper in Google/Isomorphic. Pharma informatics teams should recognize this hierarchy when considering their long-term strategy.

Another consideration is data standards and interoperability. As the CodePhusion analysis emphasizes, **open-source LIMS** approaches can mitigate lock-in ([25] codephusion.com). If Google's skills use widely adopted IDs and exchange formats, data portability might improve. For example, outputs citing UniProt IDs can feed into internal models. But if Google extends into proprietary knowledge (e.g. proprietary fine-tuned models not exportable), then moving away becomes harder. It's too early to tell where Science Skills will be on this spectrum, but any evaluation should watch for:

- Are skill outputs easily exported (CSV, JSON) or only obtainable via Google interfaces?
- Does Google offer transition paths (e.g. containerized Gemini agents) if a company wants in-house versions?
- Will open APIs be accompanied by strict user agreements?

Impact on Existing Vendors and Workflows

To illustrate how Google's actions might disrupt existing moats, consider some examples:

- **Genomics and Sequence Analysis Tools**: Before, a lab might use local BLAST servers (NCBI) or paid services for searches, and standalone apps for sequence alignment. Now, an agent can do "Find homologs of TP53 and summarize differences" via Science Skills. Vendors who sold sequence-annotation pipelines (e.g. DNASTAR, Geneious) may find their user base diminished if Gemini becomes the interface of choice.

- **Structural Biology Software**: Legacy tools like PyMOL or Chimera are used for visualization, and suites (Schrödinger, MOE) for modeling. If Gemini can fetch and view structures via conversational prompts, this could complement or replace some manual steps. However, for deep tasks like molecular dynamics, specialized software will remain relevant. Interestingly, Google's announcement signals cooperation: they mention the open AlphaFold Server can be used by anyone, and Isomorphic Labs (a Google spinout) is already collaborating with pharma on structure-guided design (blog.google) ([6] time.com).

- **Chemical Databases and Tools**: Tools like SciFinder (CAS) and Reaxys hold chemistry data, often behind costly subscriptions. Science Skills linking to PubChem/ChEMBL might not replace SciFinder (which also has integrated reaction chemistry), but it does provide some overlap for finding known compounds or bioactivity. The mention of PubChem and ChEMBL (www.antigravity.google) suggests agents can query chemical structures and assay results quickly. Smaller biotech might leverage that for initial screens, potentially reducing SciFinder usage or complementing it.

- **Literature and Knowledge Discovery**: Google is arguably already a de facto literature tool via Scholar. The novelty here is agentic synthesis. Co-Scientist and others show that Google's models can read and propose insights from literature. Traditional "knowledge discovery" platforms (like Causaly, Incite) could face competition if Gemini can answer complex questions directly from literature. On the other hand, academic publishers may partner (as hinted by AI-paper review tools) because AI demands scholarly content.

- **Cloud and Infrastructure**: Many pharma are already on major clouds (AWS, Azure, GCP). Google's entry specifically for science could tip the balance for those on the fence. For example, if companies start storing and analyzing data via Antigravity, they incur cloud consumption. This increases Google's overall cloud revenue and deepens dependency. A "cloud-provider moat" is already recognized in enterprise IT; Google now bolsters its life-science moat. Competing clouds will likely respond (AWS expanding health AI) but Google has a head start on agentic interfaces.

- **Integration Partners**: A subtle impact is on partners in the analytics ecosystem. Many companies integrate tools via APIs. If Google agents can do end-to-end tasks, the business case for middleware might change. Conversely, integration could be easier: if an ELN exports all new data to Google Cloud, AI agents could automatically analyze it. Some LIMS providers might partner with Google (for example, Bolt et al developing connectors) to stay relevant.

In summary, Google's Science Skills lowers the technical barriers to advanced informatics, which can erode some existing lock-ins. Traditional vendors may see reduced demand for purely manual tools (especially in bioinformatics and literature mining). However, the competitive moats of vendors with proprietary data (e.g. private clinical assays, patented compounds) remain intact unless Google strikes specific deals for them. It is a nuanced shift: Google is **expanding the frontier** of what's possible, and users will naturally gravitate to the most efficient tools. The bar for others will be to integrate similar agentic capabilities or to offer unique data/services that Google does not cover.

For pharma procurement and strategy, the lesson is to anticipate this shift. More cross-vendor collaboration or openness may be required. Some industry observers even note a trend towards consolidation (where tech giants acquire smaller AI-focused bio companies). Time will tell if Google acquires more pharma tech arms (beyond DeepMind/Isomorphic), or if incumbents form alliances with AWS/Microsoft to respond. Regardless, any company evaluating informatics moats should revisit assumptions: the "data is king" principle may still hold, but the "access is king" principle (i.e. who provides organized access) is now critical.

Future Directions and Implications

Looking ahead, the implications of Gemini, Science Skills, and associated technologies can be framed in terms of potential developments and challenges:

- **Expanding Data and Models**: We expect Google (and others in AI) to continue expanding these integrated systems. Possible future steps include adding more specialized skills (e.g. for cryo-EM maps, clinical trial registries, patient data anonymization). The underlying models (Gemini) will evolve (3.5 Flash to higher versions) with greater reasoning ability. On the DeepMind side, new Alpha tools (beyond AlphaFold3) may be released, covering protein-protein interactions, or even cell-level models. Each such advance could be folded into future Science Skills updates. For pharma, the roadmap might involve skills for target identification, virtual screening, or translational research (linking biomarkers to outcomes).

- **Regulatory Evolution**: As AI becomes a deployed part of drug R&D, regulators will adapt. We already see discussions of AI/machine learning in regulatory guidance (FDA's pledge to modernize evaluation of software-driven devices). In Europe, the upcoming AI Act may impact how "high-risk AI" is certified (if say AI suggests diagnostic or therapeutic decisions). Pharma companies will need alignment: if Gemini outputs inform critical path decisions (e.g. IND-enabling studies), regulators may ask for evidence of validation. The field of *explainable AI (XAI)* is likely to grow, with specialized tools for validating biomedical AI claims.

- **Collaborative Research and Open Science**: Google's partnerships with institutions (Stanford for fibrosis, Imperial for AMR, Crick for cancer) indicate a collaborative approach (blog.google). It's plausible that open challenges (akin to CASP for protein structure) will emerge for agentic science, inviting the community to stress-test these tools. If Google Labs opens some Gemini for Science experiments publicly, there could be crowdsourced insights. Projects like the WHO's AMR initiative (co-authored by Google) suggest a future in which science is increasingly data-sharing and multi-stakeholder. Pharma might join consortia where models trained on combined pharma/academic data yield mutual benefits.

- **Ethical and Societal Implications**: Beyond efficacy, the societal role of AI in science matters. If agents can propose drug targets, who bears responsibility if things go wrong? The interplay with IP also raises questions: if Gemini suggests a molecule, does the idea itself become a patentable invention? If many companies use the same base models and data, will innovation crowd around new delivery mechanisms (like novel drug formulations)? Additionally, equity issues arise: will low-resource labs benefit (via open access) or will only well-funded companies exploit these tools (especially if commercial access is restricted)? Early hints: Google's emphasis on "responsible development" and building community "trusted testers" (blog.google) suggests awareness of these issues, but concrete frameworks are yet to crystallize.

- **Competition from Other AI Paradigms**: Google's approach is one of agentic AI built on foundation models. Parallel lines include BioGPT (specialized LLMs trained on biomedical text), and AI-driven robotics for wet labs (company Opentrons with automated protocols + AI planning). Also, quantum computing promises long-term impacts on molecular simulation. Pharma must keep an open mind: Gemini may be powerful, but not the sole route. That said, Google's integration of multiple modalities (text, sequence, structure) potentially offers an advantage over single-focus LLMs.

- **Commercial Strategy and Market Dynamics**: Possible futures include (1) deepened Google-pharma partnerships: we saw AstraZeneca, Bezos, others form AI partnerships pre-2025; maybe more will work with Google Cloud/DeepMind. (2) M&A: Google/Alphabet may acquire promising startups (e.g. Insitro, Recursion), or competitors might buy out AI companies to bolster their own offerings. (3) Pricing models: Google might tier services (free for bench scientists, paid enterprise version with higher throughput). Companies like AstraZeneca may decide to build internal AI teams versus outsourcing to Google – a strategic decision influenced by how these tools perform.
- **Cultural Change in Pharma**: On the human side, AI like Gemini could change the role of researchers. Some may fear automation of tasks they currently do. Others will become "AI supervisors" – drinking in quick AI analyses then applying domain judgment. Training and workflows will shift. For instance, R&D teams may need PhDs fluent in prompting and verifying agent outputs. Similarly, traditional bench labs may incorporate in silico AI experiments as preliminary steps.
- **Open Data and Standards**: A positive possibility is that Google's emphasis on open knowledge push others to cooperate more. Pharma often keeps data private, but initiatives like Accelerated Access and OpenData portals may multiply if the AI demand grows. Interoperability standards (HL7 FHIR for health data, GA4GH for genomics) could become more critical for feeding AI pipelines. This could ultimately benefit science: better data sharing and FAIR (Findable, Accessible, Interoperable, Reusable) principles would feed better AI outcomes.

In conclusion, the **Google for Science launch at I/O 2026** marks a potential inflection point in how AI interfaces with drug discovery. It demonstrates impressive technical prowess and ambitions, but also confronts us with strategic challenges. The future will likely see an ever-deepening role for AI agents – but also a need for vigilant evaluation, flexible strategies, and ethical oversight. As with any transformative technology, outcomes will depend on responsible use and wisdom in integration. The citations and analysis in this report are intended to help science and informatics leaders navigate this new terrain.

Conclusion

Google's unveiling of Science Skills, together with the Gemini agent platform, brings a powerful new capability to the life sciences arena. By embedding major biological databases (UniProt, AlphaFold DB, etc.) and AI models (AlphaGenome, Co-Scientist) into an accessible, agentic interface, Google aims to accelerate scientific workflows dramatically (blog.google) (blog.google). Early results – such as minute-scale analysis leading to novel genetic insights and successful drug repurposing discoveries – are promising indicators that agentic AI can act as a *multiplier* for human researchers (blog.google) (deepmind.google).

For pharmaceutical R&D, the **technical implications** are significant. Tools like AlphaFold DB put structural predictions in every researcher's hands (alphafold.ebi.ac.uk). AlphaGenome opens up interpretation of the vast non-coding genome ([7] spectrum.ieee.org). Agents combining these resources could cut years off target discovery cycles. Our analysis has provided a framework to **evaluate** these tools in context – asking how they perform against benchmarks, how they integrate with compliant workflows, and what costs/risks they entail. We have also highlighted that rigorous testing and regulatory due diligence are essential before relying on AI outputs in critical drug development decisions.

At the same time, Google's moves carry **strategic consequences**. By bundling data and AI, Google strengthens a new competitive moat around its platform. Conventional vendor moats based on proprietary data may weaken as agents access open data through Google's interface. Pharma companies and informatics vendors must consider how to adapt: either by leveraging Google's capabilities in-house or by ensuring their own platforms offer similar integrative AI functions. Partnerships and standards (e.g. open APIs, cloud agreements) will be key to maintaining flexibility.

Looking forward, we see a landscape where multi-agent AI tools become part of the standard toolbox in pharma. Companies that thoughtfully adopt and govern these tools could see large productivity gains. Our recommendations include: (1) **pilot and benchmark** agentic AI features in non-critical projects before broad adoption; (2) ensure **transparency** by having the AI cite its data sources; (3) negotiate terms that allow data portability and define commercialization rights; (4) retrain workflows to incorporate human-AI collaboration; and (5) engage in community efforts to establish best practices for AI use in medicine.

Finally, the broader scientific community will judge tools like Gemini and Science Skills by their outcomes. The hope is that, ultimately, these innovations will enable faster development of drugs, personalized therapies, and medical insights that save lives. But history tells us that any powerful tool requires wisdom in use. Ongoing collaboration between tech developers, scientists, and regulators will be vital. The work of evaluating Google's science skills – as we have begun here – must continue dynamically as these tools evolve.

References: Inline citations have been provided throughout to substantiate each claim, drawing on official Google announcements (blog.google) (blog.google), peer-reviewed publications (deepmind.google) ([12] spectrum.ieee.org), industry analyses ([2] pharmaphorum.com) ([1] www.insideprecisionmedicine.com), and expert reports ([25] codephusion.com) (www.imperial.ac.uk). Each reference is linked to the original source for verification.

External Sources

[1] https://www.insideprecisionmedicine.com/topics/precision-medicine/deepminds-isomorphic-inks-3b-worth-of-deals-with-lilly-and-novartis-in-one-day/#:~:With%...

[2] https://pharmaphorum.com/digital/biopharmas-ai-rally-readiness-not-hype-2025#:~:The%2...

[3] https://blog.scriptome.ai/2025/04/20/ai-agents-in-biotech-and-pharma/#:~:The%2...

[4] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:Enter...

[5] https://time.com/6975934/google-deepmind-alphafold-3-ai/#:~:Resea...

[6] https://time.com/6975934/google-deepmind-alphafold-3-ai/#:~:Alpha...

[7] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:Accor...

[8] https://www.alphagenomecommunity.com/t/alphagenome-nature-publication-weights-for-academic-use/752#:~:Alpha...

[9] https://cloud.google.com/blog/ja/products/ai-machine-learning/running-alphafold-on-vertexai#:~:%E6%9...

[10] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:Alpha...

[11] https://github.com/google-deepmind/alphagenome#:~:The%2...

[12] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:Even%...

[13] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:%E2%8...

[14] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:Alpha...

[15] https://www.alphagenomecommunity.com/t/alphagenome-nature-publication-weights-for-academic-use/752#:~:A%20c...

[16] https://pharmaphorum.com/digital/biopharmas-ai-rally-readiness-not-hype-2025#:~:risks...

[17] https://pharmaphorum.com/digital/biopharmas-ai-rally-readiness-not-hype-2025#:~:Small...

[18] https://pharmaphorum.com/digital/biopharmas-ai-rally-readiness-not-hype-2025#:~:Looki...

[19] https://www.tomsguide.com/news/live/google-io-2026-live-news-updates#:~:,grow...

[20] https://www.insideprecisionmedicine.com/topics/precision-medicine/deepminds-isomorphic-inks-3b-worth-of-deals-with-lilly-and-novartis-in-one-day/#:~:Isomo...

[21] https://www.insideprecisionmedicine.com/topics/precision-medicine/deepminds-isomorphic-inks-3b-worth-of-deals-with-lilly-and-novartis-in-one-day/#:~:%E2%8...

[22] https://time.com/6975934/google-deepmind-alphafold-3-ai/#:~:,%28S...

[23] https://spectrum.ieee.org/alphagenome-ai-gene-regulation#:~:%E2%8...

[24] https://codephusion.com/blog/lab-software-data-lock-in#:~:The%2...

[25] https://codephusion.com/blog/lab-software-data-lock-in#:~:Data%...

[26] https://codephusion.com/blog/lab-software-data-lock-in#:~:Then%...

[27] https://aws.amazon.com/health/biopharma/#:~:AWS%2...

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.