

Autonomously Reduce Bad Data Risk with DataBuck

Catch Data Errors Before Your Business Partners Do

Challenges

Growing data volume, microservices, number of platforms and data complexity makes traditional data validation solutions costly to scale and difficult to manage.

Data errors spread like cancer and cause serious problems:

- Lack of institutional trust in data team
- Non-usable Lake, Cloud, Warehouse
- Enormous loss of investment in advanced infrastructure and tools
- Quasi-reliable reports, analytics, BI, predictions
- Extra time and money for rework

Status quo – old methods are not scalable:

- Too many essential rules to code
- Unknown data-risks slip through
- Expensive, takes too long & too many people to implement
- Needs hardware w/more horsepower

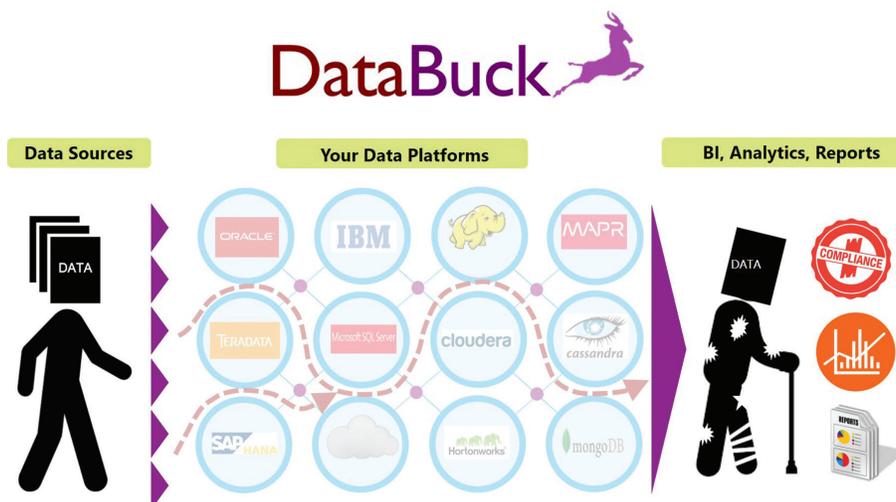
Solution: Validate Data Upstream with DataBuck

Catch bad data by autonomously sensing changes in the DNA of data using AI/ML. Continuously monitor the health of 1,000's of data sets in a few clicks in a highly scalable process.

When is DataBuck most powerful?

Current data monitoring and validation tools and processes fare very poorly under these conditions. A high rating on at least two of these metrics implies a high risk of data errors spreading in your systems. DataBuck is perfect for:

- Cloud/Lake use
- Dynamic data (for e.g., operational and transactional data have constantly changing boundaries)
- High volume of data
- New sources or changing structures of data
- New uses for data (unvalidated data for analytics, customer insights, etc.)
- Tremendous growth in microservices



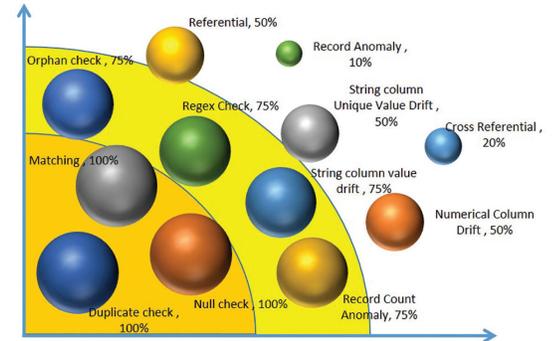
Data Validation Challenge

Only 5- 10% of data in any organization gets validated and trustable. A large volume of data in a variety of formats flowing via old and new platforms (on-prem, Cloud, Lake, Hadoop, RDBMS, etc.) has decreased the reliability of data.

Data Engineers are struggling to answer these basic questions:

- (Lake) "How do I reduce the bad-data risk for a large volume of data?"
- (Cloud) "Is there a Cloud-native tool to identify bad-data?"
- (Complexity) "How to ensure we can catch bad-data from unexpected risks, the unknown-unknowns?"
- (Cost) "How do we reduce data-rework, with fewer h/w resources, with fewer people, in a shorter time and at lower cost?"

Traditional tools were architected for on-prem sources and for regular/ small data. Their older architecture cannot process bigger and complex data efficiently in new plat-forms. Lake and Cloud data are not easy to handle and need a natively built solution .

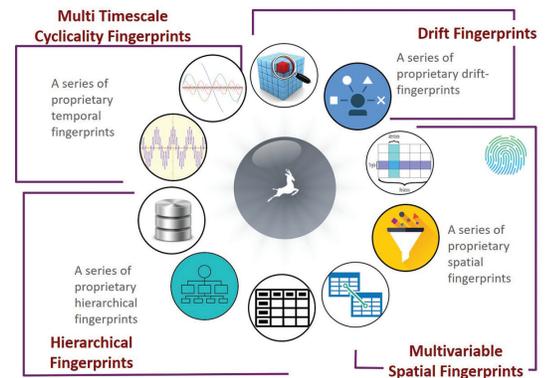


New Paradigm

- **Future-Proof Data Architecture** – Tool must not be propped up on old architectures. Even if data volumes are low today, have the flexibility to scale as needed.
- **New data source compatibility** – Must natively handle Cloud, Hadoop, SQL, RDBMS, etc.
- **Intelligence** – Must detect known and unknown risks from bad-data, by monitoring 100's of thousands of columns in 1,000's of Tables continuously and autonomously.

DataBuck – Autonomous DNA Fingerprint Monitoring of Data

- Just like a biological DNA, DataBuck captures the core essence of data characteristics
- Eliminates the hard to catch unanticipated data errors ("unknown, unknowns") and the anticipated errors
- AI/ML autonomously tracks 1,000's of Data DNA Fingerprints
- In-situ data validation minimizes data movement. Takes the validation rules to the data.
- In a few clicks, w/o coding reduce bad-data risk in 1,000's of tables.
- Cloud native, up and running in hours, not months as with other tools
- Spark and specialized algorithms make it >10x faster than any other tool.



Build Natively for Cloud & Lake
Designed for handling Cloud & Lake.
Architected for extremely fast processing, at low cost



Machine Learning
Leverages machine learning algorithms to establish & monitor data DNA fingerprints and identify bad data issues autonomously



Code Free
Easy to navigate wizard-based user interface eliminates need for sophisticated technology resources. Data's DNA fingerprint and validation checks can be set up within minutes

DataBuck – Be up and Running in a Few Hours!



DataBuck – Benefits



People
Productivity
Boost >80%



Reduction in
Unexpected
Errors: 70%



Cost Reduction
>50%



Time Reduction
to Onboard data
set ~90%



Increase in
processing
speed >10x



Cloud
Native

DataBuck – User Reviews

“What took my team of 10 Engineers 2 years to do, DataBuck could complete it in less than 8 hrs”

– VP Technology, Enterprise Data Office, Major US bank

“Streamlining the DQ monitoring and validation process w/ DataBuck has reduced our time-to-market. With fewer resource we auto discover DQ rules, which also self-heals as the data evolves.”

– Head of Enterprise Data Quality Monitoring, Major US bank

“AML is on the rise. We have data from 10 countries in different formats and standards that need to be validated. We could not keep up doing it manually. DataBuck has automated and streamlined our data pipeline.”

– Senior Executive Technology Office, Top-3 African bank

“DataBuck’s Data Quality automation does 80% of the heavy lifting for us with just 5% of the effort.”

– CIO of US Financial Services firm

“DataBuck can really add a lot of headcount efficiency for us. This tool makes it easy for us to not only profile and discover the rules, but also to operationalize them and auto-heal as the data evolves over time.”

– VP, Enterprise Information Management, Information Governance Leader, Insurance Company

“In the last 3 years we’ve had a 100x increase of APIs and microservices on the Cloud. This proliferation is beyond what Data Stewards can manage. As Cloud-native tool designed for Data Engineers, DataBuck autonomously validates data upstream and tremendously eases the burden on Stewards.”

– Senior VP Data Mgmt and Analytics, US Investment Bank

“Monitoring and validating files and data at ingestion directly impacts our revenues. DataBuck gives us the reliability, intelligence and speed we need to eliminate revenue-leakage.”

– VP Technology, Enterprise Data Office, Telehealth provider

“Aggregating weekly sales data from many dozens of sources and validating them is laborious and error prone. With DataBuck’s AI/ML-driven DQ automation we got more accurate data with less than 10% effort.”

– Director, Commercial Data Operations, US pharmaceutical

“With the traditional Data Quality tools, we could not thoroughly audit the financial data for the Street w/in our audit window. DataBuck’s performance has reduced data validation times from 11 hrs to 2 hrs, and w/higher accuracy.”

– Director, IT – Data Strategy, Financial Planning, Fortune-50 Hi Tech manufacturer



www.FirstEigen.com
contact@firsteigen.com

