

WHY MEDTECH AND PHARMACEUTICAL COMPANIES MUST OWN THEIR AI INFRASTRUCTURE

The case for vertical integration over marketplace dependency



CONTENTS

Abstract	03
Introduction	03
The Evolution and Importance of Intelligent Building Management	04
What Vertical Integration Means and What It Does Not	05
The Compounding Advantage of Proprietary AI Infrastructure	06
The Regulatory Argument	07
The Talent Argument	08
The Cyient Perspective: Building the Infrastructure	08
Cyient's Engineering Intelligence Platform	09
Agentic AI as the New Operating Model	10
Conclusion	13
About the Authors	13
About Cyient	14

Abstract

MedTech and pharmaceutical companies acquiring AI capabilities through marketplace vendors are solving an access problem while creating a structural dependency. This white paper argues that vertical integration of AI infrastructure, owning the data architecture, model development environment, validation and compliance layer, and deployment stack, produces durable competitive advantages that marketplace dependency cannot. The evidence is now concrete enough to act on. The window for making this investment at reasonable cost is narrowing.

Introduction

Most MedTech and pharmaceutical companies acquiring AI capabilities today are making a decision that feels pragmatic in the short term and will prove expensive in the long term. They are buying access to foundation models, analytics platforms, and regulatory intelligence tools from marketplace vendors whose business model depends on those companies never building the capability themselves. The result is an industry generating AI activity without building AI infrastructure and confusing the two.

This white paper makes a specific argument: vertical integration of AI infrastructure owning the data architecture, the model development environment, the validation and compliance layer, and the deployment stack produces durable competitive advantages that marketplace dependency structurally cannot. The evidence for this claim is now concrete enough to act on, and the window for making the investment at a reasonable cost is narrowing.



The Evolution and Importance of Intelligent Building Management

When a MedTech or pharmaceutical company deploys a third-party AI platform - a regulatory intelligence SaaS tool, a commercial drug discovery engine, a vendor-supplied imaging algorithm, it is solving an access problem. It is getting AI capability into production faster than it could build one internally. That is a legitimate short-term gain.



What Vertical Integration Means and What It Does Not

Vertical integration of AI infrastructure does not mean building a foundation model from scratch. It does not mean hiring 200 data scientists and competing with technology companies on their own terms. Those framings set an impossible bar that conveniently justifies the status quo.

Practically, it means owning four layers of the stack:



Data infrastructure

The pipelines, governance frameworks, and storage architecture that make a company's proprietary data systematically available for AI applications. Clinical trial records, real-world evidence from connected devices, manufacturing process historian data, regulatory submission histories, post-market surveillance findings, and compound screening databases are all potential AI fuel that most companies currently manage as operational records rather than strategic assets. The companies that structure, clean, and govern this data create a foundation that compounds in value over time.



Model development and fine-tuning environment

The infrastructure to take foundation models (whether open-source or licensed) and adapt them to proprietary data and domain-specific tasks. A regulatory intelligence model fine-tuned on a company's own submission history, reviewer feedback, and deficiency letter patterns will consistently outperform a generalized commercial tool on that company's submissions. A pharmacovigilance model trained on a company's specific product portfolio, patient population, and adverse event history will detect signals that a generic model misses or flags as noise.



Validation and compliance layer

Where life sciences AI diverges most sharply from AI in other industries. In a regulated environment, an AI system supporting regulatory decision-making, clinical trial monitoring, or medical device functionality must be validated under 21 CFR Part 11, ISO 13485, IEC 62304, and the FDA's evolving AI guidance. Validation is not a one-time event, it requires continuous monitoring, drift detection, revalidation protocols, and audit-ready documentation. A company that owns this layer controls its AI compliance posture; one that relies on a vendor is accepting someone else's interpretation of regulatory requirements for processes it is ultimately responsible for.



Deployment and integration architecture

The APIs, connectors, and workflow integrations that embed AI outputs into operational systems where decisions are made. This means connecting AI inference into TrackWise, Veeva, SAP, Jama, Windchill, and clinical data management systems. Marketplace tools typically offer standard integrations; custom integration into a company's specific configuration, data model, and approval workflows is almost always either unavailable or prohibitively expensive through a vendor relationship.

The Compounding Advantage of Proprietary AI Infrastructure

The core economic argument for vertical integration is that proprietary AI infrastructure compounds in value in a way that marketplace access cannot.

Pharmacovigilance

A company with vertically integrated AI infrastructure begins with a fine-tuned model on its own adverse event history, product portfolio, and patient population. As post-market surveillance data accumulates, the model improves, detecting more signals earlier, generating fewer false positives, adapting to new product lines as they launch. The regulatory intelligence embedded in the system improves with every submission, every deficiency letter, every reviewer interaction. After three years, the gap between this system and any commercially available pharmacovigilance tool is substantial. After seven years, the gap is structural.

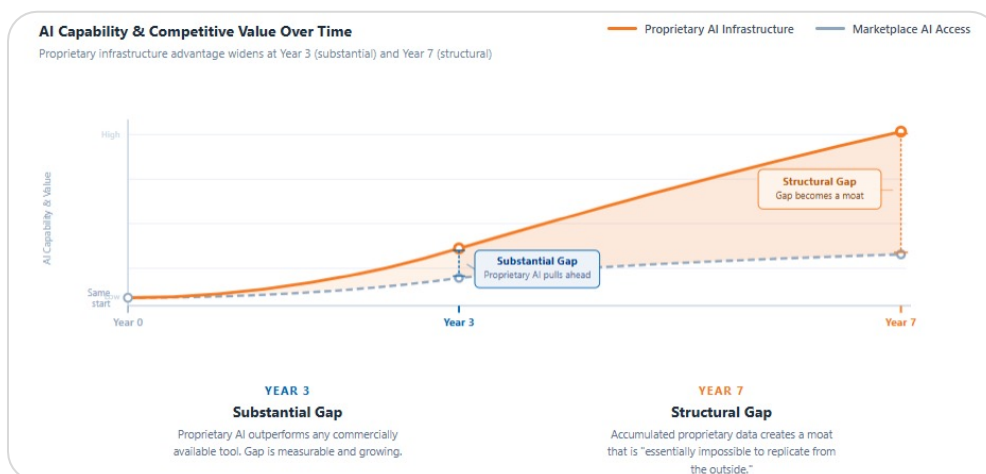
Drug Discovery

A company whose AI discovery platform is trained on its own compound screening data, failed trial records, and internal biomarker research is working from a biological map of its therapeutic areas that no external vendor has access to. Each new experiment, each clinical result, and each failure enriches that map. The accumulated proprietary biology becomes a moat that is essentially impossible to replicate from the outside.

Manufacturing

Processing historical data from a company's own facilities - equipment signatures, environmental variables, batch genealogy, deviation patterns, trains predictive models specific to that company's processes in a way that generic models cannot be. A batch deviation prediction model trained on 10 years of a company's own manufacturing data will outperform a commercial tool trained on industry-wide data for that company's specific processes.

McKinsey's analysis of AI leaders across industries consistently finds that companies capturing disproportionate AI value share two characteristics: they have invested in proprietary data infrastructure, and they have built internal capability to develop and operate AI systems rather than solely consuming them. In life sciences, where proprietary data is among the most valuable in any industry, this finding has particularly sharp implications.



After seven years, the gap between a vertically integrated pharmacovigilance system and any commercially available tool is structural.

The Regulatory Argument

There is a second argument for vertical integration that deserves separate treatment—because it is less frequently made and arguably more important for life sciences specifically: regulatory control.

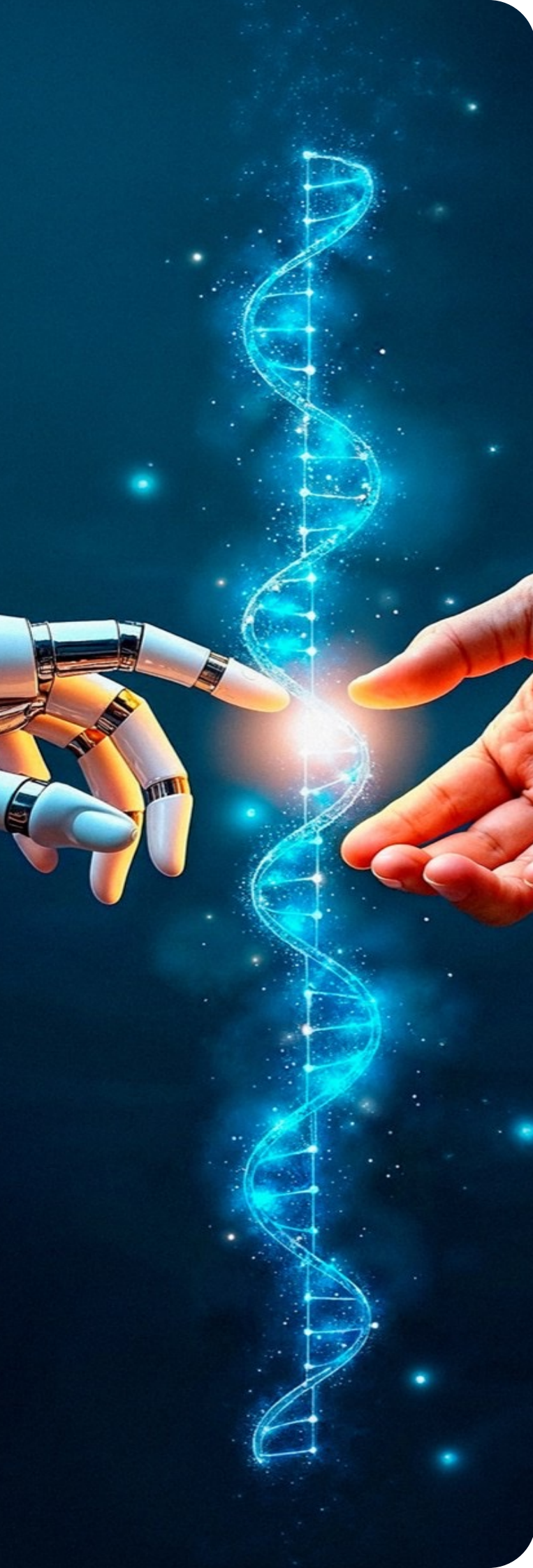
The FDA's January 2025 draft guidance on AI supporting regulatory decision-making, and the joint FDA-EMA guiding principles published in January 2026, establish a clear direction of travel. AI systems used in regulated contexts - drug development, medical device functionality, manufacturing quality, post-market surveillance will face increasing levels of scrutiny, documentation, and transparency requirements. The agency expects companies to demonstrate understanding of their AI systems: how they were trained, on what data, with what validation, and under what change management protocols.

A company that built its AI infrastructure internally can answer those questions with precision. It has full visibility into training data provenance, model architecture decisions, validation methodologies, and deployment configurations. When a regulatory query arrives about an AI-assisted submission or a device algorithm, the response is drawn from internal documentation the company owns and controls.

A company that deployed a marketplace solution faces a fundamentally different situation. The answers to those questions sit with the vendor. "What training data was used?" The vendor decides what to disclose. "How was the model validated for this specific use case?" The vendor's standard validation may or may not map to the company's specific regulatory context. "What happens when the model is updated?" The company may not even know.

This is not a hypothetical risk. As regulatory scrutiny of AI in life sciences intensifies and the trajectory from both FDA and EMA is unambiguous - companies with opaque, vendor-dependent AI infrastructure will find themselves unable to answer basic questions about systems they are operationally dependent on. The regulatory relationship risk that creates is material, and it is one that vertically integrated competitors will not share.





The Talent Argument - Often Overlooked, Consistently Decisive

Companies that build proprietary AI infrastructure develop a class of talent that marketplace consumers cannot attract or retain: engineers and scientists who work at the intersection of deep domain knowledge and applied AI development.

A data scientist who fine-tunes regulatory intelligence models on real submission histories, understands the difference between substantial equivalence arguments under 510(k) and the clinical evidence requirements of a PMA, and has built pharmacovigilance signal detection systems reviewed by the FDA exists in a small talent pool and commands significant market value precisely because the combination of domain and technical depth is rare.

Companies that consume marketplace AI solutions have limited need for this profile. They need people who can configure and manage vendor tools, a narrower and more easily substitutable capability. The organizational consequence is that their AI capability is permanently bound by what their vendors choose to build, while their vertically integrated competitors continuously extend the frontier of what is possible with their specific data and domain context.

The talent flywheel is underappreciated. Proprietary infrastructure attracts domain-AI talent. That talent builds better infrastructure and better models. Better models generate more distinctive outputs and more proprietary insights. The accumulated capability makes the organization a more attractive destination for the next generation of domain-AI talent. Marketplace consumers are outside this flywheel entirely.

The Cyient Perspective: Building the Infrastructure That Makes the Advantage Real

The argument for vertical integration is straightforward. The execution challenge is real, and it is where most organizations get stuck, not because they disagree with the logic, but because they lack a credible path from current state to the integrated architecture they need.

This is where Cyient's position is distinctive. Two decades of embedded engineering partnership with the world's leading MedTech OEMs have produced a detailed operational understanding of the data environments, regulatory workflows, and technical architectures that proprietary AI infrastructure must integrate with. Cyient has worked inside TrackWise and Veeva implementations, built DHF and technical file structures to ISO 13485, written and validated software under IEC 62304, and supported 510(k) and PMA submissions for imaging, diagnostics, and connected care devices across multiple regulatory jurisdictions.

The Cyient AI platform architecture for MedTech and Pharma is built on this foundation. It is vendor-agnostic by design - built on open standards, deployable on client infrastructure or private cloud, and integrated directly into the operational systems where regulatory, quality, and engineering decisions are made. The QARA AI platform brings AI-native workflows to regulatory submissions, post-market surveillance, and quality management without requiring clients to route sensitive data through external systems. The validation framework is built to 21 CFR Part 11 and ISO 13485 from the ground up, not retrofitted.

Critically, Cyient works alongside clients to build the internal capability that makes their AI infrastructure genuinely theirs. The goal is not a Cyient-dependent system. It is a client organization with data infrastructure, model development environment, validation processes, and deployment architecture to extend and operate its AI capability autonomously. That is what durable competitive advantage looks like in this space and it is what marketplace dependency structurally cannot deliver.

The goal is not a Cyient-dependent system. It is a client organization with the capability to extend and operate its AI infrastructure autonomously.



Building the Data Infrastructure That Makes AI Real - Cyient's Engineering Intelligence Platform

The first and most consequential obstacle most MedTech and pharmaceutical companies face is not a shortage of AI ambition; it is a shortage of AI-ready data. Clinical trial records, adverse event histories, device performance telemetry, regulatory submission archives, manufacturing process historians, and compound screening databases exist in every organization of scale. What does not exist, in most cases, is the infrastructure to make that data systematically available, governed, and machine-readable at the level of precision that regulated AI applications demand.

Cyient's Engineering Intelligence Platform (EIP) is purpose-built to close that gap. The platform's data engineering and ingestion layer - Layer 7 of the EIP architecture, ingests multi-modal data from the operational systems that already exist in a client's environment: QMS platforms such as TrackWise, regulatory and clinical data repositories in Veeva, manufacturing historians in SAP, and unstructured institutional knowledge locked in documents, shift notes, and engineering manuals. Critically, EIP does not require data migration or a multi-year transformation program before value is realized. The platform is designed to layer over what the client already has, deliver measurable insight within days of deployment, and enrich its own knowledge base continuously as new data flows in.

What makes EIP genuinely differentiated for life sciences is what sits above the data layer: a proprietary domain ontology and knowledge graph architecture built specifically for engineering-intensive, regulated industries. In most AI deployments, data is treated as flat text, fed into models as documents or rows without structured context. The EIP ontology encodes the semantic relationships that drive real decision-making in life sciences:

- A compound maps to its screening history, which connects to trial outcome data, which links to regulatory submission precedents, which traces to post-market surveillance signals.
- An adverse event is not just a record, it is a node in a graph connecting patient population characteristics, device configuration, clinical setting, and corrective action history in a form that agents can traverse and reason over.

For a pharmacovigilance team, this means signal detection that contextualizes anomalies against the full evidence graph of the product portfolio. For a regulatory affairs team, it means submission intelligence that draws on the company's own history of reviewer interactions, deficiency letter patterns, and predicate device arguments, not a generic model trained on public filings. For a medical device manufacturer managing post-market obligations under the EU MDR, it means a living compliance map connecting device variants, clinical evidence, real-world performance data, and PMCF requirements in a single, auditable knowledge structure.



Agentic AI as the New Operating Model - From Workflow Automation to Patient Outcomes

The business case for AI infrastructure in life sciences ultimately rests not on operational efficiency alone, but on something more material: the ability to improve patient care and strengthen patient safety at scale. Agentic AI - autonomous, multi-step AI systems that reason, plan, and act across enterprise workflows is the mechanism through which proprietary data infrastructure converts into those outcomes.

The EIP deploys a coordinated architecture of specialized agents, each trained on domain-specific data and governed by human-in-the-loop validation gates, that collectively transform how MedTech and pharmaceutical organizations operate across their most consequential processes:

A pharmacovigilance agent trained on a company's own adverse event history, product portfolio, and patient population characteristics does not merely automate signal detection, it detects safety signals earlier, generates fewer false positives, and routes actionable alerts to the right clinical and regulatory decision-makers faster than any manual process can.

A quality and regulatory compliance agent spanning the full product lifecycle - from deviation detection in manufacturing through CAPA management, regulatory impact assessment, and audit-ready evidence generation does not merely reduce the cost of compliance; it closes the gap between a quality event and a corrective action in days rather than weeks, directly reducing the window of patient exposure to unresolved product risks.

A medical device post-market surveillance agent that continuously ingests real-world performance data from connected devices, maps it against the clinical evidence base, and flags emerging safety signals against regulatory thresholds is not performing a reporting function, it is functioning as a continuous patient safety system.

What the EIP agentic model delivers is not a set of disconnected automations—it is a new operating architecture in which agents collaborate across functions, each enriching a shared knowledge graph that becomes more precise and more organizationally specific with every engagement. A risk identification agent flags a potential device performance trend; an action tracking agent ensures the corrective response is owned, time-bound, and closed; a reporting agent generates the regulatory notification with full traceability to source data; a compliance validation agent confirms the evidence package meets the applicable standard before submission.

This is the difference between AI as a productivity tool and AI as an operating model: the former reduces manual effort within existing processes, while the latter redesigns the processes themselves around autonomous, continuously improving intelligence.

For MedTech and pharmaceutical leaders, the imperative is to recognize that this transition is not a future state, it is available now, it is being implemented by the organizations that will define the competitive and clinical standard of the next decade, and it requires the foundational investment in owned data infrastructure and domain-specific AI capability that marketplace solutions are structurally incapable of providing.

What makes Cyient's position genuinely differentiated at this moment is the breadth of what Cyient brings to this problem. Most technology partners arrive at AI infrastructure from one direction—software, consulting, or systems integration. Cyient arrives from all of them simultaneously and adds something most cannot: end-to-end hardware engineering and data center services capability spanning physical infrastructure design, network architecture, server provisioning, and the facility-level engineering that underpins a private AI compute environment.

This matters because the infrastructure question is not only a software question. A MedTech or pharmaceutical company building a vertically integrated AI architecture needs to make decisions about where computing lives - on-premises, private cloud, hybrid and those decisions have direct implications for regulatory data sovereignty, latency in real-time manufacturing and pharmacovigilance applications, and the total cost of ownership over a 10-year horizon. A partner that can design and deliver the physical infrastructure layer, the network and security architecture, and the software and AI stack in an integrated engagement removes a coordination problem that most organizations underestimate until they are in the middle of it.

The Cyient Thought Board

Question	Key Points
Why is marketplace AI insufficient for life sciences?	- Designed for generalization, not competitive specificity - Exposes sensitive proprietary data to external systems - Creates operational dependency on vendor continuity
What does vertical integration actually require?	- Owned data infrastructure and governance - Model fine-tuning environment on proprietary data - Internal validation and compliance layer - Custom deployment into operational systems
Why does proprietary AI compound in value?	- Models improve with every new data point from company operations - Pharmacovigilance, drug discovery, and manufacturing advantages widen over time - McKinsey: AI leaders consistently own their own data infrastructure
What is the regulatory case for ownership?	- FDA Jan 2025 draft guidance; FDA-EMA joint principles Jan 2026 increase AI scrutiny - Internal ownership provides full traceability and auditability - Vendor-dependent companies cannot answer basic regulatory questions about their own systems
How does the talent flywheel work?	- Proprietary infrastructure attracts domain-AI talent - That talent builds better models and deeper capability - Marketplace consumers are structurally outside this flywheel
What makes the timing urgent?	- Companies investing in data infrastructure in the mid-2010s have a 10-year compounding advantage - The same dynamic is unfolding now for AI infrastructure broadly - The window to begin at comparable cost and position is measurable in months, not years

Conclusion

The case for vertical integration is not that marketplace solutions are without value. Some are excellent, and a well-designed proprietary infrastructure will selectively incorporate external capabilities where they are genuinely superior. The case is that the architecture, where data lives, who controls the models, how validation is governed, how AI outputs are integrated into operational decisions must be owned. Everything built on top of that architecture can be sourced flexibly. The architecture itself cannot be borrowed.

Companies that invested in proprietary data infrastructure in the mid-2010s, structuring their clinical trial data, instrumenting their manufacturing processes, building longitudinal real-world evidence databases from their connected devices are now sitting on training assets that took a decade to accumulate. Competitors starting that same investment today face a 10-year compounding disadvantage on the data layer alone, regardless of how sophisticated their modeling capability becomes.

The organizations making the foundational investments in 2026 in data governance, model fine-tuning environments, validated deployment architecture, and domain-AI talent, are building a compounding advantage that will be structurally difficult to replicate in 2030. The organizations waiting for the technology to mature, or for vendor solutions to become sufficiently sophisticated, are deferring a cost that will be significantly higher when they finally pay it - both in dollars and in competitive position.

The organizations that choose to build now, with the right partner, will reach patients faster, protect them more reliably, and make better decisions at every stage of the product lifecycle. The window to start at comparable cost and competitive position is measurable in months, not years.

About the Authors



Harjott Atrii

Chief Business Officer –
Strategic Initiatives, Cyient

Harjott Atrii is a global business leader with over 27 years of experience driving digital transformation, AI, cloud, and strategic growth initiatives across global markets. At Cyient, he leads strategic technology partnerships, GTM initiatives, and special growth programs.



Umesh Kuppuraj

Solutions Practice Head - Healthcare
& Life Sciences, Cyient

With over two decades of experience spanning engineering, quality, and enterprise solutions, Umesh brings a rare blend of hands-on product lifecycle expertise and executive-level business acumen. He drives strategic solutioning, innovation-led growth, and largescale transformation for global MedTech and Life Sciences clients, partnering closely with CXOs to address mission-critical challenges across regulatory compliance, operational scalability, and AI-driven intelligent automation initiatives.



About Cyient

Cyient (Estd: 1991, NSE: CYIENT) delivers intelligent engineering solutions across products, plants, and networks for over 300 global customers, including 30% of the top 100 global innovators. As a company, Cyient is committed to designing a culturally inclusive, socially responsible, and environmentally sustainable tomorrow together with our stakeholders.

For more information, please visit
www.cyient.com



Contact Us

North America Headquarters

Cyient, Inc.
99 East River Drive
5th Floor
East Hartford, CT 06108
USA
T: +1 860 528 5430
F: +1 860 528 5873

Europe, Middle East, and Africa Headquarters

Cyient Europe Limited
Apex, Forbury Road,
Reading
RG1 1AX
UK
T: +44 118 3043720

Asia Pacific Headquarters

Cyient Limited
L14 3 Parramatta Square,
153 Macquarie St,
Parramatta NSW 2150,
Australia
T: +61 2 8887 8600

Global Headquarters

Cyient Limited
Plot No. 11
Software Units Layout
Infocity, Madhapur
Hyderabad - 500081
India
T: +91 40 6764 1000
F: +91 40 2311 0352

Follow us on:  