

The Era of Big Data is coming to an end as the focus shifts from how we collect data to processing that data in real-time. Big Data is now a business asset supporting the next eras of multi-cloud support, machine learning, and real-time analytics.

The Era of Machine Learning stands out in its focus on analytic models, algorithms, model training, deep learning, and the ethics of algorithmic and deep learning technologies. Machine Learning requires much of the same work needed to create clean data for analytics, but also requires additional mathematical, business, and ethical context to create lasting and long-term value.

Third, machine learning and data science are the next generation of analytic analysis and will require their own new data management efforts. The creation of testing data, synthetic data, and masked data at scale as well as the lineage, governance, parameter and hyperparameter definitions, and algorithmic assumptions require efforts beyond traditional Big Data assumptions. The most important consideration here is to use data that does not serve the business well due to small sample size, lack of data sources, poorly defined data, poorly contextualized data, or inaccurate algorithmic and classification assumptions. In other words, don't use **data that lies**. Lying data leads to outcomes that are biased, non-compliant, inaccurate, and can lead to issues such as Nick Leeson's destruction of Barings Bank in 1995 or Societe Generale's \$7 billion trading loss based on well-manipulated trades by Jerome Kerviel. AI is now the new potential "rogue trader" that needs to be appropriately governed, managed, and supported.