

Data Does Not Lie — Until It Does

By John W. Moran, Ph.D.¹ & Amanda McCarty²

March 2026

Recently, someone confidently stated, “Data does not lie.” It is an appealing phrase — concise, reassuring, and seemingly objective. But it is incomplete. Data may not lie intentionally, yet it can mislead profoundly when it is poorly collected, poorly defined, misinterpreted, or stripped of context.

In the age of artificial intelligence (AI), the stakes have changed. Data quality is no longer merely a technical concern — it is a leadership responsibility with ethical, operational, and strategic consequences.

Data is often perceived as neutral — an unbiased reflection of reality. In truth, data reflects the systems, processes, assumptions, and human behaviors that produced it. The long-standing principle of “garbage in, garbage out” in information science reminds us that flawed inputs inevitably produce flawed outputs³. AI systems are no exception. If those systems are flawed, the data will faithfully reproduce those flaws.

When organizations rely on inaccurate, incomplete, or siloed data, they do not simply make small errors. They make strategic misjudgments — allocating resources incorrectly, misreading trends, or responding to symptoms instead of root causes. The real danger is not that data lies. The danger is that leaders believe it cannot.

Most poor data is not malicious; it is accidental. It emerges when organizations:

- Measure activity instead of outcomes
- Fail to define metrics clearly
- Rely on ambiguous categories such as “other”
- Collect or enter information inconsistently
- Neglect to define units of measure
- Change calculation methods without documentation
- Separate data across siloed systems

The Data Management Body of Knowledge (DAMA-DMBOK) framework emphasizes that clear definitions, governance, and stewardship are essential to preventing

¹ Senior AI and Quality Advisor, Public Health Foundation

² Performance Improvement Specialist, Public Health Foundation

³ <https://www.ebsco.com/research-starters/computer-science/garbage-garbage-out-gigo>, Accessed 3/5/2025

inconsistency and misinterpretation⁴. Without these safeguards, organizations create variability at the source.

Consider a simple metric: *Number of Flu Shots Administered Per Day*.

What question is this answering? Is it measuring population protection? Staffing productivity? Inventory management? Without clarity of intent, a metric becomes noise.

Data integrity begins at the point of capture. Organizations must:

- Walk the process being measured
- Identify appropriate collection points
- Establish clear collection rules (who, when, where, how often)
- Train data collectors or automate when possible
- Assign ownership for data entry and stewardship
- Standardize calculation rules and rounding conventions

Data governance is not clerical work. It is a foundational strategy.

AI: The Great Amplifier

AI does not create bias. It scales it. As the National Institute of Standards and Technology notes in its AI Risk Management Framework, risks associated with AI systems frequently stem from data quality, representativeness, and contextual limitations⁵.

AI systems learn from the data they are given. When that data is incomplete, siloed, or distorted, AI does not correct the error — it reinforces it at speed and scale.

Organizations increasingly rely on AI to:

- Forecast demand
- Allocate resources
- Identify risk
- Detect patterns
- Guide operational decisions

But AI tethered to fragmented application data sees only narrow slices of reality. It cannot produce enterprise-wide insight if it lacks integrated, trusted data. Once flawed

⁴ DAMA International. (2017). DAMA-DMBOK: Data management body of knowledge (2nd ed.). Technics Publications.

⁵ National Institute of Standards and Technology (NIST). (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). U.S. Department of Commerce. <https://www.nist.gov/itl/ai-risk-management-framework>, Accessed 3/5/2026

data is embedded in a model, it becomes difficult to unwind. Poor inputs become automated outputs — and flawed guidance becomes institutionalized.

Organizations that wish to make sound, data-driven decisions must move beyond simple data accumulation. They must build trusted data systems.

Trusted data is:

- Accurate – free from systematic error
- Consistent – collected and calculated uniformly
- Relevant – aligned with strategic questions
- Rich in context – interpretable within operational reality
- Accessible – available when and where decisions are made

Well-designed dashboards can provide early warning signals, detect emerging trends, and support root cause analysis. But dashboards built on unreliable data create a false sense of confidence — a dangerous illusion of control.

The organizations that will lead in the AI era are not those with the most data — but those with the most trusted data.

Building a successful AI framework does not begin with algorithms. It begins with data integrity. Breaking down data silos, clarifying definitions, training staff, standardizing processes, and establishing governance structures are not technical afterthoughts. They are strategic imperatives.

Data does not lie. But poorly designed systems distort the truth.

In the AI era, trusted data is no longer optional. It is essential.