

WHITE PAPER

Looking to Leverage Powerful Artificial Intelligence Tools?

Get Your Data Governance Ducks in a Row First

PRESENTED BY:



PUBLISHED BY:



Looking to Leverage Powerful Artificial Intelligence Tools?

Get Your Data Governance Ducks in a Row First



CONTENTS:

Why healthcare organizations need to address data issues before developing and adopting AI.....	3
The disconnection dilemma	3
The importance of interoperability	4
The common concerns	6
The governance and AI equation	10
References	11
About.....	11

Why healthcare organizations need to address data issues before developing and adopting AI

In science, a dependent variable is what is observed and measured in an experiment. It is called “dependent” because it changes based on the alterations made to another variable. When it comes to the healthcare industry’s current foray into artificial intelligence (AI), success is based on these dependent variables. Indeed, effective AI implementations are dependent on good data—and good data is, in turn, dependent on good data governance.

The problem? Data governance in healthcare is not as advanced as it needs to be.

“Traditionally, in health we have not done a good job of data governance. When I think about domains that have done a really good job, I think about FinTech or MarTech, where [these industries] have tight control over data. They can tell you to the millisecond what’s going on. We don’t have that ability in health,” said Shenita Freeman, director of solutions leadership at Mathematica, as she joined Ngan MacDonald, data innovations director at Mathematica, and Mike Jennings, senior director of data governance and architecture at Walgreens Boots Alliance, during an online virtual event focused on data governance.

The disconnection dilemma

The healthcare industry deals with a higher level of fragmentation than many other sectors, according to MacDonald. Consider the following: In addition to provider organizations, commercial healthcare entities such as Blue Cross and Blue Shield and United Healthcare, federal programs such as Medicare, state-based programs such as Medicaid, and government agencies that serve specific populations such as the Indian Health Service and Department of Veterans Affairs, all work—with their own systems and nuances—to enhance patient and population health.



Each of these systems owns its own data and then is sometimes called upon to share this data. The data, therefore, needs to carry a benefit for both the original owner and the receiver. The upshot? Healthcare organizations need to be especially cognizant of how they gather, access, leverage, and ultimately share data, according to Freeman.

Optimal data usage, in fact, is especially important in public health. In this segment of the industry, data is intertwined with information from other sources (for example, water, sewage, workforce, and education) in the overall infrastructure. In addition, public health agencies rely on data to understand specific patient populations such as patients who suffer with hepatitis C, HIV, or rare diseases. “All of these layers impact our communities and public health, and data underpins all of it,” Freeman noted.

Currently, however, many public health entities struggle to share data. “What if somebody [or some organization] asks for their data back? In the public health space, that would be a nightmare. Someone comes knocking on the door and says, ‘I want every single piece of my data, no matter how it’s been transformed or permuted back.’ There’s no way to trace that and give the data back,” Freeman explained.

The importance of interoperability

To address this issue, healthcare leaders should keenly focus on interoperability as part of their data governance initiatives. Interoperability is especially important in the current digital environment. Consider the following: 96% of hospitals and 78% of physicians use electronic records.¹ The ability to share and integrate data across these systems, however, is lacking. Unfortunately, data is often hidden in isolated databases, incompatible systems, and proprietary software, making the data difficult to exchange, analyze, and interpret. This fragmentation slows down medical progress, as technologies that rely on data—AI, big data, or mobile applications—cannot be used to their full potential, according to a study published in *NPJ Digital Medicine*.²



What if somebody [or some organization] asks for their data back? In the public health space, that would be a nightmare. Someone comes knocking on the door and says, ‘I want every single piece of my data, no matter how it’s been transformed or permuted back.’ There’s no way to trace that and give the data back.

Shenita Freeman
Director, Solutions Leadership
Mathematica

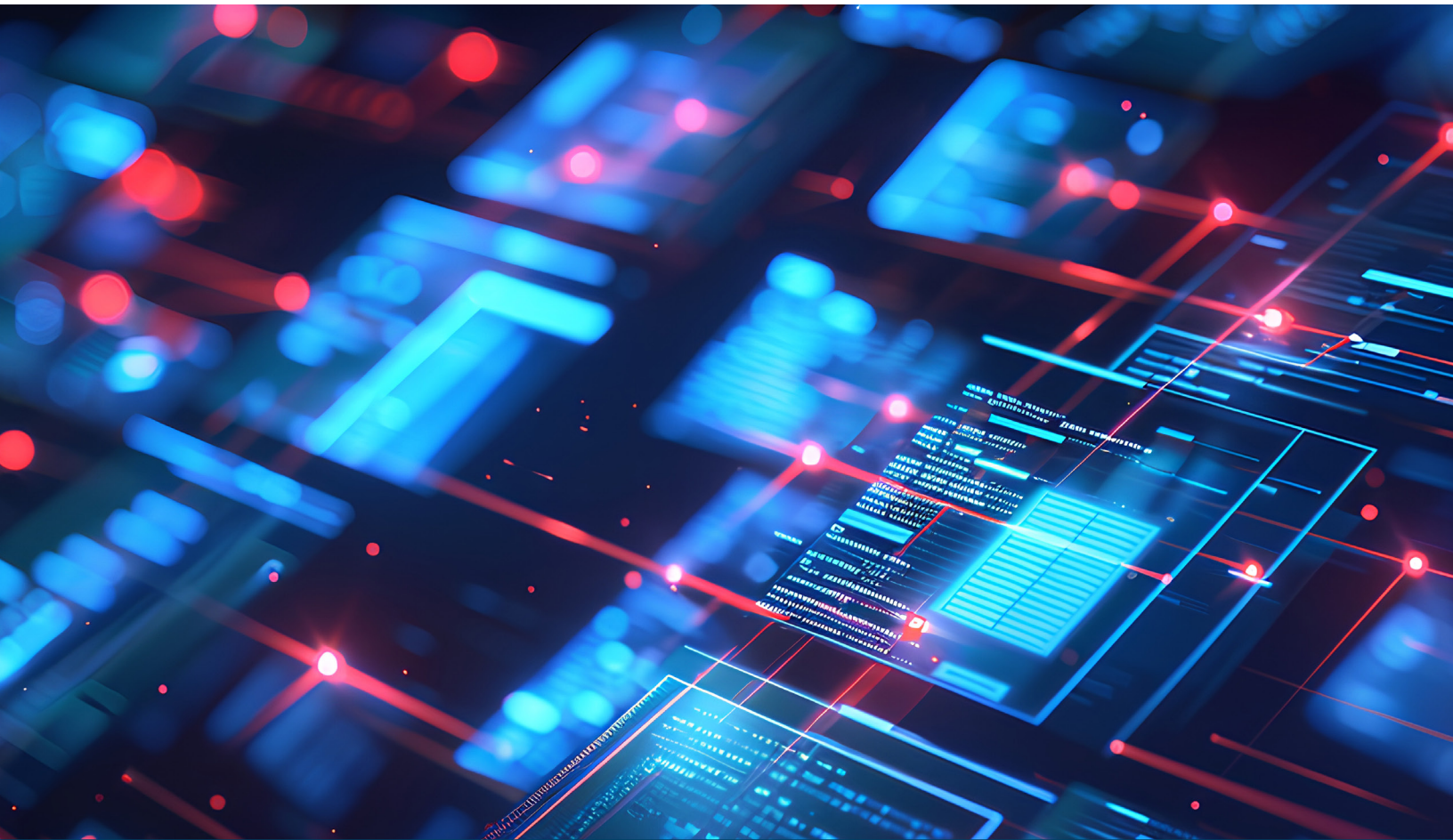


The challenge is to establish interoperability that can support strong data exchanges between health entities, making it possible to improve patient outcomes, enhance operational efficiencies, and drive innovation to benefit providers, health systems, public health agencies, and the private sector.

Such interoperability is especially important when it comes to AI. Reliable interoperability improves AI by providing algorithms with clear data structure and semantics; ensuring validity of analysis results; and creating trust in digital technologies.

To establish this interoperability, healthcare organizations need to:

- Understand the origins of data (a major step for data governance). With a good idea of the origins of information coming into a system, it is possible to subsequently create workflows from this data.
- Prioritize understanding of the values of people and institutions helping to execute the interoperability. To accomplish this, leaders need to talk to all creators, consumers, and owners to better understand the data source.



- Apply the right technologies and capitalize on emerging innovations. For example, AI-assisted data mapping can help enable the connecting and collecting of data needed to achieve interoperability. Healthcare organizations can also use AI to recognize quality defects and assist in mapping similar concepts together.³

The common concerns

While interoperability is important, it's not the only issue that healthcare leaders need to focus on, as they strive to better govern data. To establish the effective data governance that will produce high quality data and, in turn, support the responsible implementation of AI, leaders also need to:

Understand exactly what data governance is. Data governance is the framework of policies, standards, processes, people, and tools that enable an organization or enterprise to manage its data effectively, safely, and within regulatory requirements. The overall goal is to ensure that data is accurate, consistent, secure, and compliant with various regulations, according to Jennings.

In essence, “data governance is the rules of the road and the related enforcement. So, it’s about all the aspects of sharing, storing, using, and destroying data,” Freeman explained.

Variations on this definition, however, often focus on just one aspect of governance (e.g., data sharing) and, therefore, muddy the waters. “People struggle with this because they can’t pinpoint a solid definition. But the truth is that all of the [various interpretations reflect] various aspects of governance. They are not wrong. They are talking about one sliver of governance,” Freeman said.

In addition, data governance definitions can vary from organization to organization and often hinge on data usage maturity. For example, one organization’s governance might focus on creating a data model, while another might zero in on quality issues and a third on data cataloging, Jennings pointed out.



People struggle with this because they can’t pinpoint a solid definition. But the truth is that all of the [various interpretations reflect] various aspects of governance. They are not wrong. They are talking about one sliver of governance

Shenita Freeman
Director, Solutions Leadership
Mathematica

Acknowledge the heightened importance of data governance in healthcare.

Healthcare organizations must govern data to stay in compliance with various regulations and avoid fines. For example, personal health information (PHI) is protected for privacy. At the same time, such data needs to be available to multiple users to improve patient care.

“[Walgreens Boots Alliance is] a [retail] pharmacy and healthcare company... PHI data can’t be mixed between the two. So, we have to understand where the data originates from and where it’s been... and make sure we have the right usage of that data for the right areas,” Jennings said.

In addition, good data governance is more important in healthcare than in some other industries because it can “literally be the difference between life and death, if we don’t get it right,” Freeman said. “If somebody gets the wrong treatment or doesn’t get treatment it can severely impact their life.”

Good data governance also has a trickledown effect when it comes to AI. For example, an AI model that predicts which patients will get sepsis could be 80% to 90% accurate when working with a specific, local population.⁴

“And then you roll it out into the national sphere and find out it’s probably as good as a coin toss” because of shortcomings in the underlying data used to train the model, MacDonald noted.

Understand AI governance. Healthcare leaders need to realize that AI governance is part of the broader governance realm. As such, it’s important to acknowledge that the same principles of good data governance overall apply to AI data governance. Of course, these principles need to be in place before producing the data used to build AI models—as any shortcomings will simply be exacerbated when AI tools are implemented.



Ensure data accuracy. Accuracy is also an extremely important element of health data governance, according to Jennings, whose organization supports both healthcare and retail operations. “If I get something wrong on the retail side, they end up with the wrong pack of gum or whatever. It’s not the end of the world. If I give them the wrong prescription, that’s an issue... So having that accuracy of data, that’s a challenge and it’s important to make sure the right data is being used at the right time,” Jennings said.

Optimize data quality. In clinical settings, staff members often can’t determine how much a treatment will cost before patients receive the service. “That’s a data governance problem. AI is not going to fix that until we implement the governance programs that we need to ensure data integrity and quality,” Jennings said.

To address this, public health organizations need to formally track data. Such data cataloging cannot just be “in someone’s mind who’s been working for a long time in the public health agency,” Freeman said. Instead, formal documentation must track ownership, provenance, and lineage, she added.

Overcome data bias. AI systems sometimes rely on biased data. As such, it is important to feed these systems with data that truly represents patient populations holistically. If AI is created on a subset of data that only represents a small segment of the overall population, the technology will not produce optimal treatment guidance. In some situations, the guidance could even potentially cause harm to specific patients.

Strive for equity in data access. It’s important to make sure that data is available to all who need it throughout an enterprise. “And so the top-of-the-line hospitals may get access, but will our nursing homes get access and be able to use this? So it’s really important to think about that from an equity perspective,” Freeman said.



That’s a data governance problem. AI is not going to fix that until we implement the governance programs that we need to ensure data integrity and quality

Mike Jennings
Senior Director of Data Governance
and Architecture
Walgreens Boots Alliance

AI could exacerbate inequities. “Our communities which have always been marginalized or experienced structural inequity will have yet another layer of technology that will automate and speed up the harms that [they have already experienced]. So it’ll be a wash, rinse, and repeat,” Freeman said. To avoid compounding inequities, organizations need to establish good governance practices before implementing AI so the bias is not amplified.

Provide the right level of data access. Healthcare staff members across the board should not have access to all data. Instead, each staff member needs to get the minimum level of data access that will enable them to do their job optimally.

“You don’t want somebody who’s inside who has access to an EMR looking up their neighbor’s health or looking at Oprah Winfrey’s prescription,” Jennings said.

Balance data privacy and security concerns with the need to know. While seeking to improve outcomes via innovative technologies, healthcare organizations need to balance data access with the need to keep information private and secure. While doing so, leaders must consider national as well as state privacy regulations.

“We need to focus on figuring out what happens to all the data that we currently think has no personally identifiable value and rethink what that means, as it gets commingled with other data sources and is accessible by tools like AI. So as the technology advances, we need to maintain privacy across and within environments... We need to protect privacy in terms of individual elements and how they’re being used,” Freeman said.

In addition, when providing data to external parties for the development of AI, healthcare leaders need to make sure they have appropriate agreements in place to protect PHI. It is important to “understand the use case, the volume of data you’re sharing, what data you’re sharing. Making sure that you’re providing the minimal amount of data needed to do



You don’t want somebody who’s inside who has access to an EMR looking up their neighbor’s health or looking at Oprah Winfrey’s prescription

Mike Jennings
Senior Director of Data Governance
and Architecture
Walgreens Boots Alliance



the use case. You need to factor in all those type of things to determine the right balance between privacy and security [and access]. Every use case needs to be evaluated fresh because everyone is different,” Jennings said.

The governance and AI equation

In sum, healthcare leaders need to understand the connection between data governance, good data, and effective AI tools. Failing to make this connection could result in the misuse of data and subject organizations not only to fines and regulatory sanctions but reputational harm, which is hard to recover from. “Once you’ve lost your reputation, you’ve lost that for good,” Jennings said.

Good data governance could also position public sector organizations to better work with partners across the healthcare ecosystem. Indeed, public-private collaboration could lead to better data usage between the two sectors. Offerings such as [Mathematica’s Health Data Innovation Lab](#) help organizations in the healthcare industry optimally leverage data in this manner. With this digital hub that operates at the intersection of data and policy expertise, government agencies, foundations, medical centers, and other healthcare organizations can collaborate and unlock data insights for innovative health solutions. The Health Data Innovation Lab provides a safe and secure digital sandbox to help healthcare organizations meet their specific data needs.

With the Health Data Innovation Lab and solid data governance practices, healthcare organizations can develop AI tools based on high-quality, reliable data, and expert insights. And, this innovative digital hub can truly make a difference for patients. AI, for example, could be leveraged to help patients better understand their healthcare journeys.

“My mom’s 84, and she doesn’t speak English that well. She goes to the doctor, and the doctor says, reduce your sodium intake, and she has no idea what that means,” MacDonald said. “AI could break that information down and help her understand



My mom’s 84, and she doesn’t speak English that well. She goes to the doctor, and the doctor says, reduce your sodium intake, and she has no idea what that means, AI could break that information down and help her understand that she needs to eat less salt.

Ngan MacDonald
Data Innovations Director
Mathematica

that she needs to eat less salt. And, AI could further help her understand that doesn't mean not adding salt to your food but also using less soy sauce and fish sauce, which contain a lot of salt. If developed properly, AI can provide accurate information that helps patients better navigate care.”

References

1. HealthIT.gov. National Trends in Hospital and Physician Adoption of Electronic Health Records. <https://www.healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records>
2. Moritz, L, Sass, J., Essenwagmer, A., et. al. Why digital medicine depends on interoperability. *NPJ Digital Medicine*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6702215/>
3. Mathematica. Realizing the Value of Healthcare Interoperability. <https://www.mathematica.org/events/realizing-the-value-of-health-care-interoperability>
4. Smith, D. Widely used AI tool for early sepsis detection may be cribbing doctors' suspicions | University of Michigan News (umich.edu). <https://news.umich.edu/widely-used-ai-tool-for-early-sepsis-detection-may-be-cribbing-doctors-suspicions>



Mathematica is a research and data analytics consultancy driven by a mission to improve well-being for people and communities. We innovate at the intersection of data science, social science, and technology to translate big questions into deep insights. Collaborating closely with decision makers and changemakers, we're reimagining the way the world collects, analyzes, and applies data to solve urgent challenges.

For more information please visit: mathematica.org