

Building an AI-Ready Healthcare Organization with Data Intelligence to Drive Clinical Success

Preparing healthcare data for AI demands accuracy, diverse representation and governance that delivers actionable insights

"Gene," the Children's Hospital of Philadelphia (CHOP)'s branded data catalog, is transforming how professional clinicians, researchers and operational teams access and explore the data and analytic assets essential to decision-making across the organization. Drawing from one catalog, users can search, discover and explore the data and analytic assets essential to decision-making.



By working from one centralized catalog, users can access reliable, well-governed information and quickly understand key details."

BETH SENAY | Director of Data Trust for the Office | Children's Hospital of Philadelphia

"By working from one centralized catalog, users can access reliable, well-governed information and quickly understand key details such as data lineage, quality and governance implications — building confidence in the data, improving efficiency and accelerating insights across CHOP's pediatric enterprise," said Beth Senay, Director of Data Trust for the Office at the Philadelphia-based health system.

This cultural shift reflects what distinguishes healthcare organizations that successfully deploy AI at scale from others. Almost 80% of organizations have implemented AI in at least one business function,[1] yet only 11% have achieved generative AI at scale.[2]

The challenge isn't the volume of data, but rather making that data AI-ready. As Gartner notes, "AI-ready data must represent the use case, including patterns, errors, outliers and unexpected emergence."[3] This shift from conventional data quality standards requires healthcare organizations to rethink their data management strategies while maintaining governance and compliance.

Understanding AI-ready healthcare data

AI-ready data differs fundamentally from traditional healthcare analytics data. Where conventional business intelligence seeks clean, standardized datasets with outliers removed, AI models require a comprehensive representation of real-world complexity.

Healthcare AI models must account for clinical variation, not sanitized averages. Abdul Tariq, Associate Vice President of Data Science at CHOP, explained: "We want to make sure that these models are accurate and operationally relevant."

This real-world representation includes patterns and variations across diverse patient populations, errors and anomalies that may signal critical clinical conditions, outliers representing rare but important cases and unexpected emergences of new clinical patterns.

As Garland Lynn, Alation Regional Vice President, East Region, observed, "Healthcare's data tells a fragmented story — shaped by years of being captured on paper and only recently digitally for clinical, operational and research needs. Yet when the data becomes trusted and unified, it unlocks one of the most profound opportunities for AI to transform patient care, decision-making and value."

Accuracy with real-world representation

Creating AI-ready healthcare data begins with ensuring that datasets accurately reflect the full spectrum of clinical reality. CHOP's three-step process examines workflow integration, conducts statistical evaluations to identify bias across subgroups and implements mitigation strategies through either statistical methods or strategic deployment decisions.

Tariq said that a major hurdle with AI data is demonstrating its real-world utility. "By far, the biggest challenge is figuring out exactly how or why the AI or machine learning solution is going to improve upon the current state," he explained.



When data becomes trusted and unified, it unlocks one of the most profound opportunities for AI to transform patient care, decision-making and value."

GARLAND LYNN

Regional Vice President, East Region | Alation



> **By far, the biggest challenge is figuring out exactly how or why the AI or machine learning solution is going to improve upon the current state."**

ABDUL TARIQ | Associate Vice President of Data Science | CHOP

Understanding how AI will integrate into existing workflows requires data that mirrors those workflows, including their inconsistencies, variations and edge cases. Without this comprehensive representation, organizations risk building models that perform well in testing but fail when confronted with the realities of clinical practice.

Tariq cited physiologic deterioration prediction models as an example: "Even in cases where machine learning models outperform physicians, it's actually hard to deploy these models such that they provide actionable information."

The critical question isn't whether a model can predict risk, but whether clinicians can act on that prediction. "If you don't have an umbrella, even if you know it's going to rain, you're still going to get wet," he said.

This drives CHOP's implementation philosophy: incremental rollout starting with pilot studies, performance validation in smaller cohorts and scaling only after gathering sufficient evidence.

From data assets to data products

Healthcare organizations are shifting from managing data assets to creating purpose-built data products: curated datasets aligned with specific business outcomes and clinical applications. These data products are essential for building trusted AI systems that deliver measurable business value by ensuring analysts and data scientists work from reliable, governed foundations, rather than starting anew each time.

Caroline Peika, Director of Integration and Analytics at Rady Children's Health (formerly Children's Hospital of Orange County), described her team's strategic shift since adopting a new data intelligence platform. Rather than forcing data scientists to create data products from scratch with each project, they're building curated, validated and continuously monitored datasets as a foundation. "We've been building this to enable data scientists or other power users to start from there," she said.

This approach reflects a broader transformation. Lynn noted that more organizations now consider creating data products from their data assets by asking: "What is the business outcome they're trying to drive? What are the business processes required to drive that outcome? And then, what is the data needed to drive those processes?"

Building effective data products requires:

- **Representation of diverse populations:** In healthcare, data products often incorporate protected health information (PHI) to enable AI models that improve patient outcomes across diverse populations. Research excellence requires organizations to leverage comprehensive patient data while safeguarding PHI to prevent AI from perpetuating health disparities. This requires robust governance frameworks that balance innovation with regulatory requirements and ethical obligations.
- **Domain-specific stewardship:** Both CHOP and Rady Children's Health employ data stewards for specific domains such as surgery, clinical care and revenue cycle who understand the nuances of data collection and use within their respective areas. Peika emphasized that cross-functional workgroups must understand each AI model's use case (and that AI is the best tool) before even starting, "so we're not using a hammer for everything, only for nails."
- **Reusable, certified datasets:** By building curated data layers, organizations enable data scientists and power users to access trusted data confidently, accelerating innovation while maintaining quality and compliance.



> **Building [curated data sets enables] data scientists or other power users to start from there."**

CAROLINE PEIKA

Director of Integration and Analytics | Rady Children's Health

> **"The biggest mistake organizations make is failing to define the business outcome they want from AI — and then aligning the right data, trust and precision to achieve it."**
>
> GARLAND LYNN

The evolution of healthcare data governance

Data governance in healthcare has evolved from a compliance requirement to an AI enabler. CHOP's governance approach centers on what Senay called patient-centered custodianship. "We see ourselves as the custodians of data owned by our patients," she said. "Our obligation is to manage the data ethically, securely and in ways that help advance health outcomes and research excellence."

CHOP's Data Quality Center of Excellence operationalizes this through automated and retrospective quality reviews. Role-based access controls further provide users with the appropriate level of data access based on their organizational role, ensuring both data accessibility and security. By granting appropriate access to authorized personnel while maintaining strict compliance boundaries, this approach enables broad participation in data-driven innovation without sacrificing governance or compromising patient privacy.

Balancing compliance and innovation

The tension between maintaining HIPAA, GDPR and other regulatory compliance while enabling AI innovation requires careful navigation. "We select partners that allow us to work within those boundaries of compliance," Peika said.

Rady Children's Health is implementing Alation's data catalog to quickly identify sensitive data, helping teams quickly distinguish PII and PHI and enabling informed decisions about what data a project requires.

The role of data intelligence platforms

Both CHOP and Rady Children's Health emphasize the importance of unified data cataloging. As Senay explained, "Working from one trusted catalog ensures every team is using consistent definitions, understanding data assets and governance around them using certified sources. It builds confidence and speeds up insights."

Data intelligence platforms like Alation provide capabilities essential for AI readiness by providing business context integration that captures definitions and usage patterns; data lineage visibility tracking data flow from source to destination; collaborative curation where teams add annotations and reviews; and AI-powered discovery with natural language querying. "We're giving the breadcrumbs on what SQL [database] it ran and why it ran," Lynn said. "You can provide feedback to say, 'That wasn't exactly what I wanted,' and it's self-learning."

An implementation roadmap to building AI-ready data

Healthcare organizations should start by identifying high-value business problems where AI can deliver measurable impact rather than attempting to catalog all available data.

"The biggest mistake organizations make is failing to define the business outcome they want from AI — and then aligning the right data, trust and precision to achieve it. Without that approach, even the most advanced models can miss their purpose and even cause some problems," Lynn said.

This outcome-focused approach ensures that data management efforts align with specific clinical or operational needs. In parallel, organizations must build foundational governance by establishing data stewardship roles, defining accountability frameworks and implementing quality monitoring systems. These governance structures should be in place before scaling AI initiatives to prevent the proliferation of ungoverned, potentially biased models across the enterprise.

Advanced organizations can move beyond individual AI applications to orchestrate multiple specialized AI agents, each addressing specific aspects of larger clinical

> **"The most important thing is having a process for evaluating all deployments of AI."**
>
> ABDUL TARIQ

or operational challenges. This approach enables more sophisticated automation while maintaining human oversight at critical decision points.

Mature programs should also focus on scaling data product libraries, such as expanding curated dataset repositories with comprehensive documentation that enables enterprise-wide reuse. By building reusable, certified datasets rather than bespoke data pipelines for each project, organizations accelerate AI development while maintaining consistent quality and governance standards across all applications.

"The most important thing is having a process for evaluating all deployments of AI. You need technical expertise to create data products and write code, but you also need domain expertise from clinicians and providers to evaluate whether a particular tool is working or not," Tariq said.

Organizations that align data with specific use cases, qualify it to meet AI requirements and implement governance accounting for AI's unique challenges will accelerate their competitive advantage. Those that delay will find themselves outpaced by providers delivering better clinical outcomes, optimized operations and more equitable care — all enabled by AI-ready data intelligence.

To learn more, visit alation.com.

References

1. Singla, A., Sukharevsky, A., Yee, L., Chui, M., and Hall, B. March 12, 2025. The state of AI: How organizations are rewiring to capture data. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/.
2. Baig, A., Merrill, D., Sinha, M., Mead, D., and Xu, S. May 13, 2024. Moving past gen AI's honeymoon phase: Seven hard truths for CIOs to get from pilot to scale. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/moving-past-gen-ais-honeymoon-phase-seven-hard-truths-for-cios-to-get-from-pilot-to-scale.
3. Sweenor, D. Sept. 24, 2024. Gartner report: How AI-ready data drives AI success. Alation. https://www.alation.com/blog/gartner-report-ai-ready-data-drives-success/.

Alation

About Alation

Alation is the data intelligence company nearly 600 global enterprises — including 40% of the Fortune 100 — rely on to realize value from their data and AI initiatives. Customers such as Cisco, DocuSign, Nasdaq and Pfizer trust Alation's platform for self-service analytics, cloud transformation, data governance, and AI-ready data, fostering data-driven innovation at scale. Our platform brings together data cataloging, governance and AI-powered search, helping care teams, analysts and IT work from a single source of truth. Whether you're driving clinical research, improving patient outcomes or ensuring HIPAA compliance, we help you unlock the full value of your data — safely and efficiently. Headquartered in Redwood City, California, Alation has been recognized five times by Inc. Magazine as one of the Best Workplaces. To learn more, visit www.alation.com.

Produced by HIMSS®