



Harness the power of data

Now is the time to become an analytics-driven organization.
Discover how.

Imagine this...

...You are reviewing the weekly spreadsheet that's just been delivered to your email. You spot an anomaly in the financial data that you don't understand—despite the pivot table provided in the report that allows you to drill down to at least some level of detail.

You ask your operations analyst what's going on. To which your analyst responds, "I'm not sure. Let me find out."

The next day, the analyst tells you that the reason for the anomaly is the fact that productivity was way down at the manufacturing plant.

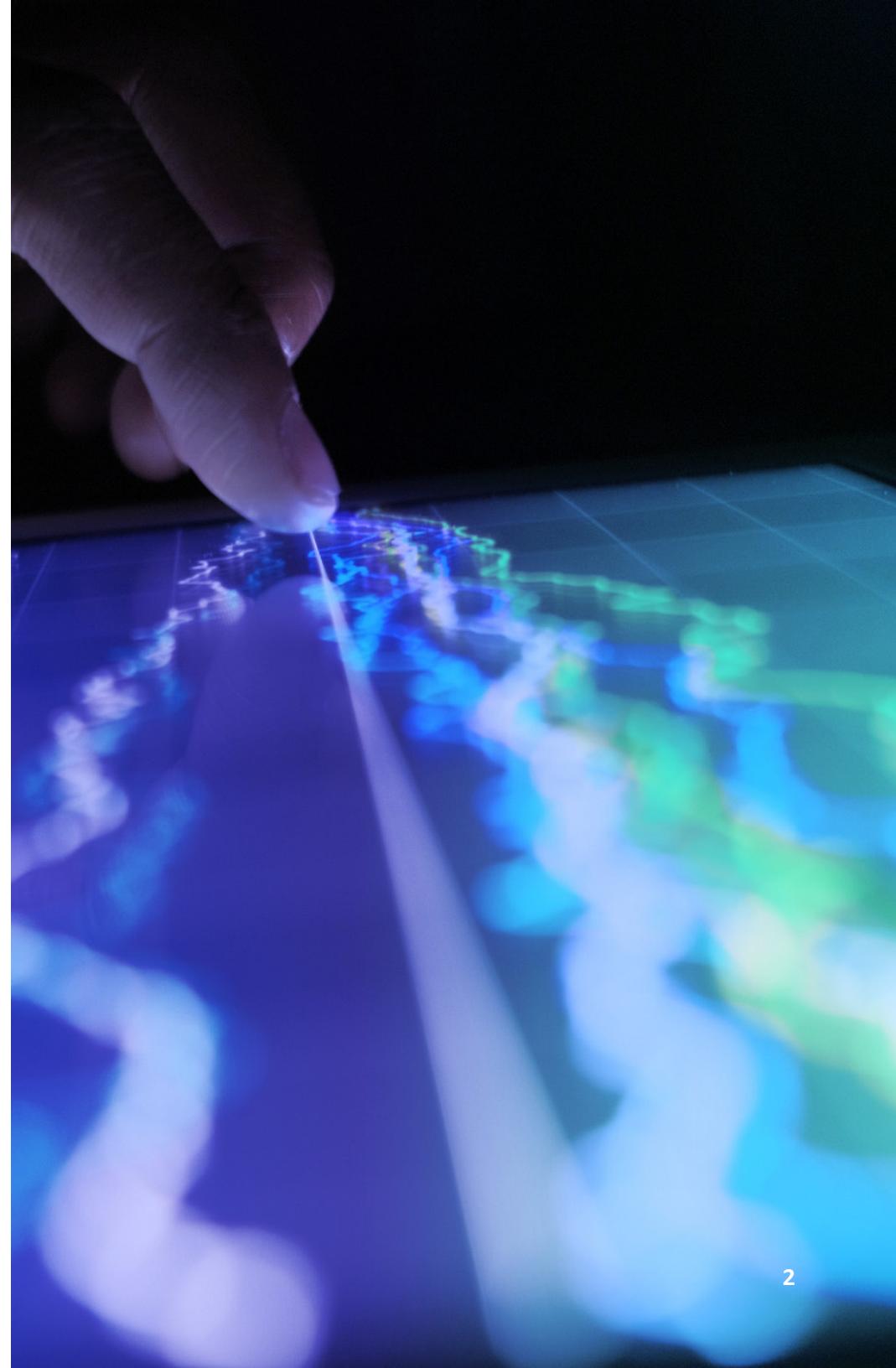
"That doesn't make sense," you say. "Can you ask Human Resources if sick days are impacting the productivity numbers? Or could it be that there was an issue with the time-capture application at the plant?"

"It will take a week to get that data and merge it with the financial data," your analyst says.

"Can't you just send me a dump of the data from the ERP and time application, and I'll work with it myself?"

The analyst responds,

"I don't have access to the data, and it will take a few days to submit the right tickets to get access to it."

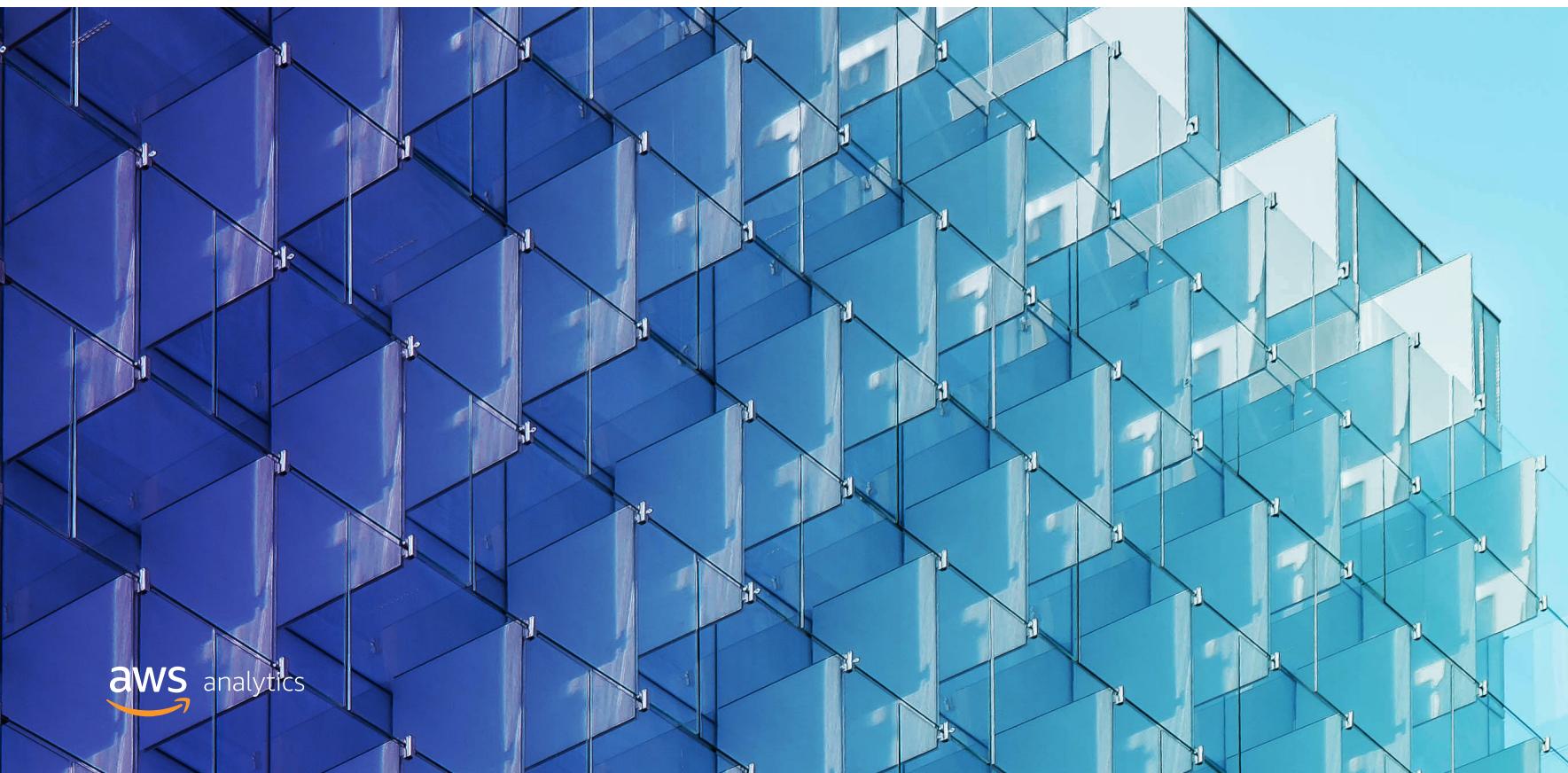


Sound familiar? Your organization is not alone.

The key fact is that every organization has a growing data problem. First reactions often lead companies to believe that these issues are related to business intelligence processes and tools. However, the real hurdle is the inability to simply assess the data that enterprises already have. With artificial intelligence and machine learning capabilities coming to fruition, now is the time to harness the power of information to become a data-driven organization.

How does data architecture behave?

Just as nerve endings extend throughout the human body, sending sensory signals to be processed by our brains, data architectures move information across modern networks. They have the ability to receive, process, and store data from anywhere, inside and outside the enterprise. The signals are processed in real time and acted upon by machine learning algorithms. These capabilities are applicable across the board—not just to specialized data scenarios.



Harnessing data means removing the dysfunctions

Big data problems are about more than just volume. Most organizations will encounter some or all of these common data dysfunctions.

Decentralized data

Siloed data exists across many applications, storage system databases, and data warehouses. Storing data requires its own access roles, rules, and processes.

How do you derive meaning from all of your data when you cannot access all of your data?

Diverse data

New data types are being created: Internet of Things sensors, images, videos, social media, text, and documents. Legacy systems have issues handling these new data types. Data can now be semi-structured or unstructured. Diverse data storage and analysis differs from that of traditional relational or transactional data.

How do you store and analyze new data types with analytics tools that are not optimized for new data types?

Dark data

You have data, but not all of it is loaded into your data warehouse. Data warehouse capacity is limited and unstructured data is difficult. Your data model does not have the right fields.

How do you get insights from all your data when all your data is not being analyzed?

Delayed data

Batch data loading is still pervasive. Nightly batch jobs involve extracting data from one system, transforming it, and loading it into another system. This transferred data can be analyzed or used for another purpose. The transfer of data creates delays for data availability.

How can you make a decision based on up-to-the-minute data if the data is not available until a later time?

Duplicate data

Data silos may store duplicate data. Different departments store and use customer data for different purposes. Each department may use unique customer data. Some data may be common across these departments.

How do you determine which system to get the duplicated data from?

Deserted data

Data may be stored for a new project. Data may be deserted when the project is cancelled, priorities shift, and people change jobs. This data can still have purpose, meaning, and be useful. However, people may not be aware of what the data is or its value.

How do you incorporate valuable deserted data into your analytics?

Deleted data

Interesting data may be deleted. Older data that may seem to provide little use today could be deleted. However, teams may need to analyze deleted data.

How can you take advantage of data if it has been intentionally or unintentionally deleted?

Dirty data

To confirm data is as clean as possible before being stored and analyzed, forms, rules, and other validations are necessary. Interesting data, such as unstructured or object-based data, may not be so clean.

How do you explore dirty data to understand what value it can provide?

Conclusion

If these points are familiar to you, it may be time to revisit your approach to data and analytics, and how your architecture enables it. You can now deploy fit-for-purpose analytics solutions (storage, processing, querying, analysis, presentation) to meet existing and future business and IT challenges.



Modern analytics platforms enable critical business insights

Once you're ready to tackle your big data problem, what can you reasonably expect to accomplish with a modern analytics platform?

Access to any and all data

Data-driven decisions necessitate access to all your data—areas, systems, applications, and data sources—by those who need it, when they want it.

Store it all in one place—an organization-wide data lake.

The advantages to data lake architectures:

- Store any data, in any format
- Durable, highly available
- Extends to petabyte and exabyte scale
- Secure, compliant, and auditable
- Handles regulatory and privacy requirements

With data lake architectures, you can run any type of analytics—from simple SQL queries to complex data warehouse, big data, machine learning, and predictive analytics. Data lake architectures decouple compute from storage so that you can scale either as needed. They can also implement storage tiering to optimize data availability and storage costs.



FORTNITE

Keeping sophisticated gamers engaged

Fortnite is an online game developed and run by Epic Games. In the world of *Fortnite*, players can cooperate on missions or attempt to be the last person standing in the game's Battle Royale mode. The game has become a phenomenon, with more than 250 million players around the world supporting over 10 million concurrent players during in-game events.

How revenue is achieved

Fortnite is free to play, with revenue coming entirely from in-game microtransactions—meaning its revenue depends on continuously capturing the attention of gamers through new content and innovation.

How Epic Games harnesses the power of data

Epic Games needs an [up-to-the-minute understanding of gamer satisfaction](#) in order to create an experience that keeps players engaged.

Epic Games collects billions of records on a daily basis, tracking virtually everything happening within the game: how players interact, how often they use certain weapons, and even the strategies they use to navigate the game universe. It's a massive amount of information, and it's all stored in their data lake for analysis. When the company introduces new elements and updates, it can use analytics to almost instantly learn how the *Fortnite* community is engaging, and then respond accordingly.

[Watch the presentation >](#)

Use purpose-built analytics tools on the same data

Why use an analytics platform? Simply stated, one size does not fit all data or use cases, and using separate tools to analyze disconnected data is complex. A modern analytics platform can handle diverse data types, wide-ranging use cases, and multiple analytics approaches on an organization-wide data lake.

What do modern analytics tools need?

They need to access and process any data from the same source—your data lake.

How does it access data?

Many analytics tools can access and process data directly from a data lake. Data in your data lake can also be queried from a modern data warehouse to take advantage of a structured and normalized data model with a high-performance query engine.

Traditional data warehousing vs. modern data warehousing

Traditional data warehousing and visualization tools may fit basic reporting, dashboards, and ad hoc queries on structured data that resides in data warehouses and data marts. Some use cases require access to data that is not in your data warehouse and analyzed in context with the normalized data in the data warehouse. A modern data warehouse can extend data warehouse queries to include data residing in the data warehouse, the data lake, and operational databases. This happens without any delay of data transformation or movement, so you always get those essential timely insights.

Can other analytics tools access the same data lake?

Other tools like big data analytics on unstructured data, real-time data stream analytics, and predictive analytics require purpose-built analytics tools that can access the same data in the data lake. This allows everyone across the organization, from business users to data scientists and everyone in-between, to have confidence in both the data and their analytics results.



Keeping customers constantly motivated

Equinox Fitness—a company with integrated luxury and lifestyle offerings centered on movement, nutrition, and regeneration—replaced its traditional on-premises data warehouse with a cloud data lake analytics platform.

The system collected data for their cloud data lake and data warehouse from:

- Exercise equipment (gamified cycling machines, cardio machines, etc.)
- Clickstreams
- Cycling logs
- Service-enhancing software

How Equinox Fitness builds on purpose-built analytics

Purpose-built analytics tools perform simple and complex queries against both structured and semi-structured data. Developers, business analysts, and data scientists throughout the company are empowered with the insights they need. Customer experience, engagement, and overall satisfaction have improved as a result.

[Watch the video >](#)



Real-time data availability for analytics

The speed at which decisions can be made is dictated by the availability of the data to be analyzed; and it needs to be extracted, transformed, and loaded before it can be analyzed. Therefore, we should expect modern analytics systems to ingest and analyze data in near real time and be responsive to changes from upstream data sources.

Key facts

- Ingesting data into your data lake or data warehouse in periodic batches, typically each day, delays time-to-insight.
- You should never have to wait for the daily batch to be completed to understand where your business stands.
- Streaming technologies allow data to be ingested at a massive scale, in real time, so you can perform analytics almost instantaneously.



Data and analytics on a massively entertaining scale

Netflix—the world's leading internet television network—uses data and analytics services from AWS to help keep its 100 million+ members reliably and happily streaming a collective 125 million hours of content each day.

Real-time response across TBs of daily data

Netflix must constantly monitor and optimize its network in order to continuously improve customer experiences, increase efficiency, and keep costs under control. By using [Amazon Kinesis Data Streams](#) to ingest, augment, and analyze multiple terabytes of daily data, Netflix can quickly identify and act on opportunities to maximize application uptime.

Creating a better streaming experience

Netflix's Amazon Kinesis Data Streams-based solution has proven to be highly powerful and scalable. It processes billions of traffic flows each day, generates updated analytics within seconds, and helps ensure high availability and great experiences for Netflix customers.

[Read the full story >](#)

Interactive insights where and how you want them

By considering the user experience when interacting with data, Business Intelligence (BI) tools have made things better. No longer do users have to suffer frustrating analytic experiences, delays, or complex searching for the data they need.

However, the number of tools provided to organizations requires users to be experts in each one. With infrequent use of tools, users experience frustration and costs can become exorbitant.

A modern analytics platform should empower users by providing interactive reporting, dashboards, as well as easy and intuitive visualizations. The platform should also have flexible licensing with pay-as-you-go terms and come with the freedom to explore the data without incurring large IT investments in infrastructure and licensing.





Keeping data visualizations in front of decision makers

The National Football League is a professional American football league consisting of 32 teams, divided equally between the National Football Conference and the American Football Conference.

Modern business intelligence (BI) tools allow the NFL to quickly build fast, interactive dashboards that seamlessly integrate with its Next Gen Stats applications. The NFL extends these secure, customized, and easy-to-use dashboards to each Club without having to provision or manage infrastructure, all while only paying for actual, pay-per-session usage.

[Read the NFL story >](#)

Intelligence embedded in the business

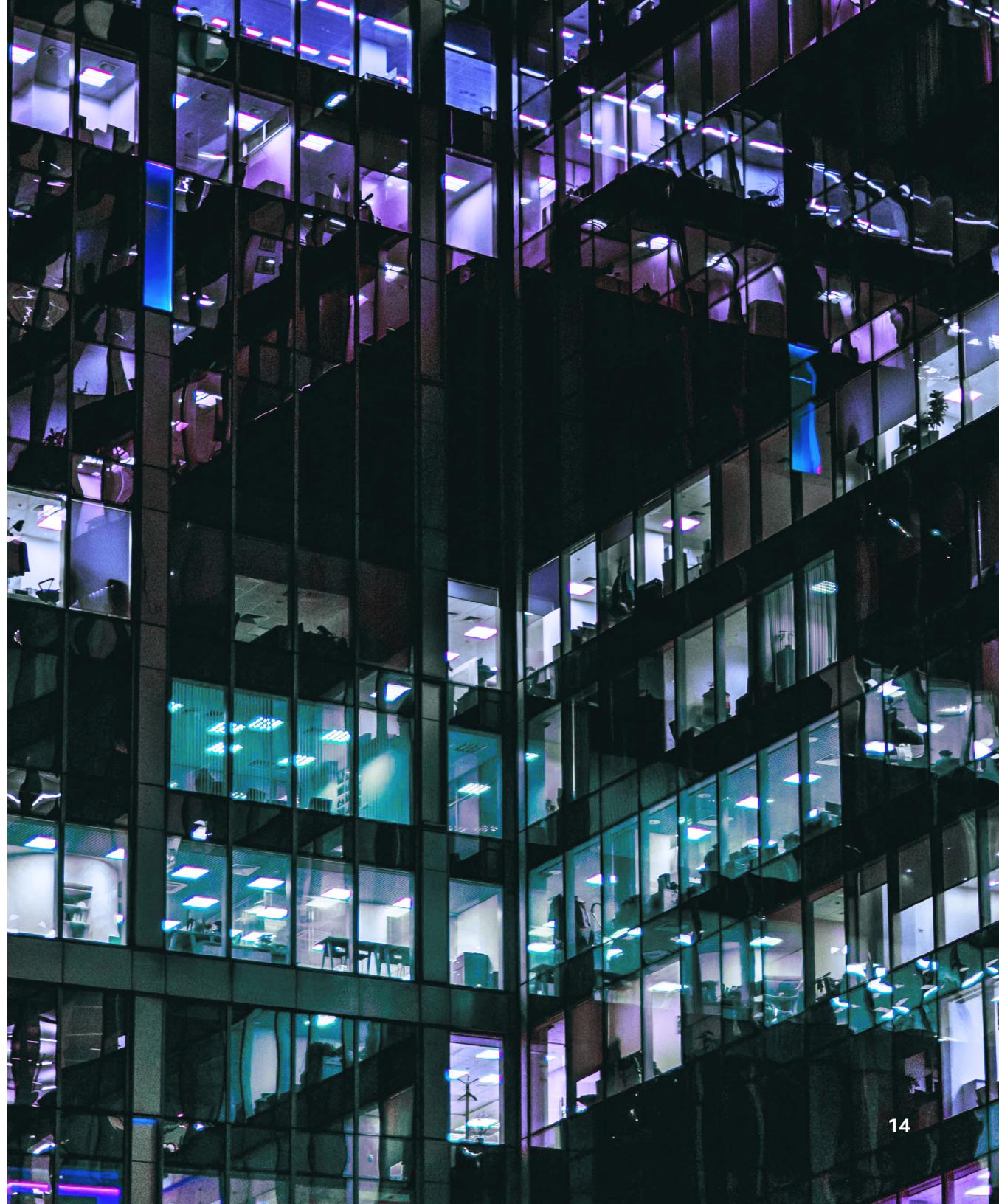
Artificial intelligence and machine learning are leading the way. Advances in machine learning frameworks, coupled with the use of specialized servers utilizing graphics processing units (GPUs), are enabling all kinds of new capabilities, like autonomous driving.

Where does data come in?

In order to train machine learning models, vast amounts of data are required. Using a data lake can make it easy to store data and make it available for AI and ML. Organizations are already starting to build these capabilities to drive new outcomes not previously possible, such as being able to better predict health outcomes based on retinal imaging.

How do you integrate algorithms with business processes?

Getting the insight or data science model created is the easy part—getting it integrated into your insurance policy engine or retail platform is the hard work. These systems do not typically have the ability to integrate outside data sources or APIs. This is a great opportunity to consider moving these systems to the cloud to take advantage of all the services available to help modernize or re-architect them.



AWS helps Zappos tie personalization and data together

Zappos began 20 years ago as a small online shoe retailer. Since then, it has grown to sell clothing, handbags, and accessories—while becoming renowned for its exceptional customer service. The company has been a subsidiary of Amazon since 2009.

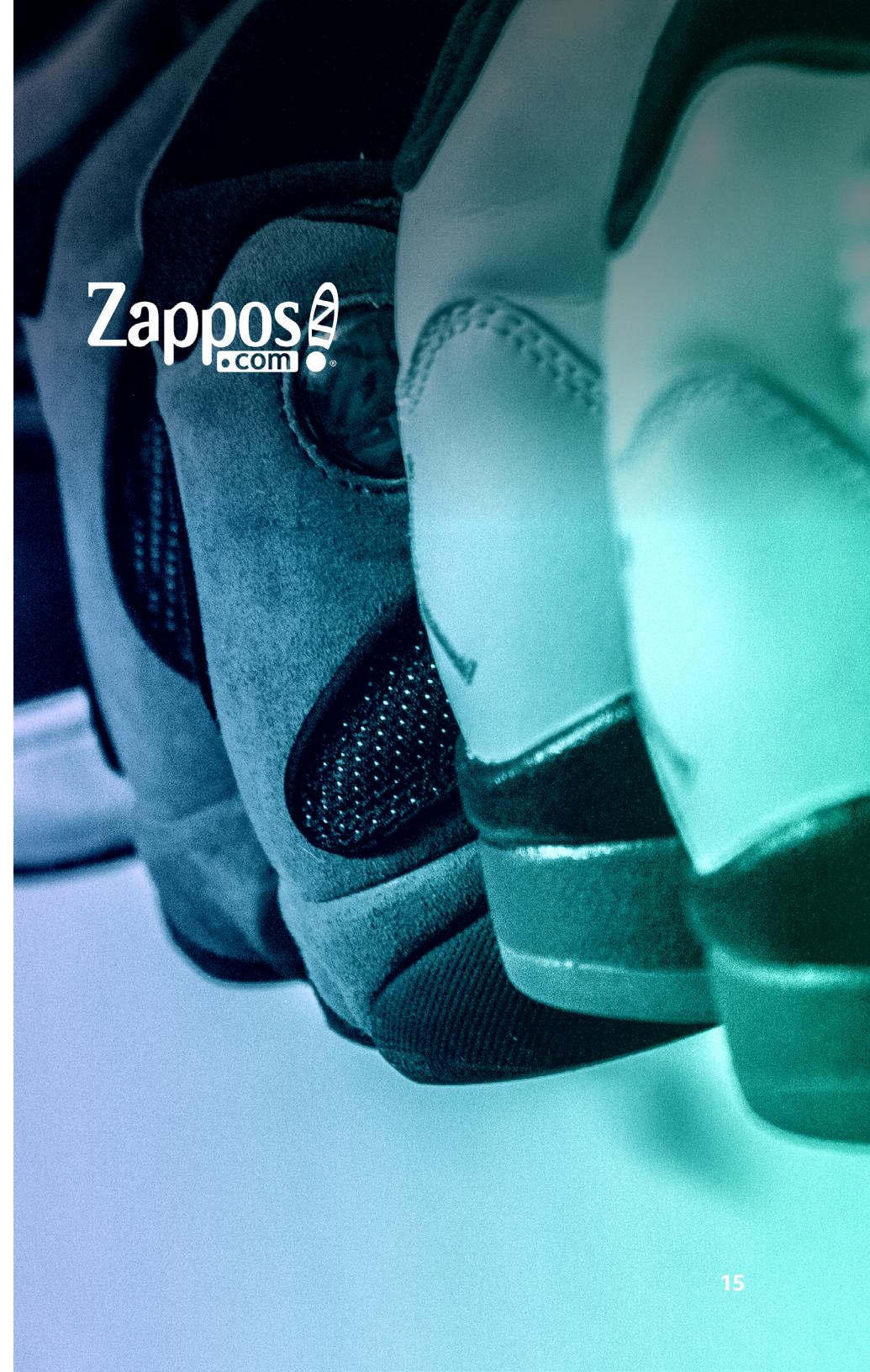
Searching for the perfect fit

Zappos provides highly fluid and responsive customer experiences by using analytics and machine learning built on AWS. These services enable Zappos to personalize sizing and search results and to offer highly relevant and enticing product recommendations based on individual user preferences.

Run, don't walk

Zappos delivers these personalized search results and recommendations with nearly undetectable latency—99% of searches complete in less than 48 milliseconds. These features have helped Zappos reduce repeated searches and product returns while increasing search-to-product-clickthrough rates.

[Read the story >](#)



Organizing for success

Technology is only one component when building an advanced analytics capability at your company. The organization—its processes, governance, and people—creates the biggest challenge. So how do you organize your analytics investment for success?

Start with an analytics Center of Excellence

Spearhead the change with a team and a leader as you move from strategy and intent to meaningful progress. Set up an analytics Center of Excellence (COE), starting small and growing to service more needs.

Large enterprises that already have established shared services organizations for business intelligence or reporting can seed the analytics COE with technical and business roles. Similar to IT infrastructure organizations, they should not only supply the talent, but also be a key driver and sponsor for the effort.

Over time, they will evolve to either adapt or become part of the analytics COE. The starter roles are often data engineers and business intelligence analysts. The leader will work across multiple organizations, business units, and back-office groups, such as finance and IT.

Serve all your customers' needs

It's all about making a mental shift from:

*"You must use our reporting solution
and you will like it."*



*"What are your analytics needs and
how can we help enable you?"*

When establishing a new analytics COE, set expectations for how the group acts and makes decisions. Position the shared services reporting organizations to answer the different questions that come from employees, business leads, and customers.

"It's not about having the latest tools in the belt; it's about making it easy for your customers to get what they need."

The analytics COE will need to service two types of customers

The data and analytics consumers

The decision makers, data scientists, business intelligence (BI) analysts, and developers who are concerned with the quality of the tools and services available to them to process and present data.

The data producers

The owners of applications, infrastructure, and devices who will supply the data into the platform. These customers need services to easily publish their data into the analytics platform and define a data contract.

This includes the domain model of the data, frequency of refresh, and definition of policies—for example, a security policy outlining who can access their data.

Those who manage the data and stand between data producers and data consumers often have a difficult mission. Data managers communicate across a potentially very large and diverse set of business units and personas—and if they do not meet expectations, then the analytics effort will not deliver the business value.

Rethink the COE

An analytics Center of Excellence presents a specialized set of cloud services focused on meeting analytics needs.

The past

Reporting and BI organizations often provided a one-size-fits-all strategy

The present

Making it easy for your customers (producers or consumers) to get what they need

Adapting to evolving technologies

This is the era of rapidly evolving technologies, including big data, rich visualizations, automated decision-making, AI, and machine learning.

In today's climate, it's just not possible to have a single technology stack. COEs run the risk of becoming concierge services. That can be fine for certain types of requests, but the COE can quickly get overwhelmed and backlogged with requests. There needs to be scalable, self-service mechanisms and transparent prioritization and governance processes in place.

Analytics COEs need to engineer and architect data platforms that are self-service, secure, operable, and scalable. The ecosystem needs to be ever evolving with technologies that can process, analyze, and present insights.

Sink or swim—What do I do today?

Put your data to work. Move it to the cloud.

Data is your strategic asset—have all your data, applications, and analytics workloads in the cloud. Start with one project and, upon its success, gain momentum to move on to the next project and the next.

Benefits

Speed, agility, and flexibility

Avoid

Planning, buying, installing, configuring, managing, and maintaining infrastructure

The steps forward

- 1.** Establish a data lake as your central, all-encompassing data repository in just a matter of days. Establish security, governance, and auditing policies in one place, applicable across the data lake to all users and all applications.
- 2.** Utilize crawlers to crawl through your existing data silos and create a centralized data catalog that can be used by multiple services.
- 3.** Connect your cloud data warehouse and other analytics tools to the data lake and data catalog.
- 4.** Empower business users, developers, and data scientists to access the same data catalog and data. Enable self-service to the data through multiple, fit-for-purpose analytics tools.
- 5.** Keep up the momentum and focus. Maintain open and frequent communication with your data and analytics consumers and your data producers.

Data harnessing starts by working modern

A modern analytics platform, with a data lake at its center, opens up a [world of opportunities](#) that can help your organization harness the power of data.

Imagine this journey...

...From the moment you begin to book your next trip, the airline, the hotel, the car rental, your data can make exclusive choices just for you.

Your favorite airline seat. The car you love to drive by the beach. The hit playlist, you know the one, that only you could find relaxing. Maybe it's the new summer read, for the beach only. That movie you didn't know you wanted to see, delivered to you on board. Or that cheese you've always wanted to try, in that perfect idyllic local restaurant, yes, it really does exist. With your data, discover the best adventures, perfect for you.

Kick-start your journey today

It all starts with capturing and storing data in a data lake and analyzing the data with a modern analytics platform. The more data you collect and keep, the greater the possibilities for using that data in new and innovative ways.

Plan out your analytics solutions, map out your data architecture to incorporate data lakes, and harness the power of your data.

Today is the day to become a data-driven organization.



Learn more about Data Lakes and Analytics today >

