Data Readiness: The Foundation for Scalable AI

HIMSS Market Insights study shows that data readiness is the real key to unlocking the potential of analytics and AI

Connecting healthcare's digital systems — interoperability — has come a long way in the past several decades. "But what we've learned is that interoperability is much more than simply getting data from one place to another. We still have a long way to go," said Julie Smith, Head of Healthcare Global Market Strategy, InterSystems.

Smith's observation is supported by a recent HIMSS Market Insights survey, sponsored by InterSystems, that assessed the state of data integration in U.S. healthcare organizations.[1] The research offered insight into why data exchange and data readiness are not the same, the role of the cloud, and how data readiness supports analytics and AI initiatives.

> **Interoperability is not just about the flow of information; the information has to be usable at the other end."**

JULIE SMITH | Head of Healthcare Global Market Strategy | InterSystems

Data exchange is not the same as data readiness

Across the U.S., most healthcare leaders (86%) agreed that interoperability without data preparation provides little value for analytics or AI (Figure 1). Two-thirds (67%) agreed that most interoperability challenges today are about data readiness, not connectivity.

Seventy percent (70%) of leaders said that data preparation should be an expected outcome of interoperability, not a separate effort. At the same time, leaders were divided between a narrow definition of interoperability — moving data reliably between systems (35%) — and a more comprehensive definition of interoperability — moving data plus cleaning, standardizing, and enriching it for downstream use (36%).

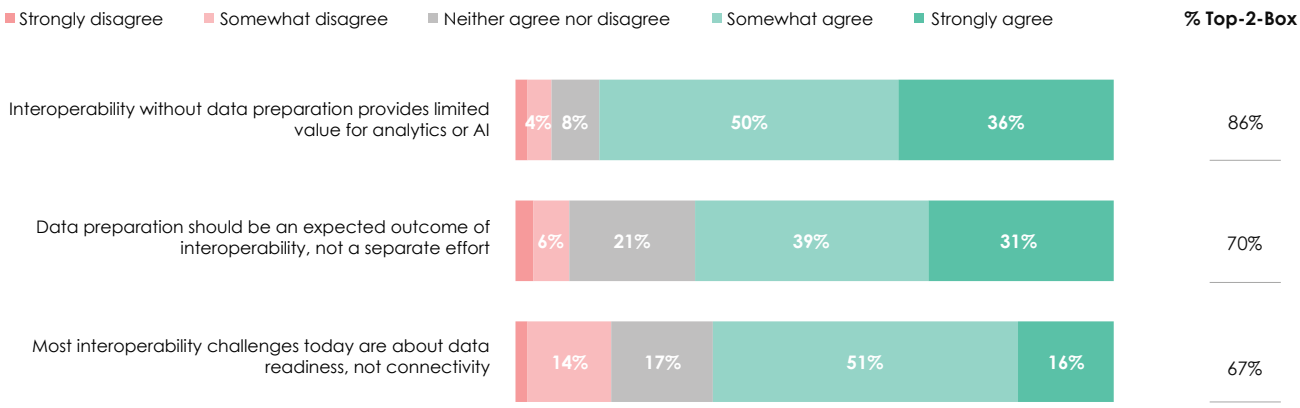
Different understandings of what interoperability should include can complicate organizational discussions about data readiness. "Originally, interoperability was focused on the individual patient in front of the provider," explained Smith. "But now, with all of this data digitized, we also have the opportunity to use it for populations."

Analytics and AI tools are ideal for this expanded use case. However, "you've got to have clean data aggregated in a usable way to make sense of it, draw correct conclusions, and define next step actions," Smith said. "What we know is that if you don't filter and clean up the data before you feed it into these tools, they don't give you the right answers."

This is why it is important to adopt a more complex definition of what interoperability is as organizations move ahead with analytics and AI applications. "Interoperability is not just about the flow of information; the information has to be usable at the other end," she emphasized. Smith offered the example of writing a letter in English and mailing it to France. If the recipient doesn't speak English, the data has technically been exchanged, but not in a useful or meaningful way.

Figure 1. Most leaders agree that interoperability and data preparation and readiness are interconnected objectives, enabling greater value from their analytics and AI initiatives.

Please indicate the extent to which you agree or disagree with each of the statements below.



The cloud helps, but doesn't ensure data readiness

Nearly three-quarters (73%) of healthcare leaders agreed that data preparation challenges persist after moving data to the cloud (Figure 2). Two-thirds (64%) agreed that cloud platforms enable better automation for data readiness. Just over one-third (34%) thought moving to the cloud improves data quality by default.

"It is a misconception that moving to the cloud improves data quality by default, because moving to the cloud doesn't change anything about the data except the location," pointed out Jeff Fried, Director of Platform Strategy and Innovation, InterSystems. Data quality only improves when organizations take intentional steps, including activating specific tools and processes, to improve data readiness.

He added that moving data to the cloud can create new opportunities for organizations to address data readiness and to use data such as:

- providing frictionless access to more tools, including automation for data readiness
- enabling more uses for data, like analytics and AI applications
- including access to experts who can help organizations work with their data

Strategies for achieving data readiness

Leaders reported that they currently use an average of two different approaches to preparing data for analytics and/or AI. Fifty-seven percent (57%) rely on custom scripts or in-house tools; 50% use third-party ETL (Extract, Transform, Load) or data integration tools; and 39% rely on manual data cleansing by IT teams. Only 9% rely on largely automated processes with minimal manual intervention.

"Unfortunately, the reality is that most of these approaches don't scale," said Smith. "The volumes of data we are dealing with are massive. If you are using manual scripts to prepare your data, there is no way you are going to be able to get to the scale you need when we are talking about analytics and AI."

As leaders considered interoperability solutions that include data readiness capabilities, they identified specific functions the solutions should be able to handle (Figure 3). Nearly three-quarters (73%) specified data normalization and mapping; 70% identified data quality checks and remediation; 68% said terminology and code standardization; 55% said data transport; 52% said analytics or AI-ready data delivery; and 46% said data enrichment or contextualization. On average, leaders believe interoperability solutions should handle four of these functions.

Figure 2. Nearly three-quarters of healthcare leaders agree that data preparation challenges persist after moving data to the cloud; however, many also agree that cloud platforms enable better automation for data readiness.

Please indicate your level of agreement with each of the following statements.

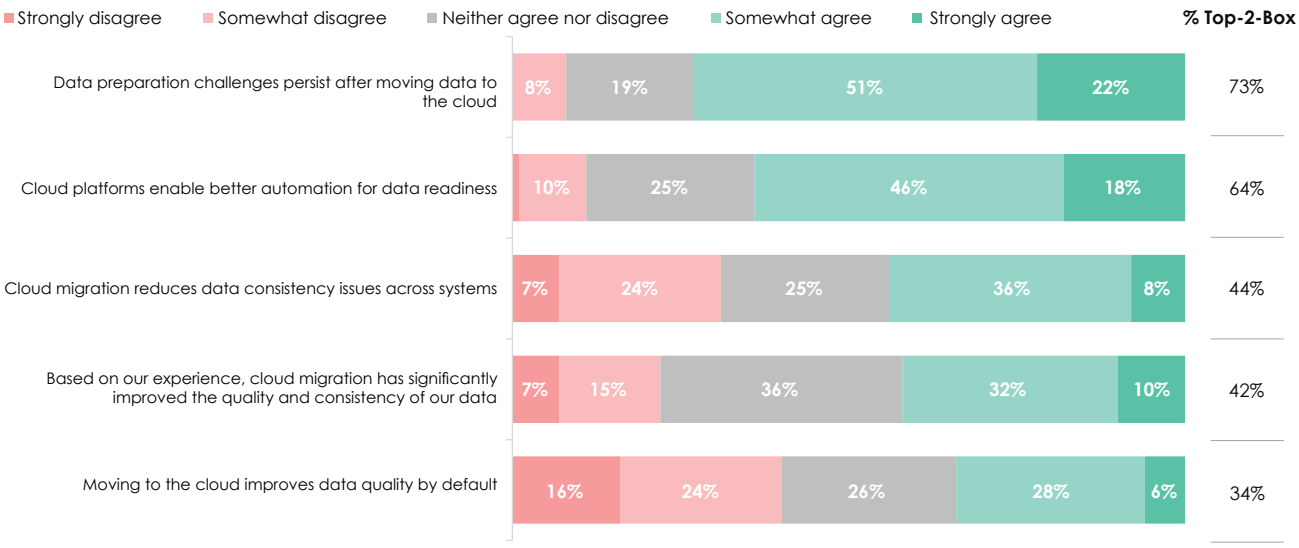
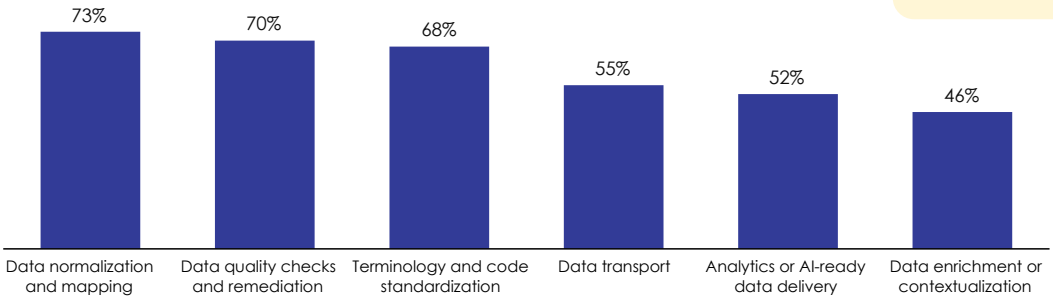


Figure 3. Leaders most often say that interoperability solutions should handle data normalization and mapping, data quality checks and remediation, and terminology and code standardization.

Which functions do you believe interoperability solutions should handle?

On average, leaders believe interoperability solutions should handle **4** of these functions.



Fried said many of these capabilities are available as point solutions. Stand-alone solutions exist that address terminology and code standardization and separate products address data enrichment and contextualization (i.e., metadata management).

"The question organizations ask is: Should I use separate products? Or should I look for a single platform that includes all of these functions?" he said. "I believe it's better to have a single substrate or platform that provides all six of these capabilities, because then you have fewer moving parts and fewer connections to break, there are fewer things to learn, and you're not moving data all over the place."

When organizations are evaluating interoperability solutions with data readiness capabilities, they should make sure the solutions are grounded in deep healthcare expertise, added Fried. "The breadth, depth, and precision that characterize healthcare data are unique in a way that make data quality checks and normalization, data enrichment, and metadata management especially challenging," he said.

Healthcare domain knowledge is essential. For instance, you need a system that understands that a diagnosis of pregnancy does not belong in the health record of a six-year-old boy. "That's why I advise healthcare leaders that they should demand healthcare domain knowledge from their solution suppliers," Fried said.

The role of real-time data

More than three-quarters (77%) of leaders agreed that real-time or near real-time data flows are essential for their organizations' next phase of analytics and/or AI (Figure 4).

Smith noted that the meaning and use of "real-time data" is contextual. For example, if someone is having a heart attack in the emergency room, that data needs to be captured in real time and the appropriate alerts generated in order to save the patient's life. On the other hand, if a researcher is trying to retrospectively determine the flu rates in a particular county, they don't need that data in real time.

"It depends on what question you are answering," she said. The volumes of data required for AI, and the cost to move and analyze that data, are considerations. "Organizations need to prioritize the use cases they need solved and then figure out what the frequency and timeframe in which to retrieve that data are."

"Moving to the cloud doesn't change anything about the data except the location."

JEFF FRIED | Director of Platform Strategy and Innovation | InterSystems

> **"Healthcare leaders … should demand healthcare domain knowledge from their solution suppliers."**
>
> JEFF FRIED

Assessing an organization's data maturity

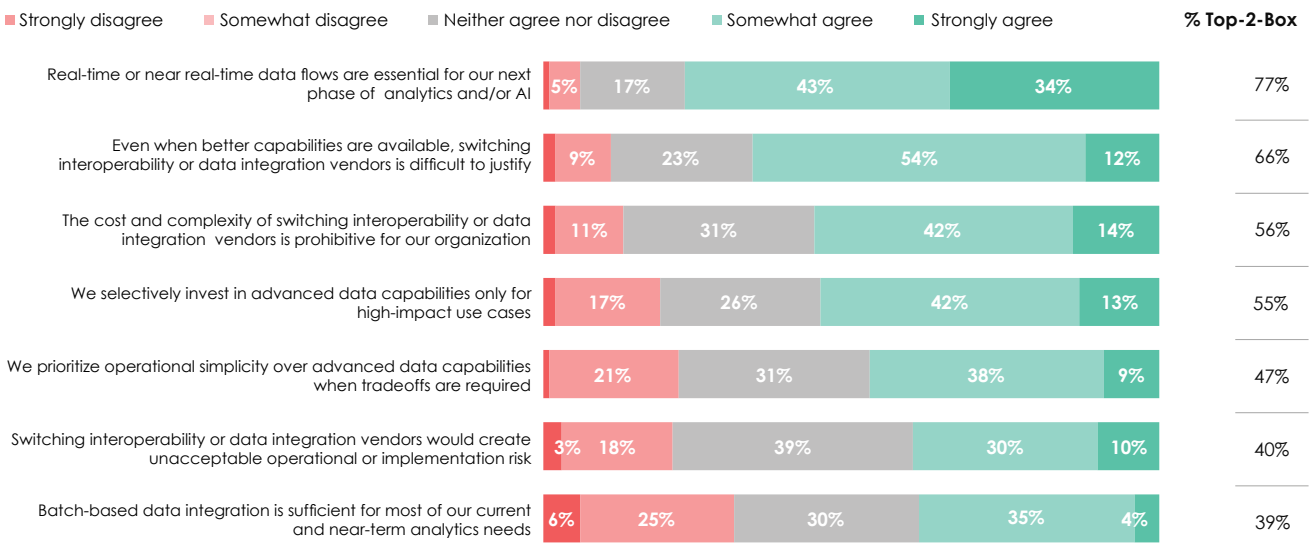
The first step in assessing an organization's data quality maturity is to do a content inventory. At a minimum, the organization needs to know and understand what data it has, where the data comes from, and where it is stored.

Once the organization has completed a data inventory, Smith suggested it can use the data readiness capabilities specified in Figure 3 to identify which capabilities are already in place and where there may be gaps:

1. **Data transport:** Is the raw data being moved from various clinical sources (e.g., EHRs, labs, pharmacy systems) to a centralized environment or data warehouse? Are interfaces in place to accelerate the connection of diverse patient information sources and formats, using industry standards?
2. **Terminology and code standardization:** Are clinical codes and terminologies (e.g., SNOMED CT, LOINC, ICD) aligned to a common standard to ensure that different systems "speak the same language"? Are terminology and code standardization automated?
3. **Data normalization and mapping:** Have disparate data structures been converted into a unified format, such as the HL7 Fast Healthcare Interoperability Resources (FHIR) standard, to ensure data is interoperable across systems? Have individual data elements been mapped to their corresponding target fields? Has disparate data been aggregated into a single, longitudinal patient record?
4. **Data quality checks and remediation:** Is there a process in place to identify and fix issues such as missing values, duplicates, or anomalies to ensure data is accurate, complete, and consistent across all sources?
5. **Data enrichment/contextualization:** Has external data such as social determinants of health (SDOH), claims data, and non-clinical information been incorporated to add value and provide better context for AI models? Has unstructured data such as clinical notes been indexed?
6. **Analytics or AI-ready data delivery:** Are the final, processed datasets available in a format optimized for model training or real-time insights?

Figure 4. Three-quarters say that real-time or near real-time data flows are essential for their organizations' next phase of analytics and/or AI.

Please indicate your level of agreement with each of the following statements.



“One of the things we know about AI is that it can produce beautiful answers that are often wrong.”

JULIE SMITH

“One of the things we know about AI is that it can produce beautiful answers that are often wrong. We are seeing study after study showing this,” Smith emphasized. “So, the results being produced that are based on the data need to be objectively measured and validated in order to have the certainty we require in healthcare.”

Although technology has a key role to play in data quality and readiness, Smith and Fried encouraged organizations not to forget the role that people play in the process. “It’s easy to say data quality is everybody’s job — because it is — but if it’s everyone’s job, it’s really no one’s job,” Fried pointed out. “There should be a person or group in each organization that is charged with data quality. Make it *somebody’s* job.”

To learn more, please visit InterSystems

Reference

1. HIMSS Market Insights. February 2026. *Data Integration in Healthcare* [research report, sponsored by InterSystems]. This research was conducted among 100 U.S. healthcare leaders (managers and above) who were screened for having a role in the selection or strategic oversight of their organizations’ data infrastructure, interoperability, or analytics platforms. InterSystems was not identified as the survey sponsor.

InterSystems®
Creative data technology

About InterSystems

InterSystems helps healthcare organizations unlock more value from their data through interoperability, integration, and data readiness across complex environments. As providers and payers advance analytics and AI strategies, InterSystems helps move beyond simple data exchange to trusted, standardized, and usable data that supports better decisions and outcomes.

Produced by HIMSS®