Humigent

POINT OF VIEW · LIFE SCIENCES COMMERCIAL · EVALUATING AI

7 Key Criteria for Evaluating AI Solutions in Life Sciences

How to separate durable advantage from another stalled pilot

A guide for commercial, medical affairs, and market access leaders

01 Domain Expertise	02 Interoperability & Flexibility	03 Usability & Adoption	04 Performance	05 Accuracy & Explainability	06 Innovation	07 Trust, Compliance & Governance

humigent.ai

info@humigent.ai

EXECUTIVE SUMMARY

The issue isn't AI adoption. It's AI placement.

Knowing where AI earns its place, and where your people stay in the lead.

Life sciences is under the same pressure as every other industry to digitize, modernize, and do more with less. But the pressure here is sharper. The industry is moving through a steep loss-of-exclusivity cycle, with a wave of its top-selling products, including some of the single biggest sellers on the market, facing generic and biosimilar competition in the second half of this decade. At the same time, Medicare price negotiation has moved from debate to reality: the first negotiated prices took effect in 2026, a second round of medicines (including the leading GLP-1s) follows in 2027, and from 2028 the program reaches physician-administered drugs for the first time. Launch windows are shorter, commercial teams are leaner, and every brand is expected to find growth in data it already owns but can't fully use.

AI has been positioned as the answer. In principle, it should help commercial, medical, and market access teams move faster, unifying fragmented data, surfacing the insight buried in dashboards and field notes, and freeing experts to spend their time on judgment instead of assembly. In practice, most of the industry is stuck. Pilots multiply, but few make it into production. The reasons are familiar to anyone who has lived through one: the tool didn't understand pharma, the output couldn't be trusted in a regulated workflow, the data never came together, or the project quietly stalled when the consultants left.

The problem usually isn't ambition. It's evaluation. With dozens of vendors claiming to do everything from harmonizing claims data to writing medical insights, it has become genuinely hard to separate what will scale from what will demo well once and disappear. And the deeper discipline isn't adopting AI, it's placing it: deciding what to automate, what to augment, and what must stay human.

This guide lays out the seven criteria that should govern any serious assessment of an AI platform for life sciences. Our criteria are rooted in the realities of regulated commercial work and designed to evaluate both the product and the partner behind it. We built Humigent around these principles, so we won't pretend to be neutral. But the criteria stand on their own, and they apply universally.

A note on scope: our examples are drawn from the commercial, medical affairs, and market access functions. The principles, however, hold across the enterprise.

AT A GLANCE

The Seven Criteria

Use these seven tests to assess both the platform and the partner behind it.

01 **Domain Expertise**
Purpose-built for life sciences, not repurposed general AI.

If it doesn't speak pharma, it won't survive contact with your data.

02 **Interoperability & Flexibility**
Fits data, systems, and deployment model already in place.

Most AI projects die in the gap between your data sources, not in the model.

03 **Usability & Adoption**
Built into daily work, with the support to make change stick.

A platform only creates value when your team actually uses it.

04 **Performance**
Consistent, reliable output that holds up under real commercial workloads.

Promises don't run commercial operations. Reproducible output does.

05 **Accuracy & Explainability**
Grounded in your data, traceable to source, and reviewable.

It isn't enough to be right. You have to be able to show why.

06 **Innovation**
A platform that compounds in value instead of restarting from zero.

The best platforms compound. The rest restart from zero every time.

07 **Trust, Compliance & Governance**
Transparent, secure, and with a human in the loop.

Trust is the gate every other criterion has to pass through.

CRITERION 01 · DOMAIN EXPERTISE

01 Purpose-Built for Life Sciences

If it doesn't speak pharma, it won't survive contact with your data.

AI can't replace the judgment of a brand lead, an analytics director, or a medical science liaison. But a platform built for the work can sharpen that judgment, driving consistency, accelerating delivery, and raising the quality of every output. The teams that win the next few years won't be the ones that bought the most AI. They'll be the ones whose AI actually understood their business.

If you've tested general-purpose tools, you've likely hit the same wall: they don't speak pharma. They don't know that a "line of therapy" is a defined construct, that an HCP affiliation has to survive a vendor refresh, or that a claims dataset and an EHR feed describe the same patient in different ways. They miss the vocabulary, the structure, and the stakes. Consider rare disease: only about 500 of roughly 7,000 known rare diseases have a disease-specific ICD-10 code, so finding patients means pattern recognition across messy, incomplete data. This is the nuance a general tool misses and a domain-built one is designed for.

WHAT TO LOOK FOR

- **Domain-native models, not generic models in disguise.** Ask what the platform is actually built on. A general foundation model wrapped in pharma terminology is not the same as a model trained on commercial, clinical, medical, and access data. The strongest platforms are built on domain-native models that understand life sciences data structures, indications, biomarkers, payer logic, and field terminology natively, rather than through prompt engineering alone. Organizations should ask vendors to demonstrate how this expertise is embedded into the platform architecture.
- **Built by people who've done the work.** Choose a partner whose team pairs deep pharma and biotech operators with deep AI engineers. Some of the hardest problems in this space (identity resolution, cohort logic, and promotional compliance) are domain problems first and technical problems second.
- **Designed to elevate expertise, not override it.** The best platforms surface the relevant signal, frame the decision, and hand your experts work they can take across the finish line. They make your best people faster, not redundant.

IN PRACTICE

A platform built for the life sciences should have solutions mapped to recurring life sciences workflows, e.g. assessing data quality, unifying fragmented HCP profiles, and building defensible patient cohorts without hand-written SQL. Each should start from how experienced teams already do the work, then accelerate it with AI, rather than forcing teams to reinvent their processes. That is how Humigent's solutions are designed.

VENDOR CHECKLIST · WHAT TO LOOK FOR

- ☐ Domain-native models, not generic models in disguise
- ☐ Designed to elevate expertise, not override it
- ☐ Built by people who've done the work

02 Fits the World You Already Operate In

Most AI projects die in the gap between your data sources, not in the model.

Commercial life sciences data is fragmented by default. Claims, Rx, EHR, CRM, specialty distribution, patient support programs, syndicated research, and a dozen internal trackers all describe the same market in different shapes, and they rarely agree. Most AI initiatives fail because these sources remain disconnected, long before anyone evaluates the model.

The right platform meets your environment as it is. It connects to the cloud, data, and BI infrastructure you've already invested in, it works across structured and unstructured sources, and it doesn't force a rip-and-replace to get value. Just as important, it should flex to your business model, not the vendor's.

WHAT TO LOOK FOR

- **Native connectivity to your stack.** Confirm the platform integrates with the data platforms and BI tools you already run (Snowflake, Databricks, Redshift, Azure, Tableau, Power BI, Qlik, and the rest) and that it handles both structured feeds and unstructured documents, presentations, and field notes.
- **Harmonization as a first-class capability.** Connecting to data isn't the same as making sense of it. Ask how the platform resolves identity, reconciles conflicting sources, and unifies records into something a team can actually trust. This is where most projects quietly fail.
- **Architectural and deployment flexibility.** Ask whether you can start with a single point solution and expand to full enterprise deployment, and whether the platform can run other foundation models alongside the domain model where it makes sense. Flexibility is what keeps a platform useful as your needs change.
- **No lock-in, by design.** Flexibility isn't only technical, it's commercial. Ask whether your data, configurations, and outputs stay portable, whether the platform can be deployed within your own environment, and whether you can scale up or step back without being trapped by proprietary formats. A platform should earn its place each year on the strength of the value it delivers, not on switching costs.

IN PRACTICE

Humigent connects and harmonizes existing disparate data sources, resolves and unifies records, then makes the result query-ready, so a brand, field, or analytics team is working from one harmonized view instead of reconciling spreadsheets.

VENDOR CHECKLIST · WHAT TO LOOK FOR

- ☐ Native connectivity to your stack
- ☐ Architectural and deployment flexibility
- ☐ Harmonization as a first-class capability
- ☐ No lock-in, by design

03 Value From Day One

A platform only creates value when your team actually uses it.

A platform only delivers value when people actually use it. Commercial and medical teams don't have time for steep learning curves or six-month implementations, and they won't adopt a tool that lives outside their daily work. The right solution feels familiar from the start, fits the way teams already operate, and comes with the kind of support that makes change stick inside a risk-aware organization.

This is the part most technology vendors underestimate. Adoption in life sciences is as much a change-management problem as a software problem. The technology is necessary; the human side (onboarding, enablement, building trust with skeptical experts) is what determines whether anything scales.

WHAT TO LOOK FOR

- **Conversational by default.** The best platforms let users ask questions in plain language ("Where are my early adopters with positive sentiment?" or "Why did sales in this account decline last month?") and get explainable answers, instead of forcing every question through an analyst queue.
- **A partner, not just a product.** Look for a provider that pairs the technology with people who understand pharma workflows and can guide deployment. Education, discovery, and a realistic roadmap are required for any successful rollout.
- **A pragmatic path to value.** Strong partners earn trust with quick wins, then build toward the harder, higher-value use cases. Humigent's execution model is deliberately staged: education and alignment first, then discovery and roadmap, then phased deployment, then internalization. The point is momentum, not a big-bang launch that overwhelms the internal team.

IN PRACTICE

Humigent positions itself as an "AI Sherpa" for life sciences, guiding teams from buzzwords to measurable impact. That means meeting teams where they are, starting with workflows that produce visible value quickly, and expanding only as confidence grows.

> "A platform only creates value when your team actually uses it.
>
> CRITERION 03 · USABILITY & ADOPTION

VENDOR CHECKLIST · WHAT TO LOOK FOR

☐ Conversational by default	YES · NO
☐ A partner, not just a product	YES · NO
☐ A pragmatic path to value	YES · NO

CRITERION 04 · PERFORMANCE

04 Consistent, Reliable Output Under Real Workloads

Promises don't run commercial operations. Reproducible output does.

Every AI vendor promises time savings. But promises don't run commercial operations, performance does. The real test isn't a polished demo on clean data. It's whether the platform holds up across many brands, territories, data types, and use cases, week after week, and still produces output your team can rely on. Performance should be tested against real-world workflows such as launch monitoring, congress intelligence synthesis, HCP segmentation refreshes, market research analysis, and field force optimization.

WHAT TO EVALUATE

- **Speed with substance.** Results should come back fast without sacrificing depth or relevance. A quick answer that misses the nuance of a payer policy or a line-of-therapy rule isn't a time savings, it's rework waiting to happen.
- **Quality at scale.** Test the platform under realistic load: multiple therapeutic areas, messy real-world data, conflicting sources, subnational granularity. Does output stay consistent, or degrade the moment it leaves the demo environment?
- **Reproducibility.** In a regulated industry, the same question asked twice should return the same answer. Reproducible output isn't a nice-to-have, it's the difference between an interesting tool and one you can put in front of compliance.

IN PRACTICE

Humigent workflows are built for reproducibility by design.

EXAMPLE OUTCOMES OBSERVED ACROSS EARLY DEPLOYMENTS

Weeks → hours	80%	3–5 hrs	70%+
Market research synthesis	Reduction in HCP segmentation refresh cycles	Analyst productivity gained per week	Faster data quality investigation

VENDOR CHECKLIST · WHAT TO EVALUATE

- [ ] Speed with substance
- [ ] Quality at scale
- [ ] Reproducibility

05 Outputs You Can Trust and Trace

It isn't enough to be right. You have to be able to show why.

In life sciences, precision isn't optional. A wrong number in a forecast, a misattributed insight, or an unsupported claim in a medical context doesn't just create rework, it creates risk. That's why accuracy, and specifically verifiable accuracy, must sit at the center of any evaluation. It isn't enough for an AI to be right. You must be able to show why it's right.

This is where the difference between architectures matters most. A standard approach retrieves some data and lets the model improvise an answer fast, but with no validation step and no reasoning trail. When that answer is wrong, no one can see why, and in a regulated workflow that's disqualifying.

WHAT TO LOOK FOR

- **Source-grounded output.** Insights, summaries, and recommendations should trace back to the underlying data and logic, not emerge from a black box. Ask the vendor to show you the lineage of an answer, not just the answer.
- **The "why" behind every output.** A score sent to the field with no reason behind it gets ignored. Surfacing the recommendation is the machine's job; the rationale, and the field judgment around it, belong to a person. The best platforms build that explanation layer between the model and the user, but most don't. It is exactly the work that needs both machine and human.
- **Reasoning you can inspect.** The strongest architectures deliberately separate the steps: retrieving the relevant data, reasoning over it with validation, and only then generating language. Data becomes a knowledge graph that encodes domain meaning; deterministic logic validates each step before any language is produced; the model explains the result rather than inventing it. Every path should be inspectable by compliance. This is how Humigent's architecture is designed.
- **Glass box, not black box.** This is a core design principle, not a feature bolted on later. Every action is auditable and every output is traceable to its sources and the logic that produced it. If a vendor can't explain how an answer was reached, your reviewers can't defend it. If an AI system recommends prioritizing an HCP, identifying a patient cohort, or adjusting payer engagement strategy, teams should be able to trace every recommendation back to the underlying evidence and business logic.

IN PRACTICE

"Glass-box auditability" is one of Humigent's foundational differentiators. Reproducible outputs, source attribution, and an inspectable reasoning path mean an output isn't just an answer, it's an answer your team and your compliance reviewers can stand behind.

VENDOR CHECKLIST · WHAT TO LOOK FOR

- ☐ Source-grounded output
- ☐ Reasoning you can inspect
- ☐ The "why" behind every output
- ☐ Glass box, not black box

CRITERION 06 · INNOVATION

06 A Platform That Compounds

The best platforms compound. The rest restart from zero every time.

A strong platform shouldn't rest once deployed. It keeps pace with the industry, shipping improvements, expanding capabilities, and getting smarter with each use case. But in life sciences there's a deeper test of innovation than release cadence: does the platform compound, or does every new use case start from zero?

Most engagements restart each time. New project, new data wrangling, new model, new learning curve, and none of the institutional knowledge from the last effort carries forward. A platform built right does the opposite. Each deployment makes the next one faster and smarter, because the system inherits everything it has already learned about your data, your market, and your standards.

WHAT TO LOOK FOR

- **A compounding architecture.** Ask whether use case five is meaningfully smarter than use case one. A modular, multi-agent design, where new capabilities inherit what the platform already knows, turns AI from a series of one-off queries into a growing institutional asset.
- **A steady, ambitious roadmap.** A regular release cadence signals a vendor invested in the platform's evolution. Look beyond incremental features to a roadmap focused on reshaping how commercial and medical work gets done.
- **Trusted, well-governed data partnerships.** Where the platform draws on external or third-party data, confirm those sources are authoritative and that their use complies with applicable privacy and contractual requirements.

IN PRACTICE

A modular, multi-agent design lets new capabilities be added to a platform that already understands your data, your market, and your standards, so capability accrues instead of resetting with each project. Humigent is built this way.

> "The best platforms compound. The rest restart from zero every time.
>
> CRITERION 06 · INNOVATION

VENDOR CHECKLIST · WHAT TO LOOK FOR

☐ A compounding architecture	YES · NO
☐ A steady, ambitious roadmap	YES · NO
☐ Trusted, well-governed data partnerships	YES · NO

07 Transparent, Secure, and Human in the Loop

Trust is the gate every other criterion has to pass through.

You're not just evaluating features; you're choosing a partner whose technology will touch sensitive data and sit inside regulated workflows. In an environment shaped by HIPAA, medical-legal-regulatory (MLR) review, the PhRMA Code, and FDA promotional standards, and with regulators now turning their attention to AI itself, signaled by the FDA's January 2025 draft guidance on the use of AI to support regulatory decision-making, trust isn't a soft attribute. It's the gate every other criterion has to pass through.

A trustworthy partner is clear about how the platform works, how it's trained, how it's secured, and how it's governed, and designs compliance into the architecture rather than treating it as a deliverable to be added later.

WHAT TO EVALUATE

- **Compliance built into the design.** Ask whether compliance is structural or cosmetic. The reproducibility and source-traceability described earlier aren't only accuracy features, they're what make an AI output defensible under review. Capabilities like real-time screening against medical and commercial guardrails should be part of how the system works, not a manual step afterward.
- **Enterprise-grade security and deployment control.** A credible vendor should meet recognized security standards and give you control over where your data lives. Humigent's solutions are designed to be deployed within your firewall, so sensitive commercial, clinical, and access data never has to leave your environment.
- **A human in the loop, by design.** The goal isn't to remove people from the decision, it's to give them better inputs. Insist on a platform that escalates ambiguous cases to humans and keeps the key decision rights with your team. High-consequence activities, anything involving a patient, an HCP, or a compliance-sensitive decision, should stay subject to human approval even when the underlying work is automated.

> "Trust is the gate every other criterion has to pass through."
>
> CRITERION 07 · TRUST, COMPLIANCE & GOVERNANCE

VENDOR CHECKLIST · WHAT TO EVALUATE

☐ Compliance built into the design	YES · NO
☐ Enterprise-grade security and deployment control	YES · NO
☐ A human in the loop, by design	YES · NO

A TEST FOR TRUSTWORTHY AUTOMATION

Before any activity is fully automated, it should clear three bars:

1
A repeated cycle
It runs again and again.

2
A checkable output
Accuracy you can verify, not take on faith.

3
An audit trail
A glass box, with every output traceable to its sources and logic.

Even when all three are met, high-consequence steps stay subject to human approval. The goal isn't simply efficiency, it's trustworthy automation.



THE AUDIT TRAIL MUST SURVIVE EACH HANDOFF

The failure is rarely inside any one step. It's at the seams, the handoffs between machine and human.

IN PRACTICE

Humigent's approach is deliberately "glass box" rather than "black box," with hybrid AI-plus-human oversight across capture and synthesis. Compliance, auditability, and human judgment are treated as design principles: the foundation the rest of the platform is built on.

DESIGN PRINCIPLE

Design for the handoff, not just the step.

That's the difference between a demo and a system compliance will approve.

★ FEATURED EVALUATION

How a Commercial Organization Evaluated AI Beyond the Demo

The strongest evaluations test the platform against criteria like these, not against a demo.

The Challenge

A specialty pharmaceutical brand was preparing for launch. Marketing, analytics, and field teams were working across fragmented datasets and struggling to generate and act on timely insights.

The Evaluation

The organization evaluated multiple AI vendors against seven criteria:

- 01 Domain expertise
- 02 Interoperability
- 03 Usability
- 04 Performance
- 05 Accuracy
- 06 Innovation
- 07 Trust & governance

The assessment included:

- ✓ Synthetic data testing
- ✓ Commercial use cases
- ✓ Security reviews
- ✓ Compliance validation

Why Humigent Stood Out

Humigent distinguished itself through:

- ✓ Native life sciences understanding
- ✓ Reproducible and explainable outputs
- ✓ Seamless integration into existing data infrastructure
- ✓ Ability to support both structured and unstructured commercial data

The Outcome

Teams reduced manual data reconciliation effort, accelerated insight generation, and established a scalable foundation for future AI initiatives.

THE DECIDING FACTOR

Their decision came down to the same seven criteria in this guide: domain expertise, interoperability, usability, performance, accuracy, innovation, and trust.

01	02	03	04	05	06	07
Domain expertise	Interoperability	Usability	Performance	Accuracy	Innovation	Trust

THE "SO WHAT"

From Stalled Pilots to Durable Advantage

The winners won't automate the most work. They'll automate the right work.

Choosing an AI platform isn't only a technology decision, it's a strategic one. The platform you choose will shape how your teams operate, how fast they deliver, how confidently they can stand behind their work, and whether your investment turns into a lasting capability or another project that quietly stalls.

The seven criteria in this guide are demanding on purpose. The industry has spent the last few years learning, expensively, that an impressive demo is not the same as a platform that scales inside a regulated, data-fragmented, compliance-bound business. The most effective organizations won't be the ones that automate the most work, they'll be the ones that automate the right work, with repetitive tasks handled by agents that can be monitored, audited, and improved, and human expertise focused on judgment, relationships, and strategic decisions.

Humigent was built for exactly this environment: purpose-built for life sciences, accurate and explainable by design, and architected so that compliance and human oversight are foundational rather than afterthoughts. From day one, it's designed to work the way your teams already work, and to get smarter every time you use it.

BRING US ONE WORKFLOW

AI creates durable advantage when it is applied to the right work.

Bring us one commercialization workflow, from launch planning and field force optimization to patient finding or market access analytics. We'll help you determine where to deploy AI, and what must stay human. The result is a focused first use case, clear governance, and a path from experimentation to measurable business impact.

Humigent · info@humigent.ai · humigent.ai

FAQS

What should you look for in an AI solution for life sciences?

Look for domain expertise, interoperability, usability, reliable performance, accuracy and explainability, continuous innovation, and strong trust, compliance and governance capabilities.

Why does domain expertise matter for life sciences AI?

Life sciences data includes specialized terminology, patient, HCP, and payer relationships, claims, EHR data, payer information and regulated workflows. AI designed specifically for the domain is better positioned to interpret these structures and support relevant workflows.