

Six cloud data challenges solved for Database Administrators



Six Cloud Data Challenges tackles common data issues you're likely facing as a database administrator. Whether it's security, scale, performance, or something else, you're under pressure from management, developers, and customers to deliver seamless data-powered experiences, and this e-book provides actionable answers for the hurdles in your path.

Who should read this e-book?

This e-book is for database administrators facing potential challenges with managing data in the cloud. In the e-book, you'll find the technical details specific to each scenario. If you're working through these issues, you'll be familiar with the concepts unpacked here and, hopefully, find value in the solutions explored.

© 2019 Microsoft Corporation. All rights reserved.

This document is provided "as is." Information and views expressed in this document, including URL and other internet website references, may change without notice. You bear the risk of using it.

This document does not provide you with any legal rights to any intellectual property in any Microsoft product. You may copy and use this document for your internal, reference purposes.

Contents

Are you getting the most out of the cloud?	4
Scenario 1: Enterprise-proven services	7
95 percent of Fortune 500 companies trust the Microsoft Cloud	9
Customer case study: GeekWire	10
Scenario 2: Multitenant SaaS applications	12
Azure SQL Database: Redefining multitenancy	13
Azure Active Directory	15
Power BI Embedded	16
Customer case study: Umbraco	16
Scenario 3: Personalization around the globe	18
Azure Cosmos DB: A globally distributed, multi-model database service	19
Customer case study: Next Games	22
Scenario 4: Big data and advanced analytics	24
A comprehensive big-data and advanced analytics platform	26
Customer case study: PROS	30
Scenario 5: AI	32
The Microsoft AI platform: Bringing intelligence to your data	33
Customer case study: renewablesAI	35
Scenario 6: Security and compliance	36
Microsoft Azure: The trusted platform	37
Customer case study: Merck KGaA	39
Innovate faster with Azure data services	40

**Are you
getting the
most out of
the cloud?**

If you're like most database administrators, you're looking to learn as much about managing data in the cloud as you can.

Sometimes your research is proactive, based on personal interest or a desire to maintain your skills. More likely it's out of necessity—to help your organization get the most out of its data.

Not long ago, data was primarily hosted on-premises, forcing you to spend considerable time on essential requirements like scalability, availability, and security. While you still must deliver on those essentials, the databases you're building need to do a lot more: support millions of users, span the globe, make sense of petabytes of data, and delight users in new ways.

Cloud services not only deliver “built-in” scalability and availability, but also provide an unprecedented range of prebuilt functionality for driving innovation into your apps. That means you're ready to evaluate your next project, without your having to worry about physical infrastructure and the related overhead.

The biggest challenge? There's so much out there that it can be hard to know where to start.

As a database administrator, you're probably already engaged with the cloud at some level. You may be asking yourself, “Am I getting the most out of the data and AI technologies that are available in the cloud, or am I missing the key ingredients to deliver a meaningful experience?”

If you're not sure of the answer to this question, you're not alone—especially when it comes to data. Microsoft internal research shows that, even among Azure users, there's a very low awareness of key features and functionalities of Azure, plus some initial anxiety around what's required to run databases in the cloud.¹

¹ Study of US database administrators conducted by phone, commissioned by Microsoft, 30 participants, November 2018.

This e-book examines six of the most common data scenarios that you're likely to face and maps them to the services you might want to consider as a starting point:

1. **Keeping everything up and running**—how the cloud helps deliver on the essentials, such as scalability and availability, so that you can focus on other things, like performance optimization.
2. **Managing data for a multitenant software-as-a-service (SaaS) app** that can effectively deliver the same services to hundreds or thousands of customers, while ensuring that they can't see one another's data.
3. **Supporting an application with global reach**—for example, an app that delivers fast, real-time recommendations to users across multiple countries and regions—without the hassles of complex configurations that span multiple datacenters.

4. **Turning mountains of data into actionable insights** through a combination of big data and advanced analytics.
5. **Employing AI within your database engines** can greatly improve performance and security.
6. **Ensuring security and compliance** across all the data you manage in the cloud.

Like most technologists, you have a myriad of tools and services at your disposal and want to use what works best for you, without compromising how you work. With Microsoft Azure, it's easy to integrate data into your company's applications using your preferred data engine and the languages, frameworks, and tools of your choice. Read on to discover how to unlock the full potential of your data—wherever it lives, in whatever form it resides.

Scenario 1:

Enterprise-proven services

No matter what your company's products do, your IT teams need to ensure that they deliver on the "essentials."

Let's face it, every user demands a flawless experience with every interaction, including fast performance, high availability, ironclad security, and uncompromised privacy. But as a DBA, while you know you need to support these core capabilities, you probably don't want to think about them—let alone be burdened with all the details necessary to ensure their delivery.

To determine whether you can trust a cloud platform to deliver on the basics—in a way that lets you stay flexible and focus on delivering new experiences that delight users—it's worth asking yourself a few questions:

How can the platform help me avoid negative user experiences like downtime? What does it offer that helps me avoid calls in the middle of the night or data maintenance that requires taking an app or service offline every weekend?

How can the platform help me handle unexpected spikes in workload and avoid complaints about poor performance—without having to pay for a lot of spare capacity that I might rarely need?

Can the platform help me expand to other geographic regions? And when it's time to do so, will I still have access to the same built-in availability, scalability, and security mechanisms that I've come to depend on?

How can the platform help me keep customer data secure? If there are regulatory or compliance requirements, does the platform have the necessary certifications?

More than any other factor, customer experiences determine whether companies thrive and profit, or struggle and fade.

"Outside In: The Power of Putting Customers at the Center of Your Business" by Harley Manning and Kerry Bodine (New Harvest, 2012).

With your short-term investment in choosing the right cloud platform for your data, you'll soon be reaping the long-term rewards of stability, efficiency, and, ultimately, happy customers.

95 percent of Fortune 500 companies trust the Microsoft Cloud

[Microsoft Azure](#) services are built cloud-first—ready to deliver all of the above, no matter where you are in the world. Azure runs on a worldwide network of [Microsoft-managed datacenters](#), across [54 \(and growing\) regions](#), ensuring global coverage.

With Azure, you don't need to worry about capacity or performance. You can easily and instantly scale up when more capacity is needed, and then scale down to save money when demand subsides. You can configure Azure to let you know when capacity or performance need attention, or you can configure your Azure storage accounts to scale up and down on their own.

Azure managed services help ensure availability through automatic patching and backups, built-in monitoring and security, and more, so you can stay focused on building resilient databases. Many Azure services are backed by [service-level agreements \(SLAs\)](#) that include uptime guarantees and downtime credit policies. To ensure business continuity and disaster

recovery, you can choose from locally redundant storage (where data is replicated locally within your primary region) or geographically redundant storage, where data is replicated to a secondary region that's at least 250 miles (400km) away but within the same geography.

What's more, because Azure is available in 140 countries/regions, you can put your data where its users are. Storage of data can be restricted to a single geography, region, or country, so you always have control over how close your data is to the people who use it.

Azure also provides comprehensive security and identity management tools and services, including advanced threat detection, [Azure Security Center](#), [Azure Active Directory](#), [Azure Key Vault](#), and [Azure Multi-Factor Authentication](#). And Azure offers the most comprehensive set of [compliance offerings](#) (including certifications and attestations) of any cloud service provider. (Security is such an important topic that we cover it separately later, in [Scenario 6: Security and compliance](#).)

Whatever you're building, Azure helps ensure that your customers get enterprise-proven levels of scalability, availability, and security—everything your customers expect. And while Azure takes care of these fundamentals, you get the space to focus on optimizing database performance.



Figure 1. Microsoft Azure runs on a worldwide network of Microsoft-managed datacenters, across 54 (and growing) regions.

Customer case study: GeekWire

Based in Seattle, Washington, GeekWire is a rapidly growing technology news site with a global readership. Started by veteran tech journalists in 2011, GeekWire quickly established a reputation for fast-breaking industry news and expert analysis.

In the past, GeekWire ran on open-source, WordPress software with a MySQL database, and the site was hosted by a managed services provider dedicated to WordPress. Although that relationship worked well in the beginning, as GeekWire's

popularity and site traffic increased, so did its performance concerns. "We rely heavily on people visiting us from social media and search engines, where performance is a ranking factor, so we need to deliver content very quickly," says Kevin Lisota, Web Developer at GeekWire.

Eventually, the time came for a better solution. Lisota and his team needed:

- **Increased control**, including visibility into the underlying infrastructure and the ability to diagnose and fix issues themselves.
- **Instant scalability**, so they could easily keep up with rapid growth, without any downtime or conversations with salespeople about “next tier” pricing.

When Lisota learned that Microsoft was launching a fully managed service for MySQL databases, moving to Azure made sense. “Azure Database for MySQL was what we were looking for,” he says. “I don’t have to deal with managing any of the details of the database—I don’t have to patch, scale, or back it up, and can rely on Microsoft to take care of that part of our infrastructure.”

GeekWire’s move to Azure delivered several immediate benefits, including on-demand scalability and reduced administrative overhead. “The biggest benefit of Azure Database for MySQL will be to have Microsoft manage and back up that resource for us so that we can focus on other aspects of the site,” Lisota says. “Plus, we will be able to scale up temporarily as traffic surges and then bring it back down when it is not needed. That’s a big deal for us.”

Our page load times are very low, and we’re able to do it on a more powerful and scalable infrastructure that costs us 45 percent less.

Kevin Lisota

Web Developer, GeekWire



Read the entire [GeekWire case study](#).



Hear from [other customers who are using Azure](#).

For more information

- Get credit to [try data services](#) with an Azure free account.
- Find out about [migrating your data to Microsoft Azure](#).
- Determine [which Azure region is right for you](#).
- Learn why [you can trust Microsoft Azure with your data](#).

Scenario 2:

Multitenant SaaS applications

Software-as-a-service (SaaS) apps are becoming more and more popular across organizations of all sizes.

This is happening for many reasons, including faster time-to-value, reduced up-front investment, and freedom from worrying about running and supporting apps.

To effectively support a SaaS app that serves hundreds or thousands of customers, delivers essential elements such as data isolation and security, and helps protect margins, you may want to start by asking these key questions:

How can I keep performance high as overall usage grows? How will I handle varied and unpredictable spikes in demand?

How can I manage customer data at scale, so that I won't need to onboard new staff linearly as my business grows?

How can I keep customer data secure— and isolate it to ensure that customers can't access one another's information? And how can I help my customers ensure secure access across their own user bases?

As the data in my company's app grows, **how can I uncover relationships in the data and gain new insights from it**—in areas ranging from app performance to customer behavior?

How can I help my customers visualize their data to gain new insights more quickly and easily— and thus add more value to my app?

If these questions resonate, then you probably already have a firm grasp of what you need from a cloud database solution.

Azure SQL Database: Redefining multitenancy

[Azure SQL Database](#), an intelligent, relational cloud database service, helps you avoid the typical tradeoffs between development efficiency, manageability, performance, and security for SaaS applications. As your app runs, Azure SQL Database employs built-in machine learning to continuously assess your app's behavior, tune performance, and automatically improve reliability and data protection—freeing you to focus on other things. And as demand for your SaaS app grows, Azure SQL Database scales on the fly, with virtually no app downtime.

With Azure SQL Database, you can take advantage of:

- **Elastic database pools**, which enable cost-effective and predictable pooling of resources to help you handle varied and unpredictable spikes in demand.
- **Elastic database jobs**, which let you streamline admin tasks by making changes to many databases at once—ultimately enabling you to manage thousands of databases as one.
- **In-memory online transaction processing (OLTP)**, which keeps active data in main memory to improve performance by a factor of up to 30—giving you more performance for your money.
- **Columnstore**, which stores table data as columns instead of rows to deliver 10 times the data compression and queries that are up to 10 times faster.
- Numerous ways to help keep customer data secure—including **row-level security**, **dynamic data masking**, and the **Always Encrypted** feature.
- Built-in intelligence to further boost app performance and data security—including features such as **Threat Detection**, **Query Performance Insight**, and **Azure SQL Database Advisor**.

- **Graph processing**, which can make it easier to express certain kinds of queries—and potentially improve query performance—by enabling you to model many-to-many relationships among relational data.
- **Managed Instance**, the best PaaS option for migrating your SQL Server database to the cloud.

Azure SQL Database supports multiple [data models for multitenant SaaS apps](#), so you can choose the one that delivers the optimal balance of isolation, cloud resource costs (compute and storage), and DevOps complexity for your unique situation. This also allows for a more flexible business model that includes a trial/free tier experience, serves a breadth of moderately active customers, and offers high-end, dedicated premium offerings for your most demanding customers.

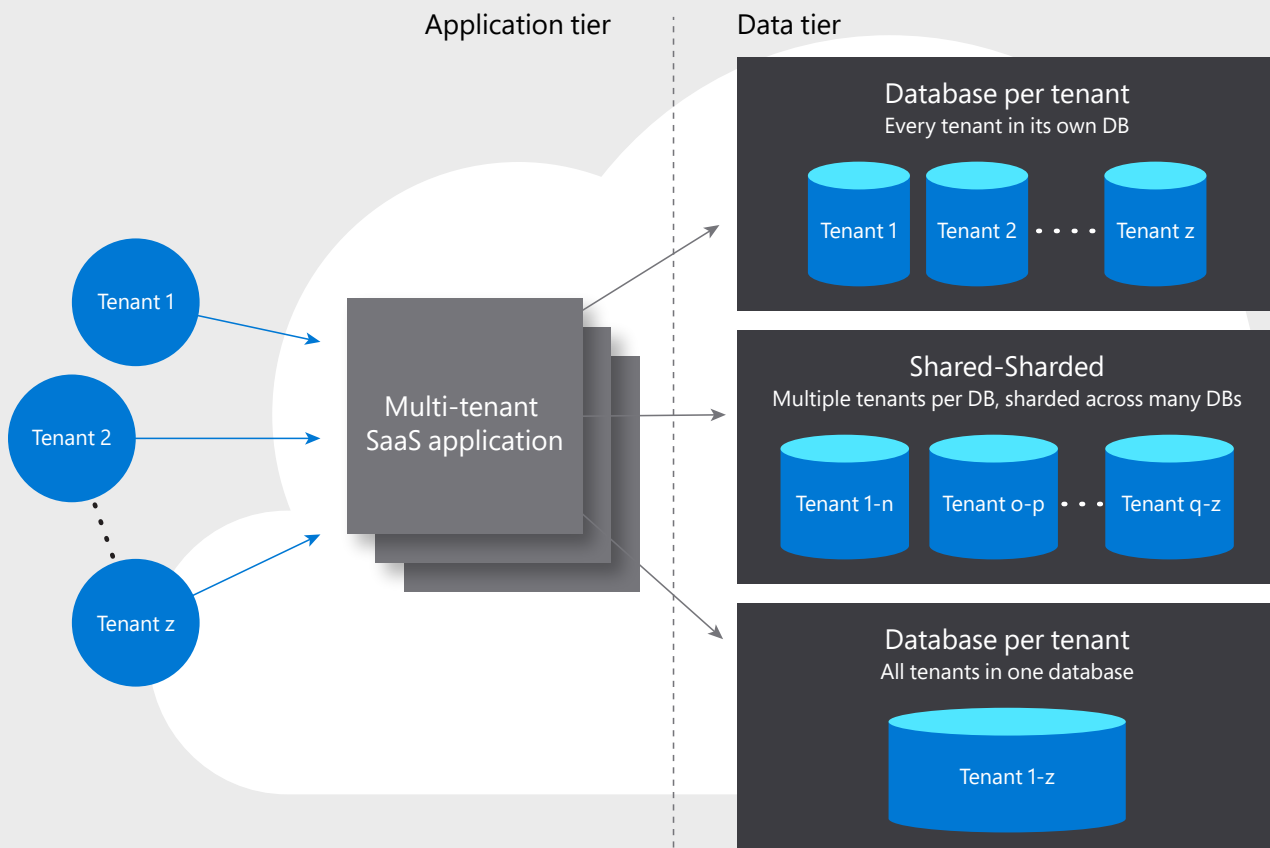


Figure 2. Azure SQL Database supports multiple data models for multitenant SaaS apps.

A list of all major Azure SQL Database features (with links to more information) can be found on the [Azure SQL Database features page](#).

Azure Active Directory

With a multitenant app, you need a way to organize and manage user accounts. More likely than not, your customers already have an identity management and authentication mechanism—and don't want to use and maintain a different one just to access your app.

You can use [Azure Active Directory](#) (Azure AD)—a multitenant, cloud-based directory and identity management service—to easily add single sign-on (SSO) to your SaaS app for greater value and differentiation in the marketplace. With Azure AD, you make it quick and simple for your customers to integrate with an existing Windows Server Active Directory, enabling them to leverage existing on-premises investments to manage access to your app.

Power BI Embedded

Few SaaS apps are complete without some sort of reporting environment. But all too often, such reports are static and deliver only limited insights. Embedding a rich, interactive analytics environment into your app is another way to increase its usefulness and help your customers make quick, data-driven decisions.

[Power BI Embedded](#) is an Azure service you can use to bring data to life within your SaaS app—without the time and expense of building and maintaining your own analytics environment. You can easily embed stunning, fully interactive reports in a cost-effective and scalable way—providing yet another source of value and differentiation, and potentially an additional revenue stream if you choose to position such functionality as a premium offering. You don't need to change the design of your app to use Power BI Embedded, and your users can sign in to your application just like before.

Customer case study: Umbraco

Umbraco is a popular open-source content management system. To deliver new customer value, Umbraco used Azure SQL Database and other Microsoft Azure services to build Umbraco-as-a-Service

(UaaS), a multitenant SaaS app that lets customers self-provision new environments.

To make new customer environments available virtually immediately, Umbraco maintains a pool of pre-provisioned databases using elastic database pools. As the pre-provisioned databases are used, new elastic database pools are created and new databases are pre-provisioned. An automated process restores databases that are no longer needed to the elastic database pool. As of October 2016, Umbraco had nearly 3,000 databases across 19 elastic database pools.

"We wanted to spend our time on solving our customers' problems, not managing infrastructure," says Niels Hartvig, founder of Umbraco. "We initially considered hosting the servers ourselves, but capacity planning would have been a nightmare."

By choosing Azure, Umbraco can provide customers with optimal performance, without the investment in IT resources that a self-hosted solution would require. "We love the ... convenience and scalability that Azure gives us, and our customers are thrilled with the features and reliability," says Morten Christensen, Technical Lead at Umbraco.

Elastic database pools are a perfect fit for our SaaS offering because we can dial capacity up and down as needed. Provisioning is easy, and with our setup, we can keep utilization at a maximum.

Morten Christensen

Technical Lead, Umbraco



Read the entire [Umbraco case study](#).



Hear from [other customers who are using Azure SQL Database](#).

For more information

- Get credit to [try SQL Database](#) with an Azure free account.
- Explore the [elastic database features of Azure SQL Database](#).
- Review the [advanced security and compliance features in Azure SQL Database](#).
- Learn [how to secure your Azure SQL Database](#).

Scenario 3:

Personalization around the globe

In today's global economy, apps and their users span the planet. To support rich, real-time personalization at global scale, your company's app needs fast access to the data that drives this personalization, regardless of where the user is located.

To achieve this, from a data perspective, you need to consider how to:

- Manage and version complex schemas.
- Scale both throughput and storage based on global demand.
- Balance the tradeoffs between the limited consistency models in a distributed database— including latency, availability, throughput, and scalability. (Most databases only offer two models, consistent and eventual, forcing you to choose from one end of the spectrum or the other.)
- Deliver customized, real-time, and highly responsive experiences.
- Ensure an always-on system around the globe.

In addition, to deliver the greatest value, you need to handle a variety of unstructured data at scale. Social media posts, mobile data, document text, application telemetry, and website content are just a few examples of today's data that doesn't always fit neatly into rows and columns. You also need to collect and analyze the data needed to drive real-time personalization before you can operationalize it for real-time access.

Ideally, you need a way to meet all these requirements without the hassle of deploying and managing servers, complex configurations of multiple datacenters across several regions, and complex big data, advanced analytics, and machine learning technology to drive personalization.

Azure Cosmos DB: A globally distributed, multi-model database service

[Azure Cosmos DB](#) is designed to provide low latency access to rich data anywhere in the world— making real-time personalization at global scale exactly the type of scenario that Azure Cosmos DB is designed to support. What's more, with Azure Cosmos DB, you can model your data the way your app requires, using familiar APIs, tools, and frameworks.

Azure Cosmos DB provides:

Turnkey global distribution—Azure Cosmos DB is currently the only fully managed database service to offer [turnkey global distribution](#), letting you put your data where your users are for fast, responsive access. Every database account can be associated with any number of Azure regions, and the data replicates automatically, synchronously, and durably.

Support for multiple modes and APIs—

Only Azure Cosmos DB empowers you with key-value, graph, and document data in one service. Azure Cosmos DB automatically indexes all data, so you don't need to worry about schema or index management. You also have your choice of APIs—including SQL, JavaScript, Gremlin, MongoDB, and Azure Table storage.

Limitless elastic scale around the

globe—With Azure Cosmos DB, you can independently and elastically scale storage and throughput anytime, anywhere across the globe, paying only for the throughput and storage you need.

Multiple, well-defined consistency

choices—Azure Cosmos DB offers an intuitive programming model and is currently the only non-relational database service to offer five well-defined consistency levels—enabling

you to build for the unique needs of your app and best navigate the tradeoffs between consistency, latency, availability, throughput, and scalability.

Guaranteed low latency—With its latch-free and write-optimized database engine, Azure Cosmos DB guarantees less than 10-ms latencies on reads and less than 15-ms latencies on (indexed) writes at the 99th percentile.

Industry-leading, enterprise-grade

SLAs—Azure Cosmos DB is currently the only cloud database service to offer industry-leading SLAs for 99.99 percent availability, latency at the 99th percentile, and guaranteed throughput and consistency.

The capabilities provided by Azure Cosmos DB make it well suited for [many use cases](#) beyond real-time personalization at global scale—including IoT and telematics, retail and marketing, gaming, web and mobile applications, banking, and other mission-critical, global scenarios.

Azure Functions:

Enabling a serverless architecture

By taking advantage of the native integration between Azure Cosmos DB and [Azure Functions](#), you can quickly and easily build and deploy event-driven, serverless, personalized apps that have low-latency

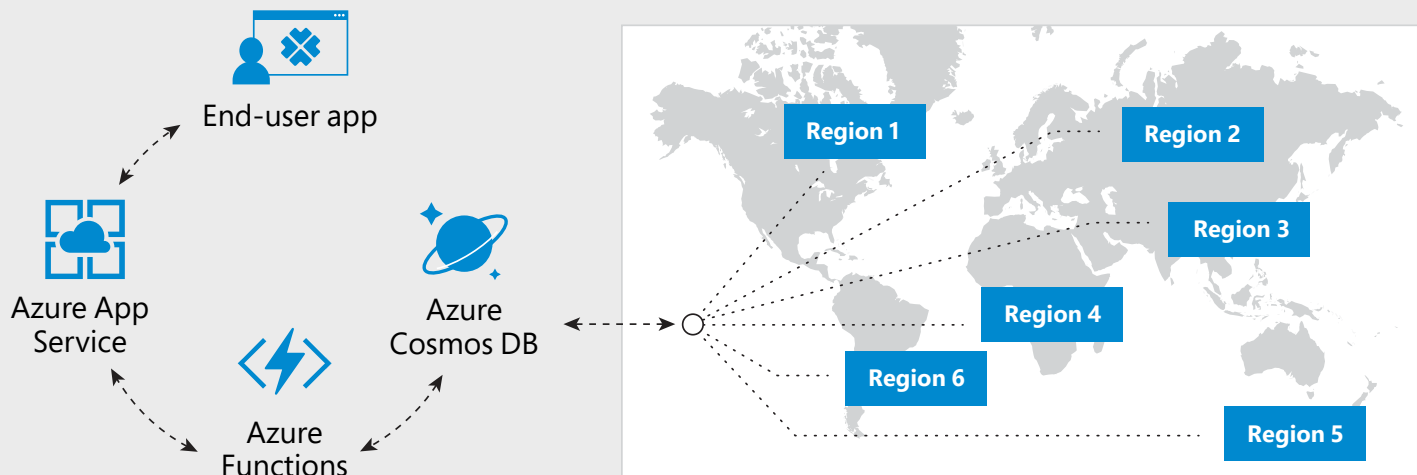


Figure 3. Native integration between Azure Cosmos DB and Azure Functions makes it easy to build fast, serverless, personalized apps at global scale.

access to rich data on a global scale. Because Azure Functions are event-driven, you can simply listen to a change feed from Azure Cosmos DB instead of creating your own listening logic. You can also bind an Azure Function to an Azure Cosmos DB collection using an input binding (which reads data from a container when a function executes) or using an output binding (which writes data to a container when a function completes).

With Azure Functions, there are no limits—you set the parameters and the functions execute in parallel, with the Azure Functions service spinning up as many as times you need. The service creates new instances of functions whenever an event fires

and closes them as soon as the function completes. This makes functions good for quick tasks and enables you to only pay for the time your functions are running.

Used together, Azure Cosmos DB and Azure Functions deliver a compelling set of benefits: event-driven, serverless computing at near-infinite global scale, with low-latency access to rich data for serverless apps, app performance that enables a real-time user experience, and freedom from infrastructure through fully managed services.

Obtaining and analyzing the data to drive real-time personalization

You must collect and analyze the data needed to drive real-time personalization before you can operationalize it and push it into Azure Cosmos DB for real-time access. The Azure big data and advanced analytics services you can use to do this—such as [Azure HDInsight](#) for analyzing data in Spark or Hadoop environments and [Microsoft Power BI](#) to bring that analysis to life—are covered later, in [Scenario 4: Big data and advanced analytics](#).

Customer case study: Next Games

Next Games, maker of the popular mobile strategy game The Walking Dead: No Man's Land, has been with Azure from the start. By 2017, No Man's Land had 16 million installations, had generated 120 GB of new data each day, and was handling 11,500 database requests per second. To support continued growth, Next Games needed more storage, increased flexibility, and the ability to maintain availability and performance for users who aren't geographically close to one another.

To meet those needs, the company adopted Azure Cosmos DB. "We want our game data to always reside in the datacenter closest to the player so that we can load the player data as fast as possible," says Kalle Hiitola, Chief Technology Officer at Next Games.

"[The Azure Cosmos DB] replication feature allows us to do this, so that we always have the data close by the player, no matter which datacenter it comes from."

Currently, all player data is in Azure Cosmos DB except for the actual saved player file, which is stored in Azure Blob storage. In addition to Azure Cosmos DB and, soon, [Service Fabric](#), Next Games is using Media Services for streaming video, Event Hubs for analytics, Notification Hubs to send push notifications to clients, Traffic Manager for load balancing, and Web Apps for the dashboard, which provides game management.

The good thing about Azure is that we can incrementally add new features Microsoft brings on to Azure into our platform and switch them behind the scenes. **You just start a new service, and off you go exploring what it can do for you.**

Kalle Hiitola

Chief Technology Officer, Next Games



Read the entire [Next Games case study](#).



Hear from [other customers who are using Azure Cosmos DB](#).

For more information

- Get credit to [try Azure Cosmos DB](#) with an Azure free account.
- Read the [technical introduction to Azure Cosmos DB](#).
- Watch the [Introducing Azure Cosmos DB](#) video.
- Learn how to [partition and scale in Azure Cosmos DB](#).
- Check out the sample [Azure CLI scripts](#) and [Azure PowerShell scripts](#) for Azure Cosmos DB.
- [Try Azure Cosmos DB for free](#) without an Azure subscription.

Scenario 4:

Big data and advanced analytics

Customers today have high expectations when it comes to speed, consistency, personalization, and more.

While failing to meet those expectations might cause customers to leave, simply delivering what they already expect might not be enough to delight them. So how can you exceed their expectations?

The answer lies in your data—or, more accurately, what you do with it. By capturing more of the data within your enterprise, augmenting it with additional data from your customers and other sources, pulling it all together (both historical and fresh), and examining it in new ways, you can extract the insights you need to deliver compelling digital experiences for every type of customer, across virtually every industry and scenario.

To get the most out of your data, you need a single, scale-out infrastructure on which you can capture and analyze a variety of data at unprecedented volumes and velocity. Specifically, you need to:

Build a highly scalable “big data warehouse” capable of ingesting and processing both structured and unstructured data. Your organization

likely has most of the structured data within your walls, such as information on customers, products, and sales. However, you can gain new insights by augmenting that structured data with unstructured data from other, often external sources—such as weather predictions, social media content, media files, streaming IoT data, or application telemetry.

Apply sophisticated machine learning and advanced analytics to find the hidden patterns in your data and derive new insights. Then operationalize those insights via machine-learning models and scoring APIs, and push them back into a globally distributed database like Azure Cosmos DB. Ultimately, the data within your traditional stores is enriched with newly derived insights for real-time personalization and deeper customer engagement.

Leverage the same infrastructure that supports the above two (batch-oriented) processes to implement real-time processing—in other words, a Lambda architecture that supports both batch (cold) and stream (hot) processing methods. With this, your possibilities are endless. Detect fraud in real time. Use IoT telematics to predict which machines on an assembly line might be about to fail. Monitor real-time physiological telemetry to identify life-threatening conditions. You get the idea.

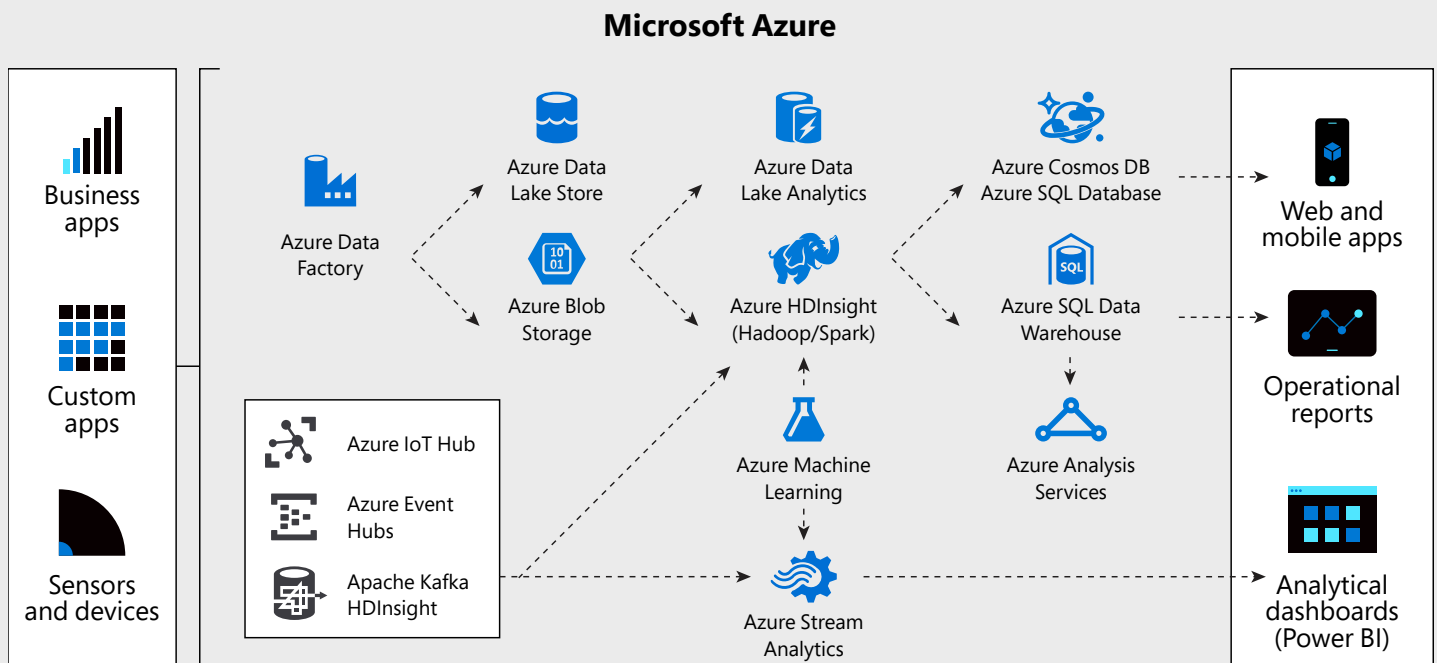


Figure 4. Microsoft Azure supports a single, scale-out architecture for capturing and analyzing a variety of data.

A comprehensive big-data and advanced analytics platform

With the big data and advanced analytics services in Azure, you have a tool set that enables you to do all the above. What's more, you can provision fast and scale even faster—paying for only what you use. You also benefit from granular security, automatic threat detection, market-leading SLAs, broad compliance certifications, a datacenter network with unprecedented worldwide coverage, and your choice of tools across both Microsoft and open-source services.

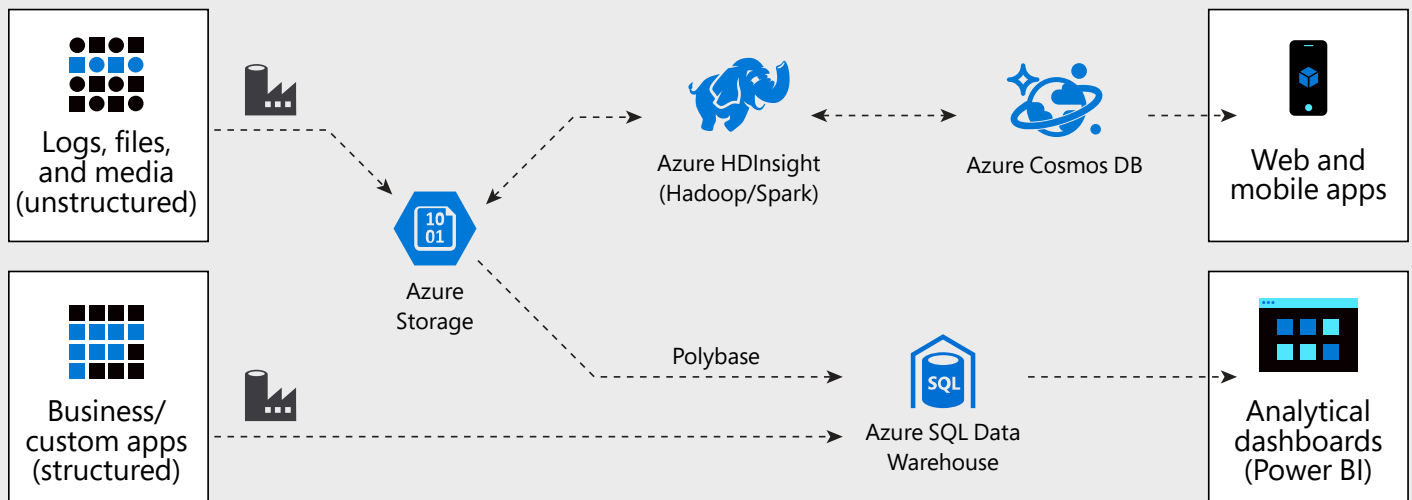


Figure 5. Big data and data warehouse.

Big-data and data warehouse

You can use the following services to store and manage all types of data at hyper-scale:

- For structured data, use [Azure SQL Data Warehouse](#) for elastic scalability with massively parallel processing.
- For unstructured data, use [Azure Data Lake Store](#) to get massive throughput and analytic performance, with no file-size limits—or use [Azure Blob storage](#) if you're cost-conscious and your needs are more modest.

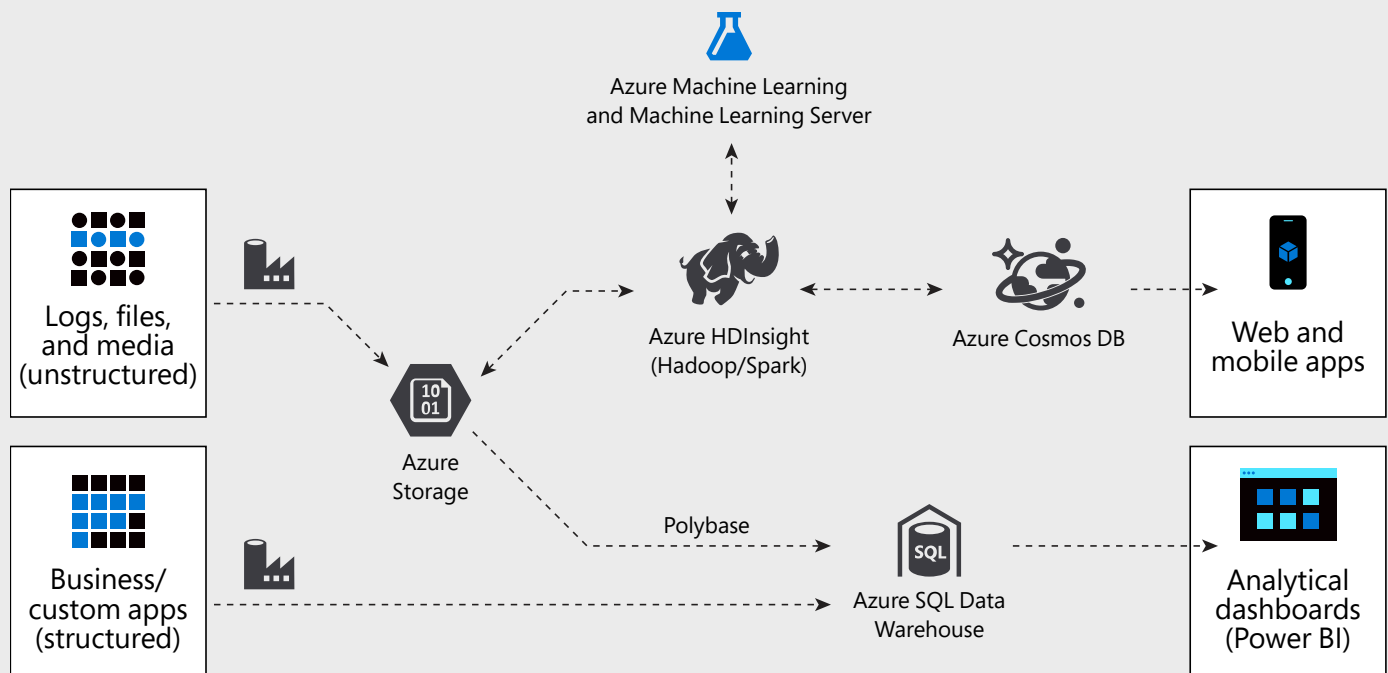


Figure 6. Analytics and data science.

Analytics and data science

You can use the following services to find the hidden patterns in your data and derive new insights:

- Use [Azure Machine Learning](#) to design and publish predictive models.
- Use [Azure HDInsight](#) to analyze data in Spark or Hadoop environments (and integrate R or Python code).
- Use [Azure Data Lake Analytics](#) for processing big-data jobs.
- Use [Azure Stream Analytics](#) for on-demand, real-time streaming analytics.
- Use [Azure Analysis Services](#) for enterprise-grade data modeling.
- Use [Microsoft Power BI](#) to create rich visualizations that bring your data to life.

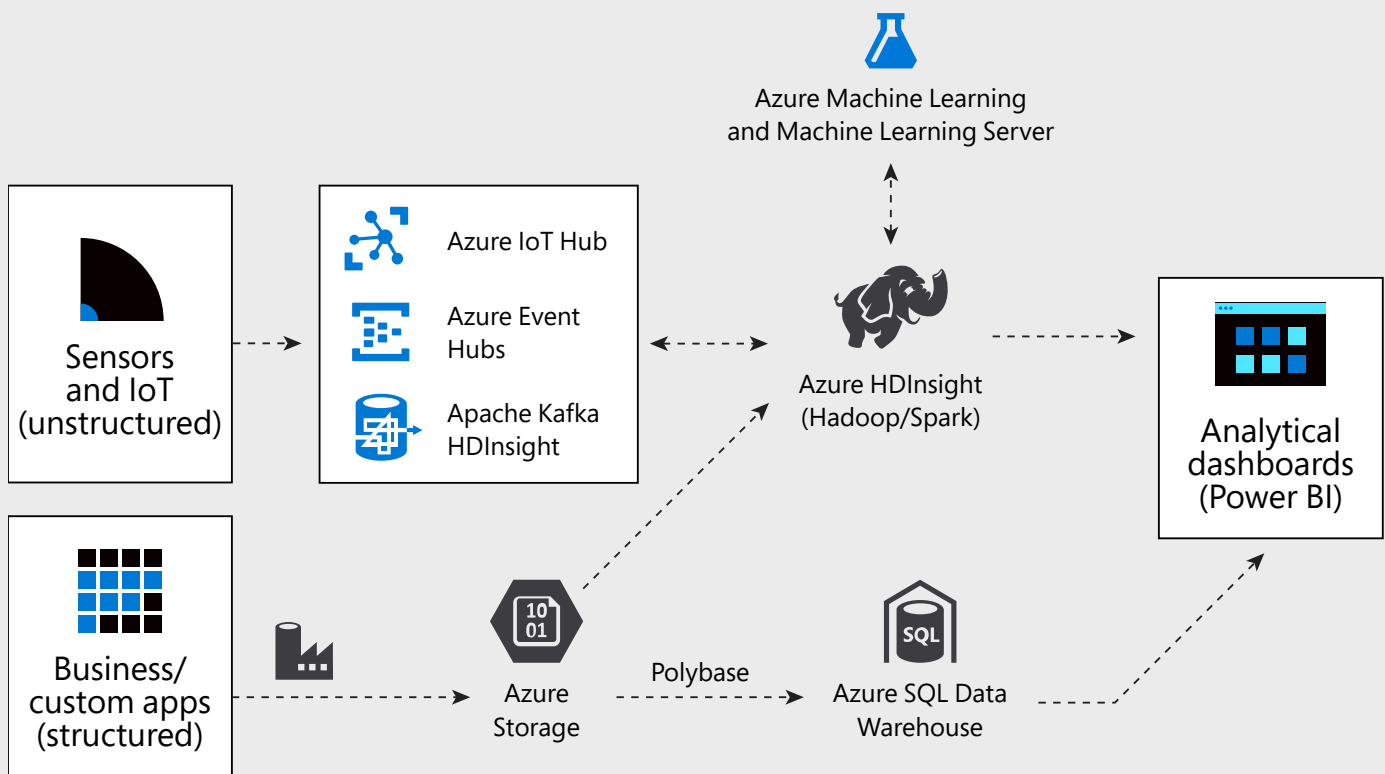


Figure 7. Stream processing and information management.

Stream processing and information management

You can use the following services to orchestrate stream processing and information management:

- Use [Azure Data Factory](#) to build pipelines and collect and orchestrate data for easier analysis.
- Use [Azure Event Hubs](#) to provide a staging area for incoming streaming data.
- Use [Apache Kafka for HDInsight](#) to buffer high-frequency, low-latency, real-time data.
- Use [Azure Functions](#) to capture change feeds from underlying operational data stores (such as Azure Cosmos DB) using an event-driven, serverless architecture.

Customer case study: PROS

PROS offers a dynamic pricing service that analyzes thousands of internal and external factors to help customers keep their pricing accurate, competitive, and profitable.

PROS achieves this by running enormously complex calculations on multiple terabytes of data. When the company made its move to the cloud, compatibility with Apache Spark, speed, scalability, and high availability were all essential requirements.

PROS chose Microsoft Azure and implemented Microsoft Azure HDInsight, which meant that the company could run its entire service platform on a single architecture. "Moving from on-premises to the cloud was essential, but we also hoped we could find a cloud system that offered us a complete computing platform," says Weiping Wang, Senior Software Architect at PROS. "This would free us up from managing the Spark cluster ourselves, which would become a much bigger job as we grew."

Apache Spark for Azure HDInsight offered the compatibility with Apache Spark technology that PROS required—including parallel processing, which multiplied the volume of data that could be processed in a given time. Apache Spark for Azure HDInsight also offered functionality specific to running large computations with vast

amounts of data. Jobs could be set up easily—regardless of their size and in a fraction of the time previously required—by spinning up an instance of Azure HDInsight. "Rather than having to configure hardware to keep up with demand, Azure HDInsight allows us to keep up with demand dynamically," says Wang.

Azure HDInsight also gives PROS an industry-leading SLA, with 99.9 percent availability, to help ensure business continuity and help protect against catastrophic events. Plus, changing the company's compute platform to HDInsight has helped the company convert manual functionality into automation in its product, making it easier for customers to update their pricing models as their businesses evolve.

In HDInsight, a process that formerly took several days now takes **just a few minutes.**

Ed Gonzalez

Product Manager, PROS



Read the entire [PROS case study](#).



Hear from [other customers](#) who are building big-data and analytics solutions on Azure.

For more information

- Learn more about [big-data and analytics solutions on Azure](#).
- Check out the [Azure Big Data blog](#).
- Learn about the [Microsoft Professional Program for Big Data](#).
- Discover [solution architectures for big data and analytics](#).
- Review GitHub code samples and labs for [Azure HDInsight](#) and [Azure Data Lake](#).
- Browse the [Azure HDInsight Frequently Asked Questions on GitHub](#).

Scenario 5:

AI

People don't want to adapt to technology. They want it to adapt to them—to help them “get things done” in a way that's easy and natural, and that complements how they perceive the world around them, process information, and interact with their surroundings.

Recent advances in [artificial intelligence \(AI\)](#)—the computer simulation of intelligent human behavior—are making this possible to an extent that, only a few years ago, seemed like science fiction. But how can you harness the power of AI within your own apps, augmenting them with human characteristics such as speech, vision, language, and knowledge to make your data even more compelling and useful?

At Microsoft, AI is all about amplifying human ingenuity with intelligent technology. That's why we're infusing AI into Cortana, our personal digital assistant; building AI into our database engines for greater performance and security; and equipping organizations with the power of AI through business solutions, accelerators, and practices.

It's also why we're building [an AI platform](#) with cutting-edge cognition technologies that empower companies to make more intelligent use of data in their apps to delight their customers. For instance, you could use the Microsoft AI platform to deploy predictive analytics on your data.

The Microsoft AI platform: Bringing intelligence to your data

You can use Azure Machine Learning and Azure Databricks, two components of the Microsoft AI platform, to run experiments on and simplify autoscaling capabilities in your own databases.

Azure Machine Learning

The [Azure Machine Learning Studio](#) is a fully-managed cloud service that enables you to easily build, deploy, and share predictive analytics solutions. It provides a simple drag-and-drop interface, so you can deploy your model into production in minutes—without writing any code. Experiments can be easily created by connecting datasets and modules in the interactive, visual workspace.

Datasets—A dataset is data that has been uploaded to Machine Learning Studio so that it can be used in the modeling process. Machine Learning Studio includes sample datasets like MPG data for various automobiles, breast cancer diagnosis data,

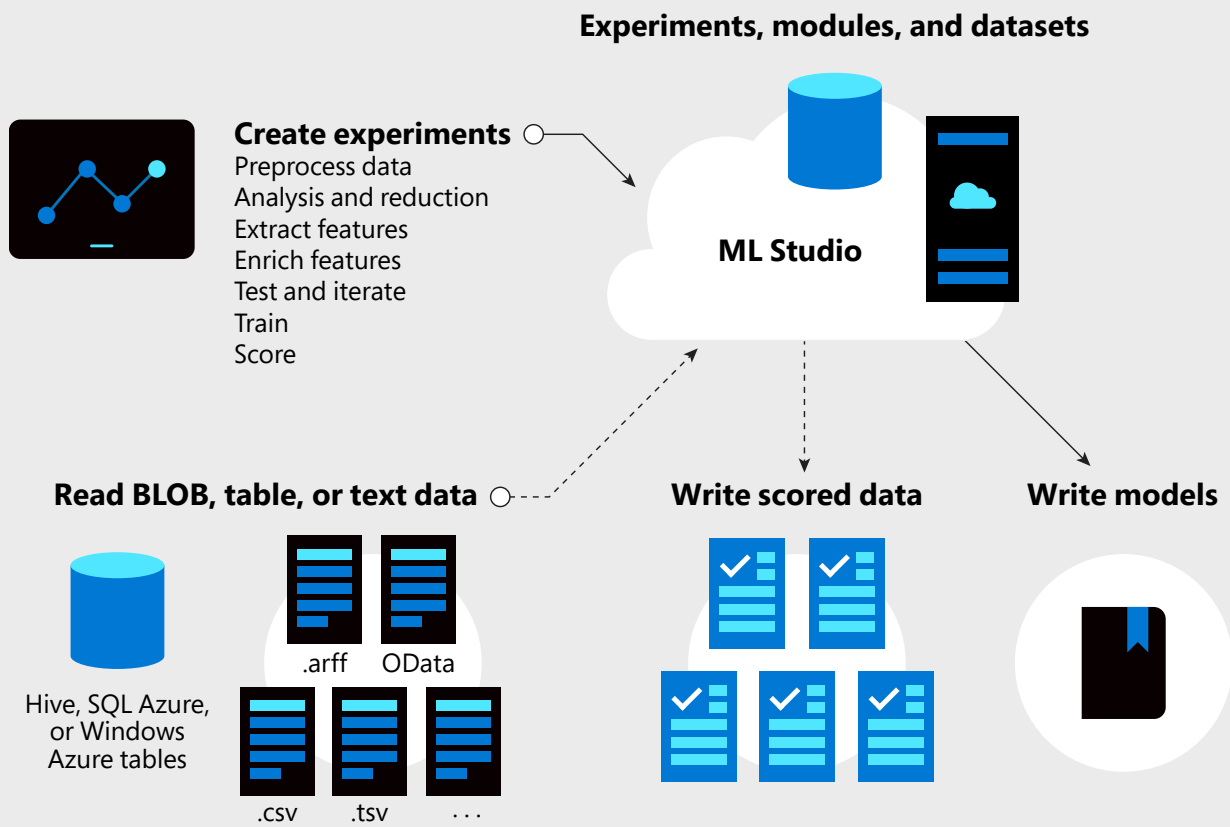


Figure 8. The Azure Machine Learning Studio makes it easy to create predictive models and run data science experiments with no coding.

and forest fires data. You can also upload more datasets as needed.

Modules—A module is an algorithm that you can perform on your data. Modules that are found in Machine Learning Studio range from data ingress functions to training, scoring, and validation processes, including converting to ARFF, computing elementary statistics, and creating linear regressions.

Try [Microsoft Azure Machine Learning Studio](#) for free—you don't even need an Azure subscription to get started.

Azure Databricks

In partnership between Microsoft and [Databricks](#), [Azure Databricks](#) is an Apache Spark-based analytics service within the Microsoft AI platform that provides one-click setup, streamlined workflows, and an interactive workspace for collaboration between data scientists, data engineers, and business analysts. It integrates natively with Azure Active Directory and other Azure services to help you build your modern data warehouse and machine learning and real-time analytics solutions.

Azure Databricks builds on the capabilities of Spark with:

Fully manage Spark clusters—Quickly create and autoscale clusters—even serverless clusters—and share them easily across teams.

Collaborative workspace—Explore data, prototype, and run data-driven applications in an integrated environment with interactive dashboards for dynamic reporting.

Databricks Runtime—Infrastructure complexity is abstracted, so specialized expertise isn't required to set up and configure your data infrastructure.

Customer case study: renewablesAI

Solar energy company renewablesAI wants to revolutionize the solar energy space through artificial intelligence. Their focus for success in solar energy is to bring down the cost to their clients. Their clients get their data from thousands of farms sending thousands upon thousands of messages per second in real time every day, so capturing all of it is a challenge.

Azure Databricks' integration with Apache Spark makes it easy to quickly spin up clusters with high reliability and performance. Cluster autoscaling lets renewablesAI and their clients ingest tens times more data than before and use all of it—not just a subsample. The notebooks in Azure Databricks also allow their teams to collaborate in real-time using R, SQL, Python, or Scala.

It's increased our productivity by **50%**. Azure Databricks is the best way to do spot-based analytics, period.

Declan O'Halloran

Commercial Director, renewablesAI



Watch the [renewablesAI case study video](#).

For more information

- Get credit to [try building intelligent solutions](#) with the Azure free account.
- Check out the [Azure Machine Learning blog](#).
- Watch the [Ignite 2018 sessions on Azure Databricks](#).
- Get the [e-book on solving big data and AI challenges with Azure Databricks](#).

Scenario 6:

Security and compliance

No matter what you're building, you need to protect users' data and ensure compliance with any applicable regulations.

You can do some of this within your app, which is why key security features that are specific to certain Azure services are covered earlier in this e-book, under the relevant scenarios for those services. However, when you rely on a cloud service provider for your infrastructure, you're also relying on that provider to help you keep your customers' data secure. It's the same with regulatory compliance; if the cloud services you're using aren't compliant, then neither is your app.

One of the best reasons to use Azure to host your data is to take advantage of its many built-in security tools and capabilities—additional layers of protection that you can use to achieve a defense-in-depth approach. For example, across all scenarios, you can use [Azure Security Center](#) for increased visibility and control over the security of all your Azure resources, use [Azure Active Directory](#) to help secure access to on-premises and cloud applications, use [Azure Key Vault](#) to safeguard cryptographic keys and other secrets used by your cloud apps

and services, and use [Azure Multi-Factor Authentication](#) to provide even more security for your data and apps.

Microsoft Azure: The trusted platform

The Microsoft Azure infrastructure is designed from facility to applications for hosting millions of customers simultaneously, providing a trustworthy foundation upon which you can meet your needs. Security and privacy capabilities are built-in from the start, beginning with the [Security Development Lifecycle \(SDL\)](#), which addresses security at every development phase and ensures that Azure is continually updated to make it even more secure. [Operational Security Assurance](#) builds on SDL knowledge and processes to supply a framework that helps provide secure operations throughout the lifecycle of cloud-based services.

Functional security areas

Depending on the cloud service model, responsibility varies for managing security at different solution layers. Azure provides a wide array of configurable security options so you can customize security to meet your unique requirements.

Security capabilities built into the Azure platform cover all major functional areas, including network security, database security, storage security, compute security, operational security, security monitoring

Responsibility	On-Prem	IaaS	PaaS	SaaS
Data classification & accountability				
Client & end-point protection				
Identity & access management				
Application level controls				
Network controls				
Host infrastructure				
Physical security				

 Cloud customer
 Cloud provider

Figure 9. Shared security responsibilities for different cloud service models.

and management, and more. The [introduction to Azure security](#) provides an overview of all these areas.

Compliance

In a world where data breaches and government requests for access to online customer information happen daily, you need a cloud platform that helps you ensure regulatory compliance. Microsoft provides the most comprehensive set of compliance offerings, including certifications and attestations, of any cloud service provider.

For example, in May 2018, a European privacy law, the [General Data Protection](#)

[Regulation \(GDPR\)](#), came into effect. The GDPR imposes new rules on companies, government agencies, non-profits, and other organizations that offer goods and services to people in the European Union (EU), or that collect and analyze data tied to EU residents, and it applies no matter where you are located. Microsoft is [committed](#) to GDPR compliance across its cloud services when enforcement begins May 25, 2018, and provides GDPR-related assurances in its contractual commitments.

- Review the [list of all Azure compliance offerings](#) on the Trust Center.
- Read the e-book on [Microsoft's Journey to Compliance](#).

Customer case study: Merck KGaA

Founded in 1668, Merck KGaA, Darmstadt, Germany, has grown to become one of today's leading healthcare, life sciences, and performance materials companies in the world. When the company first replaced several disparate internal platforms with EVA, an integrated digital workplace and social collaboration tool, employees connected to it and other internal applications via a virtual private network (VPN). As the company grew, routing all traffic through its on-premises environment became less and less tenable. And access via VPN didn't permit the use of multifactor authentication, which the company needed.

The company used Azure Active Directory Premium to give employees highly secure, single sign-on access to both the company's core cloud-based SaaS services and on-premises applications from their desktop computers or mobile devices. "By using Microsoft Azure, we are able to create a secure platform that allows us to embrace digital opportunities—SaaS applications and mobility device connectivity—that we knew would make our teams more productive," says Conor O'Halloran, Head of Identity Management for Merck KGaA, Darmstadt, Germany.

With Azure services, we have streamlined our **identity environment** and made it more secure by giving employees access to a platform we trust in the Microsoft cloud.

Marc Autenrieth

*Head of Core Solutions, Merck KGaA,
Darmstadt, Germany*



Read the entire [Merck KGaA, Darmstadt, Germany case study](#).

For more information

- > Get credit to [experience cloud-secured data services](#) with an Azure free account.
- > Learn more about Azure security at the [Azure security documentation site](#).
- > Get details on [Azure database security](#).

**Innovate faster
with Azure
data services**

With its broad portfolio of data services, Microsoft Azure can help you through any scenario, across your choice of stack, with innovative technology like AI built-in to wow your customers while making you more productive.

The Microsoft Cloud supports more than a billion customers in more than 140 countries and regions, providing a unique platform that helps you achieve your business goals.

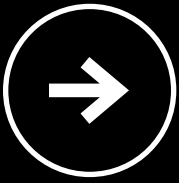
Azure supports a fully hybrid architecture, which helps you deliver the functionality you need regardless of where your data or compute power resides: your own datacenter, an Azure datacenter, other public cloud datacenters, or even a mobile device. A hybrid architecture also facilitates a phased approach. For instance, it helps you focus on harnessing the data within your walls before starting to tie-in data from customers, suppliers, and other sources.

With Azure, you can:

Easily build and deploy anywhere. Use your team's existing skillsets and tools you know and love to build intelligent apps and deploy without a change in code. Build once, deploy anywhere: in the cloud, on-premises, and to edge devices, with the confidence of global distribution to more data centers than any other provider.

Create an impact with an open platform. Choose your favorite technologies, including open source. Azure supports a range of deployment options, popular stacks, and languages, and a comprehensive set of data engines. Capitalize on this flexibility, plus the performance, scale, and security delivered by Microsoft technologies to build apps for any scenario.

Develop apps with built-in intelligence. Building intelligent apps using Azure is easy, because no other platform brings analytics and native AI to your data wherever it lives and in the languages you use. Take advantage of a variety of tools to easily introduce AI and machine learning into your databases without any coding.



Next steps

Wherever your data is, in whatever form it takes, Azure helps you unlock its full potential. Learn more about how your role will transform as your data moves to the cloud with [Azure data services](#), and get credit to try Azure with a [free account](#).