MATILLION

Close the Information Gap

How to Succeed at Analytics in the Cloud

What does a **data-driven** business look like?

Today's businesses aspire to be "data-driven," but what does that really mean? In today's terms, a data-driven business is one that uses data across the organization to:

- Quickly iterate existing product lines to address new markets
- Optimize supply chains to meet dynamic geopolitical conditions
- Provide personalized experiences to consumers at an enterprise scale
- And more

Data-driven organizations move faster and do more with analytics. For example, Fidelity Investments implemented a unified data strategy across every business unit. As a result, the company's analytics productivity has skyrocketed. Internal data teams are reporting a 60 to 80 percent reduction in the data gathering effort to onboard new analytics use cases. Data scientists report that they now spend significantly more time on solving business problems and data modeling, rather than data gathering and cleansing.[1]

1 Fast-Track Data Monetization With Strategic Data Assets, MIT Sloan, July 2021

The key is a strong data and analytics culture, producing vital information that informs decision-making and behavior throughout the business. But a failure of analytics can open up Information Gaps that divide teams and silo data, leaving enterprises in the dark and struggling to catch up as their data-savvy competitors seize new opportunities and widen their lead in the market.

Mind the Information Gap

Data + Context = Information

Data alone is not the same as information.

For example at a bank, data includes a customer's name, the number of accounts that person holds, the amount of money they save or spend, and the transactions they conduct every month. Information is what that data combined can tell you: Whether that person is a loan risk, whether they're about to take their business to another bank, whether they're a good candidate for a credit card or a better rate.

Analytics is the act of turning data into useful and timely information that is circulating throughout your organization. If all of the parts of your information engine are humming–data, technology, people, processes–analyzing and modeling data results in useful and timely information circulating throughout your organization.

If any part of that engine breaks down, you might end up with an Information Gap. You have the data, and you have users waiting for insight. But there are barriers in the middle that prevent data from becoming the information that leads to insight, including:

- Siloed data
- Poorly prepared data
- Lack of communication and collaboration between teams
- Duplicated data, or too many sources of the truth

When analytics teams serve uncurated data, there is an Information Gap. When self-served reports of questionable origin proliferate across the enterprise, there is an Information Gap.

Information Gaps can also exist where there's no data present whatsoever. But in most cases, enterprises have plenty of data, all over the place. However, it's raw data that still needs to be refined into a valuable resource that's ready for analytics.

5 signs that you might have an Information Gap

1. There's a lag time between coming up with a data product and getting it into production. Forty percent of companies say it takes a month or more to deploy a machine learning model into production.[2]
2. Your data engineering team, your data scientists, and your business analysts are all using data from the same applications… from different points in time, in different datasets.
3. Your business bases decisions on a statistic or bit of insight. No one has any idea where it came from.
4. Information comes straight from the data engineering team into a dashboard, where it gets shared selectively by those who can see it.
5. The numbers in that dashboard have zero correlation with what end users are seeing on the front lines.

If any of these things sound familiar, you have some gaps to fill.

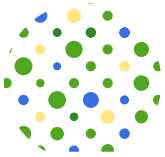
2 "The 2020 state of enterprise machine learning," Algorithmia, October 2019.

Dangers of an Information Gap

Information Gaps can ruin decisions and careers. They cause automation failure and revenue leaking inefficiencies. This lack of context and data together can have dire consequences for businesses trying to compete in digital economies.

Flying blind

If data is simply not available for use in decision making and prioritization, opinions rush in to fill the void. So many opinions, in fact, that organizations often thrash about. Strategies are often determined by those who make the most noise or wield the most power. An organization powered by loudly voiced opinion in a fast-moving digital marketplace is doomed to failure.



Data without context

This is perhaps the most tricky Information Gap of all. In this scenario, data can easily be molded into opinions masquerading as facts. When the discipline of transforming raw data into contextual information is performed incompletely, or the context comes from the inside of someone's brain, more or less raw data is spewed across the organization under the guise and backing of a formal analytics program. This creates an extremely dangerous scenario where inaccurate data can be bent to support the arguments of partisan perspectives without warning. This environment produces some of the world's worst business blunders.

Data Myths:
A Cautionary Tale

“Our business has an 80 percent customer retention rate”

That’s the statistic one business used as a proof point that informed financial decisions, marketing campaigns, business strategies, and no doubt numerous company pep talks.

It turns out, that statistic was 100 percent wrong. It was a data myth.

When the data team ran the numbers, it turns out the rebooking rate was actually only 40 percent–a huge difference. Foundations that were built on top of the data myth were suddenly extremely shaky. Unfortunately, by the time the figure was corrected, the information was already baked into company knowledge. People still see the 80 percent statistic quoted to this day.

Close your Information Gaps, or data myths may plague your business long after they are debunked with accurate analytics.

Unskilled and untrained self-service

As today’s world continues to digitize, data is more accessible than it has ever been. Ironically, this onslaught of data has reduced the capacity for analytics in most organizations precisely when it is needed the most. Traditional analytics programs–and many data teams--run on highly specialized data handling skills in the hands of a few. They’re not equipped to handle the “Three D’s” of modern cloud analytics. (see next page).

As data teams struggle under the burdens of running modern analytics programs, leading to longer delays and seemingly intractable team bandwidth issues, today’s digital savvy workforce is prone to bypassing the official analytics program altogether.

In this mode, digital-native workers simply download data directly from sources and try to stitch it together manually. The result is yet another form of data masquerading as information; an Information Gap. The charts get created and they seem conclusive. But deep underneath them lies a dataset which was not curated, cleansed and enriched by the skills of an experienced analytics professional. And worse, these datasets proliferate in silos, often propagating several half-baked versions of the truth.

The Three D's and Information Gaps

Big Data has long been characterized by the "Three V's": Volume, Velocity, and Variety of data. Those concepts still hold true, but more recently, the Three V's have given way to what IDC calls the "Three D's": Distribution, Diversity, and the Dynamic nature of data.

Distribution

Enterprise data teams used to handle a handful of large applications managed from private data centers. Not anymore. The number of data sources is exploding in the cloud, with data teams pulling from potentially thousands of data sources, integrating with an ever-increasing number of APIs.

The number of Martech solutions grew from 150 in 2011 to 8,000 in 2020, for a total growth of 5,233% over a period of 9 years.[3]

Hand coding requires 4-6 weeks of engineering effort to build a native connector, and 1 week per quarter to maintain it.[4]

Diversity

Today's data teams are required to process increasingly diverse types of data in their analytics workflows to provide next-generation insights. Audio recordings, unstructured documents, geospatial data, and IoT sensor logs are all common sources that help answer increasingly complex enterprise questions.

45% of organizations are using four or more distinct types of data in their analytics today.[5]

Dynamic Data

Teams that are used to working with static data sources are reeling against the dynamic nature of today's data, which goes far beyond simple schema changes. Constantly changing data forces them to constantly monitor and update pipelines. That's a maintenance burden that grows exponentially as they onboard new analytics projects and variable cloud data sources.

Cloud data platforms have risen to the task of handling modern analytics and the Three D's. Legacy data integration tools have not. Which leaves data teams scrambling to keep up and only widens Information Gaps.

3 chiefmartec.com
4 "How much does it cost to build an API?" Medium, April 2018

5 "Data Ops Survey," IDC, 2021, n=401

Where are the Information Gaps?

Information Gaps usually boil down to three challenges that organizations wrestle with when it comes to data: **implementation challenges, synchronization challenges, and data culture challenges**. If any one of these exist, there's likely a gap between the data that an enterprise has and the information that teams need to realize transformational business value.

The Implementation Challenge

What causes it?

The Three D's generate more data; but more critically, they generate more work for data teams. These teams need to get pipelines from experimentation to production faster. They need to quickly and easily reuse common architecture patterns or pieces of code so that each project builds on the foundation of the last. They need to take input, perform business logic reviews, and incorporate subject matter expertise from all areas of the organization. And they need to combine elements of data science and traditional analytics engineering to deliver mix-mode projects at speed and scale.

If the team attacking this challenge is small and relies on hand-coding to get the job done? The Information Gap widens. It also doesn't help that data engineering and data science teams often have different timelines and different agendas, work in different parts of the cloud, and deal with different data types (structured data vs. semi-structured and unstructured data).

The result

If it takes a week or more to even prepare data for analytics, and weeks or months to build the models that yield information, it's impossible for enterprises to quickly make insight-based decisions, create personalized experiences for customers, or rapidly deploy new products in the marketplace. Given the speed of business today and the expectations of customers, implementation challenges can render a company irrelevant in a short period of time.

The Synchronization Challenge

What causes it?

Traditionally, data teams prepared and analyzed data that most often ended up in business dashboards for a few stakeholders to interpret and act on. The days of the dashboard as the final resting place for information are over. Enriched data and insights need to go beyond the static dashboard and into the heart of the business, where they are accessible to all.

Data teams need to treat their data as a continuously recirculating asset and make it available to drive automation in the tools and processes that run their daily operations. IDC predicts that the most data-forward organizations will adopt "headless analytics" by 2025.[6]

But moving that information back into operational systems, data models, and new products can be complicated and time-consuming for data teams who lack the right tools.

6 IDC Futurescape 2021

The result

If information can only be seen by a few, it can't affect change where it needs to happen: on the front lines of the business. End users can't use it to score leads, decide on next best actions, surprise and delight customers, or plan content and campaigns that will move the needle for the enterprise.

The Data Culture Challenge

What causes it?

As the strategic value of data and analytics increases in an organization, so does the need to make it accessible to those who can transform it into useful information. However, legacy tool sets were not built with democratization in mind; they exist for specialized operators who endured hundred-hour training courses and years of hands-on experience to gain full mastery.

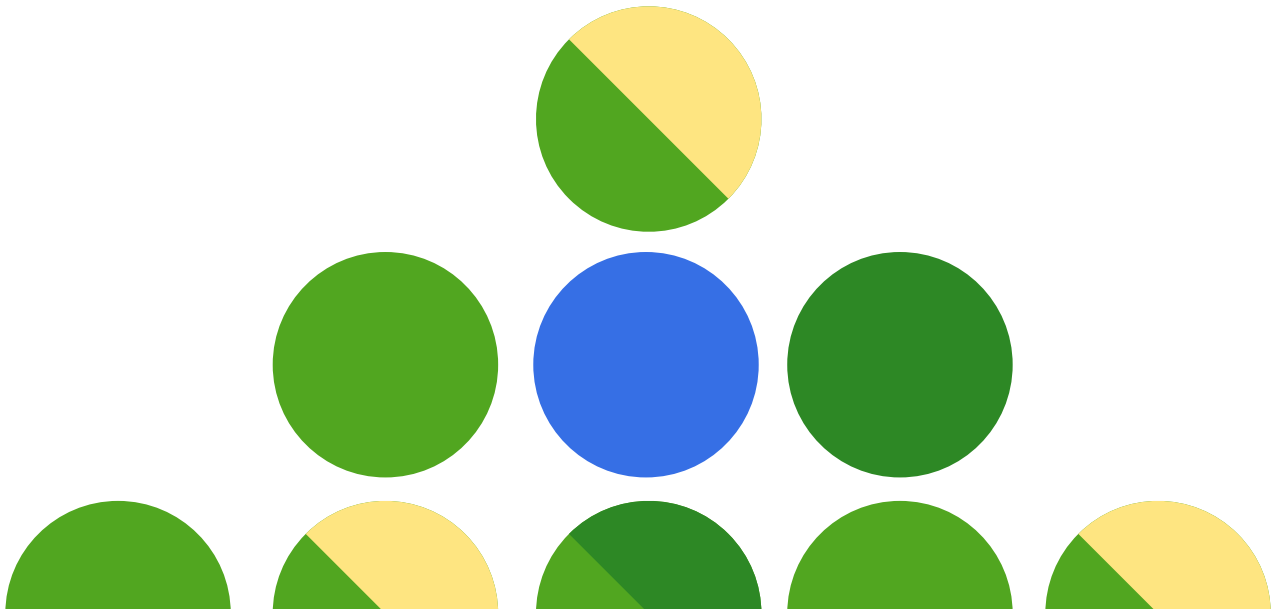
That won't cut it at an enterprise that needs to move faster, and where data-savvy workers that are a part of "Gen-D" want to access and analyze the data themselves. Data teams need to enable these users to work with data in a controlled and curated environment and champion statistical best practices.

The result

Unfortunately, data teams relying on legacy technology can never truly sustain the culture needed to create a truly data-driven organization. And as Gen-D workers continue to bypass the best practices of curated analytics, Information Gaps continue to flourish in unsanctioned reports across the business.

Closing the Information Gap

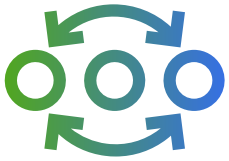
If your organization has an Information Gap, how do you close it? By enabling shared, secured, and connected data throughout the organization, facilitated by cloud data platforms, cloud-native tools, and processes that facilitate collaboration and data-driven operations.



Closing the Information Gap

Improving implementation

- Align your data teams and empower them to work together with shared data and shared tools.
- Decrease the learning curve and make it easier to onboard new data team members with highly visual, easily understood tools that provide a simplified way to work with complex logic. For example, low-code/no-code ELT software can streamline and automate many of the steps involved in integrating and transforming data and put valuable skills into the hands of many instead of a few.
- Choose tools that enable asynchronous collaboration with auto-documentation, real-time shared workflows, and support for Git and other platforms.



Synchronizing data across the business

- Think beyond the dashboard and send enriched data back out into operational systems where end users can make more informed decisions on the front lines.
- Invest in a cloud data platform such as a lakehouse or a data cloud that enables teams to work with data in both structured warehouses and unstructured environments like a data lake. A unified environment can help ensure that teams are working from the same enriched, prepared, analytics-ready data.
- Move away from data pipelines to data lifecycle management, with data flowing in and around the organization and constantly evolving.
- Choose data integration tools that re-integrate enriched data back into operational systems via two-way connectors.

Closing the Information Gap

Promote a true data culture

- Instill data best practices throughout the organization with everyone from technical teams to executives to data end users.
- Establish data governance that doesn't restrict data, but instead enables secure access for people who need it, as well as clear data lineage.
- Create shared environments where everyone on your team can work from their preferred interface and toolsets while all using the same datasets

Don't forget about security

- Accelerate implementation with a fast and secure infrastructure and automated governance
- Meet new-world regulations without adding workload to the team
- Understand the role of shared responsibility in cloud security:

  Know what security measures your organization needs, ensure your data is properly encrypted at every stage of migration, and have all necessary security controls in place.

MATILLION

Matillion ETL:
Helping organizations close the Information Gap

Matillion ETL can help data teams close Information Gaps and promote strong data culture throughout an organization.

- **Bring teams and data together in the cloud and make data useful for anyone who needs it with cloud-native ELT.**
- **Decrease the risk of data latency and bottlenecks with a low-code/no-code interface and workflow automation features.**
- **Support a multi-cloud strategy with products designed for native support on multiple cloud data platforms.**
- **Sync enriched, analytics-ready available data back into operational systems where it's accessible to users across the organization**

To learn more about how Matillion ETL can help bring data and data users closer together in your enterprise, talk with our team of experts today.

Request a demo