

Data Auditing: Building Trust in Artificial Intelligence

Do you have something to say about this article?

Visit the *Journal* pages of the ISACA® website (www.isaca.org/journal), find the article and click on the Comments link to share your thoughts.

<https://bit.ly/2oGLYte>

亦有中文简体译本

www.isaca.org/currentissue

While IT risk management driven by standards/frameworks such as the International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) ISO/IEC 27001:2005 *Information technology—Security techniques—Information security management systems—Requirements*,¹ British Standard (BS) 7799-3 (2006) and ISACA's Risk IT (2009) have been around for more than a decade, in reality, IT risk management can trace its roots back to the very

reason for the establishment of ISACA® 50 years ago. More recently, the advent of regulatory risk data reporting regulations such as Basel's BCBS239² and the US Dodd-Frank Wall Street Reform and Consumer Protection Act (through Dodd Frank Annual Stress Testing [DFAST] and the Comprehensive Capital Analysis and Review [CCAR])—both introduced this decade and both outcomes of the financial crisis of 2008^{3,4}—have shone light on data risk. Indeed, BCBS239 requires that the controls around key bank data should be as rigorous as those applied to accounting data.⁵

These regulations formalized data control for risk data in terms of data quality (accuracy, integrity, completeness, timeliness) initially only for Globally Systemically Important Banks (GSIBs), with the goal of controlling data risk. The regulations also introduced requirements for metadata and the articulation of enterprise roles and responsibilities for key enterprise data.⁶

While these data controls help protect the stability of the global financial system, as a whole, data controls determine the extent to which data are fit for purpose for applications of data such as reporting and analytics, and also for contemporary applications of data such as in artificial intelligence (AI) and machine learning (ML), because poor or biased data are a primary risk to quality AI outcomes.⁷ While these applications depend on quality, unbiased data to provide reliable outcomes, there is little that provides assurance to the end user of the quality of the inputs into the algorithms that produce the outcomes. Perhaps there is a means to close that gap.

Data Audits as a Response to the Trust Gap in AI

The trust gap is unfortunate. Indeed, the matter of trust in AI is highly topical, not only in business,^{8,9,10} but also in government.^{11,12,13} With headlines such as "If Your Data Is Bad, Your Machine Learning Tools Are Useless,"¹⁴ and "Data Quality and Artificial

Guy Pearce, CGEIT

Has served on various enterprise boards and as chief executive officer of a multinational retail credit operation. This experience provides him with rich insights into the real-world expectations of governance, risk, IT and data. Capitalizing on two decades of corporate digital transformation experience, he instructs a course at the University of Toronto (Ontario, Canada) targeting boards and the C-suite on digital transformation, based on a governance gap he identified while researching an article published in the *ISACA® Journal*. He is the recipient of the 2019 ISACA® Michael Cangemi Best Author Award and consults on digital transformation, with a special interest in its governance, risk, compliance and data aspects.

Intelligence—Mitigating Bias and Error to Protect Fundamental Rights,”¹⁵ it is clear that the fitness of data for emerging technologies such as AI and ML is in the spotlight, adding to the already global laser focus on the ethical uses of these technologies.^{16, 17, 18}

Fortunately, data audits can help close this gap by helping provide end users assurance about the quality of the outcomes from those technologies, thus helping improve the reliability of those outcomes for decision-making (figure 1).

Key Concepts

For the practitioner, governance instruments such as audits can be better understood when their activities are contextualized in terms of the subject’s problem statements. In this respect, the following ideas are important for understanding the context and structure of a data audit.

Data Fit for Purpose

Data are fit for purpose when they are in a state sufficient to perform the role expected of them. For example, the data in a bank statement are expected to be flawless; they must present an accurate record of the inflows and outflows of cash into an account. If they do not, then the data are not fit for purpose.

Just like a financial audit, a data audit reconciles back to source systems. In so doing, it provides assurance that data underlying, for example, the bank statement are fit for purpose.

The “human bias, generalizability, conflicts of interest, politics, and prejudice” implicit in data,¹⁹

however, already begin to beg the question of whether data are fit for purpose for AI.

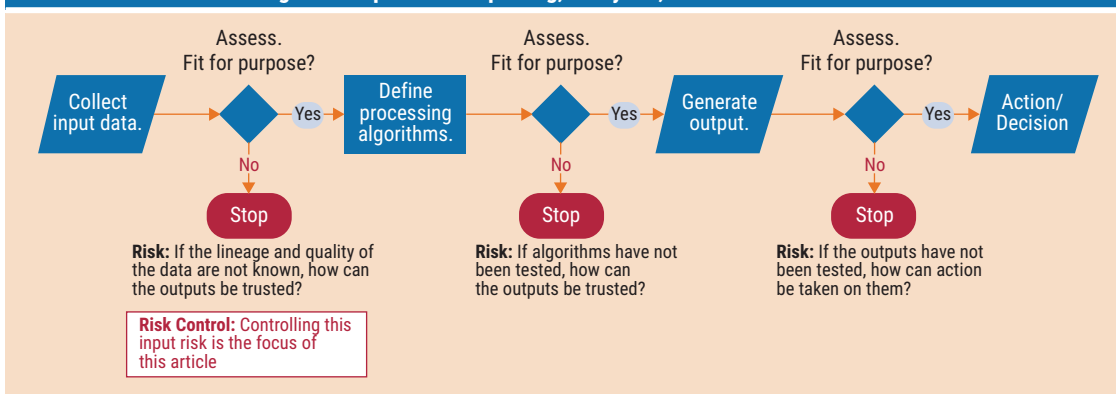
Data as a Risk

If data are not fit for purpose, they expose an organization/individual to risk as soon as a decision is based on them. Given the attention paid to the role of emerging technologies in the future, there is little attention paid to the requirement for quality input data. Is this oversight perhaps the source of another ethical dilemma that could negatively affect humans when AI becomes ubiquitous?

“ IF DATA ARE NOT FIT FOR PURPOSE, THEY EXPOSE AN ORGANIZATION/ INDIVIDUAL TO RISK AS SOON AS A DECISION IS BASED ON THEM. ”

If risk can be defined informally as the difference between the expected outcome of an event and its actual outcome, and if a bank statement reflects a financial position different to the actual position, the deviant outcome has incurred risk for the organization—a gap that could be closed by instituting sustainable controls. Because poor data incur risk, data risk needs to be managed in the same way recommended for any risk (by the process of identifying, assessing, controlling and monitoring it), with the goal of increasing the transparency of the quality of the input data.²⁰

Figure 1—Inputs Into Reporting, Analytics, AI and ML Procedures



“THE CHALLENGE FOR THE DATA AUDITOR IS TO SEEK EVIDENCE THAT THE DATA IN A DOWNSTREAM SYSTEM REMAIN AN ACCURATE REFLECTION OF THE DATA AT SOURCE.”

In an AI context, the kinds of data risk that should be considered include errors of representation (data that “do not well cover the population they should cover”) and errors of measurement (“data that do not measure what they intend to measure”).²¹ Also, the kinds of questions that should be asked concern metadata that often do not readily and accessibly exist in practice, such as metadata about the source of the data, the coverage of the data, the nature of the missing data, the time frame applicable to the data and the geography in which the data were collected.²² Coupled with bias, these ultimately need to be mitigated in the interests of “protecting fundamental rights” when pursuing AI.²³

Risk: Data Statics

Data attributes such as quality, most business and technical metadata, security, and privacy say something about a data element at a point in time. They can be called “statics.” The relationship between statics such as quality and AI is symbiotic.²⁴ The higher the quality of data, the higher the quality of (human and) AI. The better the AI, the greater the demand for quality data.

It is easiest to perform static data audits, because data either do or do not meet a specific quality control. These audits, however, need the enterprise to have defined its most important data (it is not feasible in large organizations to audit all the organization’s data) and the thresholds (controls) of the various measures of quality for those data.

From a static audit perspective, the organization needs to identify the data elements that are critical to the organization and define quality measures and thresholds applicable to these. Given this, the auditor would then:

- Determine whether the portfolio of critical data elements is reasonable
- Determine the extent to which each data element met or exceeded the defined threshold
- Determine whether controls were established and exercised in cases where the threshold was not met and controls were deemed necessary to mitigate risk in a given time frame

Risk: Data Flows

Data elements such as data lineage and data transport validation (DTV) are flow attributes associated with the data extract-transfer-load (ETL) processes. (Lineage is a type of technical metadata; it provides information on where the data come from and the path they took to get there.)

Lineage plots the path of data as they move, by ETL, from their source all the way downstream, often over many hops, to their destination. While there are many traditional tests for the success of individual ETLs—unfortunately, the most popular of which seems to be the ubiquitous, but rudimentary, row-count test—DTV specifically measures lineage quality (i.e., the extent to which data maintain their original value as they flow from source to target systems across ETLs). While many measures of ETL performance are statics, both lineage and DTV tell something about a data element over time. They can be called “flows.”

Data lineage is key to effective AI in areas such as neural networks, natural language processing (NLP), ML and deep learning.²⁵ For AI to be effective, it needs the data feeding its algorithms and models to be well-understood, and data lineage is a key part of creating that understanding.²⁶

Given the importance of lineage in AI, the role of lineage in creating quality AI outcomes is enhanced by DTV. Data in the organization’s operational systems—at source—is the purest form of the organization’s data, whether they are of good or bad quality. It is an unadulterated reflection on everything that is right and wrong with the organization’s data discipline.

The challenge for the data auditor is to seek evidence that the data in a downstream system remain an accurate reflection of the data at source. An audit volume challenge exists both in the number of paths data can take as they move downstream and in how many hops they take between databases. So, a sampling strategy might be needed as with other audits. Types of DTV include:

- Delta ETL testing (testing the quality of incremental data loads)
- Data transformation tests
- Data derivation tests
- Technical metadata tests (e.g., tests for the maintenance of data types, lengths, precision)

While DTV is ideally performed at the column level, in practice, it is more often performed at the table level given the effort involved in performing DTV for every critical data element, potentially more than hundreds of hops each in the world's largest banks.²⁷

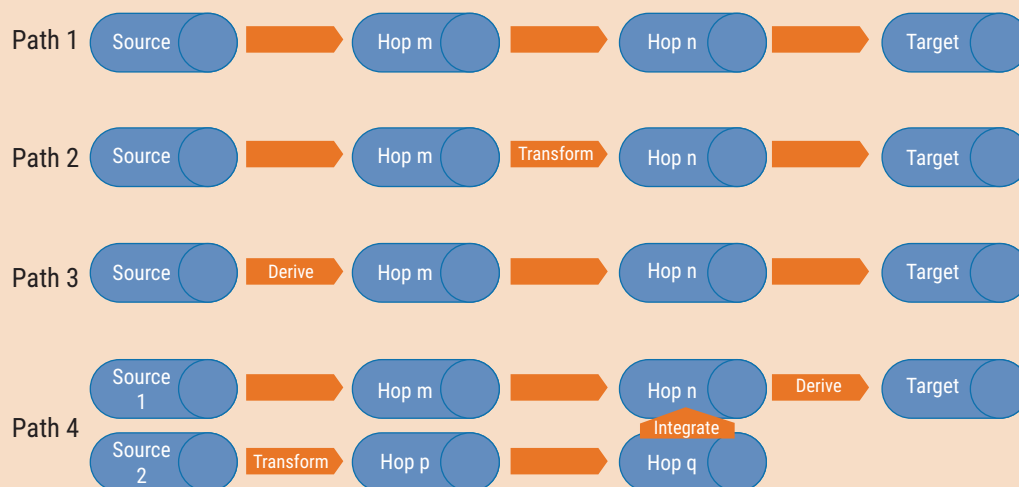
From an audit perspective, flow audits are more difficult than static audits, and they are made even more complex given that data may have undergone transformations or even been used to derive new data during their source to target flows (**figure 2**). Some examples include:

“ POOR DATA MIGRATION IS A MAJOR CAUSE OF THE POOR PERFORMANCE OR EVEN THE FAILURE OF A NEW SYSTEM. ”

- **Transformation**—The male/female flags in legacy systems were often recorded as 0/1, respectively. Many are transformed into M/F or even male/female, respectively, to make them more meaningful to humans.
- **Derivation**—A system may store a customer's date of birth, but often a customer's age is required, so it is calculated and sometimes stored in a downstream table.
- **Mass flows**—Application upgrades involve flowing data from the old system to the new system. Poor data migration is a major cause of the poor performance or even the failure of a new system. Much of this is driven by, for example, poor data quality,²⁸ poor business (semantic) metadata and/or poor subject matter expertise,²⁹ and matter generally unaddressed in AI deployments.

It should be noted that while static audits provide a single layer of insight, they can be converted into

Figure 2—Simplified Examples of the Paths Data Can Take From Source to Target



The letters m, n, p and q indicate multiple hops between source and target. Large organizations could have more than 100 hops.

Enjoying this article?

- Read *Auditing Artificial Intelligence*. www.isaca.org/auditing-AI
- Learn more about, discuss and collaborate on audit and assurance in ISACA's Online Forums. <https://engage.isaca.org/onlineforums>

flows by plotting their performance over time (**figure 3**). In **figure 3**, while the third quarter (Q3) 2019 completeness measure for data element "addr_state" may be above a threshold (a pass mark in a static analysis), it should be noted that the flow analysis shows that the measure has started to dip, which creates a situation that may need management attention.

Getting the Most Benefit From the Interview Stage of the Audit

A simple question and answer session can easily provide a high-level view of the major static and flow attributes of the required data (**figure 4**).

Like most other audits, incredible insights can be gained at the interview stage, provided good questions are asked. It does not take much additional effort to see how deeper answers can be gained by examining the static or flow measures (column 3 in **figure 4**) relevant to the question.

Conclusion

As discussed, it is important to outline how data audits could help build trust in AI. The role data regulations have had on increasing organizational data discipline, especially in financial services, and

their role in serving as critical input into the quality of data-driven decision-making and even autonomous decision-making via AI is impactful.

“ COUPLED WITH A MEANS TO VALIDATE THE AI MECHANISM (ANOTHER CONTEMPORARY CHALLENGE), DATA AUDITS ARE THE PERFECT COMPLEMENT TO QUALITY APPLICATIONS OF AI. ”

Various concepts introduced herein—such as data fit for purpose and data as a risk, and the difference between static and flow measures of data quality—highlight the contemporary social need to improve trust in applications such as AI and emphasize that data audits could play a meaningful role in building this trust. However, this discussion does not consider processing algorithm quality, AI or otherwise.

Figure 3—Adding a Flow Context to Static Measures

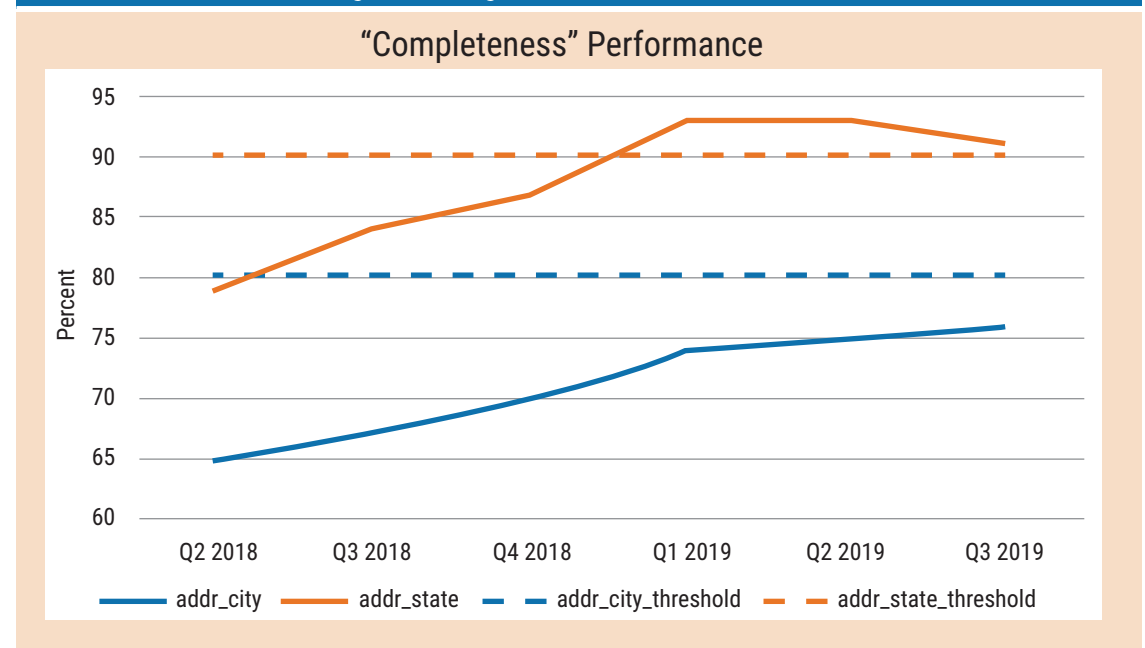


Figure 4—Examples of Questions That Help Identify Data Risk

Question	Type	Attribute	Examples of Associated Questions
Who?	Static (point in time)	Data stewardship	Who in the organization understands the data element best, and what is this person's role in this?
What?	Static	Metadata	What is the nature of the data dictionary used to describe the data? Does it exist at all, is it superficial, and where it exists, is it current? What is its relationship between data elements (conceptual architecture)?
When?	Static	Data operations (data life cycle management)	When are the data available? When are they created, updated and/or deleted? What are the dependencies on these times?
Where?	Flow (across time)	Lineage	Where do the data originate and what is their route to their current location? What assurance is there that the data have not been wittingly or unwittingly altered before they reached their destination?
Why?	Static	Metadata/conceptual data model	Why are the data important and, as a result, what do they depend on and, in turn, what depends on them?
How?	Static	Analytics, AI, ML, business intelligence (BI)	How are the data being used? Are they used in absolute terms, or are they the basis for a derived (calculated) field? What is the quality of the associated master and reference data?
How much?	Static	Quality	What are the measures of, for example, data accuracy, validity, completeness, uniqueness and timeliness, and what is the performance against these?

Data audits present strong business benefits beyond regulatory compliance to improving business operations and safeguarding data integrity.³⁰ An untold story is the data audit's potential to help build trust in AI—particularly from the point of view of validating the quality of the input data—thereby increasing trust in AI's outputs. Coupled with a means to validate the AI mechanism (another contemporary challenge), data audits are the perfect complement to quality applications of AI.

While a standard way to attest the AI mechanism will go a long way to creating trust in AI, presenting the pairing of a data attestation with the outcome of a digital transformation project involving AI to clients improves their ability—willingness even—to explore the implications of those outcomes rather than to debate its inputs.

Ultimately, “[s]tarting an AI project without checking the data first is like building an F-1 race car without understanding what quality, type, specification, or sustainability of fuel you want to use,”³¹ a sure recipe for an untrustworthy performance.

Endnotes

- 1 International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC), ISO/IEC 27001:2005 *Information technology—Security techniques—Information security management systems—Requirements*, Switzerland, 2005, <https://www.iso.org/standard/42103.html>
- 2 Bank for International Settlements, *Principles for Effective Risk Data Aggregation and Risk Reporting*, January 2013, <https://www.bis.org/publ/bcbs239.pdf>
- 3 TechTarget, Dodd-Frank Act, <https://search.financialsecurity.techtarget.com/definition/Dodd-Frank-Act>
- 4 *Op cit* Bank for International Settlements
- 5 *Ibid.*
- 6 *Ibid.*
- 7 Korolov, M.; “AI’s Biggest Risk Factor: Data Gone Wrong,” *CIO*, 13 February 2018, <https://www.cio.com/article/3254693/ais-biggest-risk-factor-data-gone-wrong.html>
- 8 McKendrick, J.; “Learning to Trust Artificial Intelligence: An Optimist’s View,” *Forbes*, 9 June 2019, <https://www.forbes.com/sites/joemckendrick/2019/06/09/learning-to-trust-artificial-intelligence-an-optimists-view/#6484c1131169>

- 9 IBM, "Building Trust in AI," <https://www.ibm.com/watson/advantage-reports/future-of-artificial-intelligence/building-trust-in-ai.html>
- 10 Rao, A.; E. Cameron; "The Future of Artificial Intelligence Depends on Trust," *Strategy+Business*, 31 July 2018, <https://www.strategy-business.com/article/The-Future-of-Artificial-Intelligence-Depends-on-Trust?gko=ffcac>
- 11 Government of Canada, *Accountability in AI—Promoting Greater Social Trust*, 4 December 2018, <https://www.ic.gc.ca/eic/site/133.nsf/eng/00005.html>
- 12 Future of Life Institute, "AI Policy—United States," <https://futureoflife.org/ai-policy-united-states/?cn-reloaded=1>
- 13 Organisation for Economic Co-operation and Development, "OECD Creates Expert Group to Foster Trust in Artificial Intelligence," 13 September 2018, www.oecd.org/innovation/oecd-creates-expert-group-to-foster-trust-in-artificial-intelligence.htm
- 14 Redman, T. C.; "If Your Data Is Bad, Your Machine Learning Tools Are Useless," *Harvard Business Review*, 2 April 2018, <https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless>
- 15 European Union Agency for Fundamental Human Rights, *Data Quality and Artificial Intelligence—Mitigating Bias and Error to Protect Fundamental Rights*, Austria, 2019, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-data-quality-and-ai_en.pdf
- 16 Bossman, J.; "Top Nine Ethical Issues in Artificial Intelligence," *World Economic Forum*, 21 October 2016, <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>
- 17 Metz, C.; "Is Ethical AI Even Possible?" *The New York Times*, 1 March 2019, <https://www.nytimes.com/2019/03/01/business/ethics-artificial-intelligence.html>
- 18 Levin, S.; "Bias Deep Inside the Code: The Problem With AI 'Ethics,' in Silicon Valley," *The Guardian*, 29 March 2019, <https://www.theguardian.com/technology/2019/mar/28/big-tech-ai-ethics-boards-prejudice>
- 19 Car, J.; A. Sheik; P. Wicks; M. S. Williams; "Beyond the Hype of Big Data and Artificial Intelligence: Building Foundations for Knowledge and Wisdom," *BMC Medicine*, vol. 17, article no. 143, 17 July 2019, <https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-019-1382-x>
- 20 Deloitte, "Making Data Risk a Top Priority," *The Wall Street Journal*, 23 April 2018, <https://deloitte.wsj.com/riskandcompliance/2018/04/23/making-data-risk-a-top-priority/>
- 21 Moltzau, A.; "Facebook vs. EU Artificial Intelligence and Data Politics," *Towards Data Science*, 20 July 2019, <https://towardsdatascience.com/facebook-vs-eu-artificial-intelligence-and-data-politics-8ab5ba4abe40>
- 22 *Ibid.*
- 23 *Ibid.*
- 24 Sykes, N.; "What to Know About the Impact of Data Quality and Quantity in AI," *SmartData Collective*, 17 November, <https://www.smartdatacollective.com/what-to-know-about-impact-data-quality-quantity-in-ai/>
- 25 Harris, J.; "Data Lineage: Making Artificial Intelligence Smarter," *SAS*, https://www.sas.com/en_us/insights/articles/data-management/data-lineage—making-artificial-intelligence-smarter.html
- 26 *Ibid.*
- 27 Guru99's article "ETL Testing or Data Warehouse Testing Tutorial," <https://www.guru99.com/ultimate-guide-etl-datawarehouse-testing.html>, is a good introduction for those interested in different ETL tests and the scenarios in which to perform them.
- 28 Miller, K.; "Seven Reasons Data Migrations Fail," *Premier International*, <https://www.premier-international.com/articles/7-reasons-data-migrations-fail>
- 29 Bertolucci, J.; "Ten Big Data Migration Mistakes," *InformationWeek*, 9 August 2012, <https://www.informationweek.com/big-data/software-platforms/10-big-data-migration-mistakes/d/d-id/1105724>
- 30 Mazer, M. S.; "Ten Business Benefits of Effective Data Auditing," *Enterprise Systems Journal*, 18 February 2004, <https://esj.com/articles/2004/02/18/ten-business-benefits-of-effective-data-auditing.aspx>
- 31 Chan, B. K.; "Why Data Governance Is Important to Artificial Intelligence," *Medium*, 10 January 2019, <https://medium.com/taming-artificial-intelligence/why-data-governance-is-important-toartificial-intelligence-fff3169a99c>