Your AI Future Depends on Your Data

John W. Moran, Ph.D.

Senior AI and Quality Advisor, Public Health Foundation

November, 2025

When an organization starts using artificial intelligence (AI), the following question should be asked: *Will we shape the future of AI in our organization or will AI shape it for us?* As organizations adopt AI tools and AI technology accelerates, many are deploying AI applications without centralized governance or consistent oversight, creating serious AI organizational and compliance risks.

The potential of AI is immense, but without a data integrity strategy, you are building on sand. If you do not have **Data Integrity,** this means your organization is accelerating AI adoption without the data foundations to support it. This exposes your organization to higher risks of bias, wrong decisions, AI hallucinations, and compliance failures.

Organizations are using AI to optimize their data usage to make better decisions to:
- Increase revenue/control costs
- Attract and retain customers
- Provide better customer service
- Improve efficiency
- Minimize risks to their business

The 2025 Data Integrity Trends & Insights Report[1] shows just how far most organizations have to go:
- **60%** say AI is a key influence on their data programs
- **12%** report that their data are of sufficient quality and accessibility for AI
- **62%** identify lack of data governance as the top challenge holding back AI initiatives

If your organization does not have a trusted data foundation, what is at stake when you use AI?
- Biased practices
- Inaccurate guidance
- Solving the wrong problem quickly

---

[1] https://www.lebow.drexel.edu/sites/default/files/2024-09/drexel-lebow-precisel-data-integrity-trends-insights-2025-outlook.pdf, Accessed 11/01/2025

- Flawed decisions made by AI

Using good data is a key ingredient to making your use of AI successful. You need to have **Data That Can Be Trusted** to fuel AI applications to make good decisions.

What is Trusted Data? These are data that are:
- Accurate
- Consistent
- Relevant
- Rich in context

Having trusted data requires cleaning existing data and databases. This will help ensure that AI is using correct information. Data cleaning is the process of fixing or removing incorrect, corrupt, or incomplete data to prepare it for analysis. Because data and databases differ, there is no single correct method for cleaning these up. Microsoft offers a guide – "Top ten ways to clean your data"[2] – which provides information on the process of cleaning your data and databases. New data are coming in all the time. These should be cleaned before adding to a database.

Once you have clean, trusted data, you want to make sure the data are secure. There are several ways to help make data secure. Here are a few:
- Data encryption
- Identity and access management
- Regular audits and monitoring
- Protection against external and internal attacks
- Secure storage of AI models
- Employee training
- Incident management
- Secure data sharing

Summary:

Most organizations have information spread across the enterprise. There are documents, spreadsheets, data trapped in silos, files sitting in Google Drive, emails in inboxes, patient and customer data, and more. For AI to provide correct analyses that can lead to sound decision making, it requires use of accurate and clean data. This requires all data to be cleaned and integrated together before starting to use AI for specific tasks within an organization. If your organization's data are not clean, AI will not hide your data flaws – it will amplify them, leading to wrong decisions.

---

[2] https://support.microsoft.com/en-us/office/top-ten-ways-to-clean-your-data-2844b620-677c-47a7-ac3e-c2e157d1db19, Accessed 11/01/2025

Remember data fuels AI. You need to fix the existing and incoming new data before using an AI application. Otherwise, you will have "garbage in" to your AI applications and get "garbage out" decisions that are not to be trusted. If the inputs to AI are corrupt, then the AI outputs will be unreliable at best, damaging at worst.

If you do not have **Data Integrity,** this means your organization is accelerating AI adoption without the data foundations to support it. This exposes your organization to higher risks of bias, AI hallucinations, irrelevant results, and compliance failures along with loss of integrity within your organization and community.

Develop a regular cleaning process and develop a culture of data excellence throughout the organization. Trusted data lets you shape the future of AI in your organization, rather than AI shaping it for you.