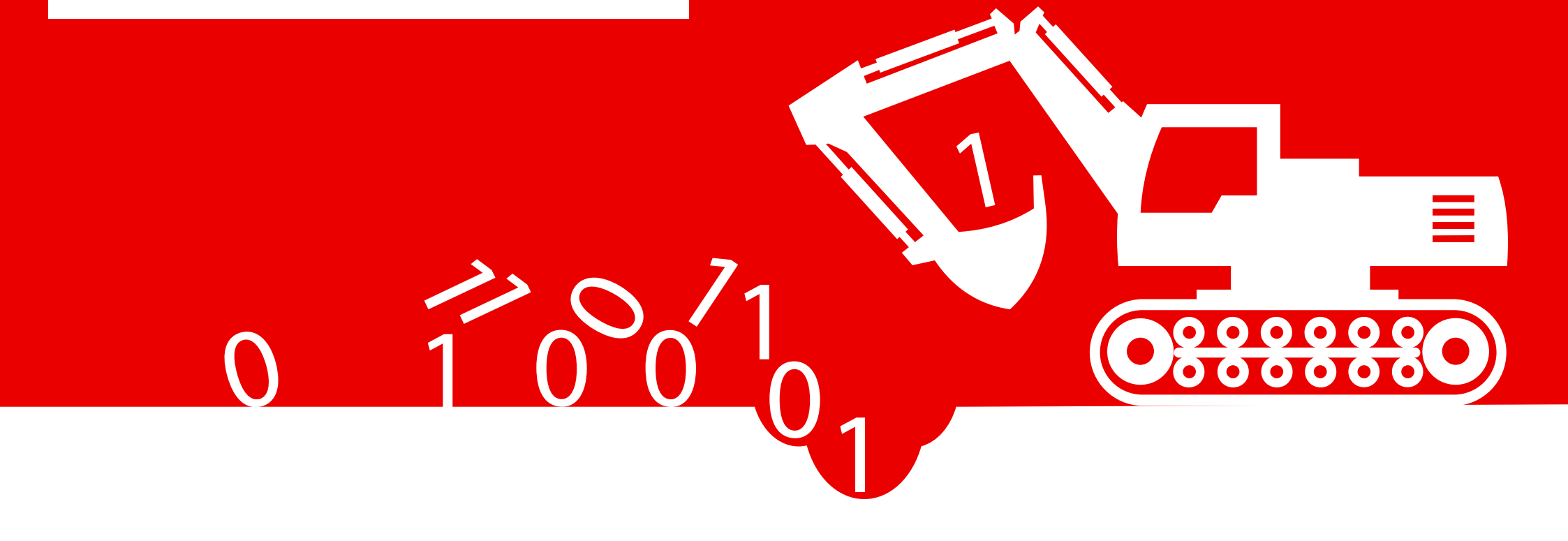
E-book

Building a Modern Data Estate on Microsoft Azure

The complete guide to becoming a data-driven organisation.

rackspace technology | Microsoft Azure



Introduction

The fourth industrial revolution is upon us, arriving in the form of digital transformation.

The leaders — and laggards — of the new era will in large part be determined by a single factor: the effectiveness of their approach to data management. What worked for businesses in the past may no longer be enough to meet today's need for speed, flexibility and innovation. A new approach is required that prioritises business intelligence and analytics to support greater agility and responsiveness. This means embracing emerging technology to democratise data and empower not just your C-suite, but also your entire organisation, with the right data and analytics tools to enable timely, insight-driven business decisions.

In today's ever-changing and increasingly competitive marketplace, organisations are challenged to find differentiation that sets them apart and maintains their momentum. The companies that can unlock and activate the vast amount of data will see several topline benefits.

According to recent **research from McKinsey**, data-driven companies are:

- 23 times more likely to acquire a customer
- 19 times more likely to retain that customer
- 19 times more likely to be profitable as a result

If you want to completely understand your business environment, including your customers and your value chain, you need a modern data architecture, and the ability to continuously enhance and optimise it. This new data architecture becomes the foundation for intelligent cloud-native applications that can incorporate all data that's needed, whenever it's needed.

However, building a modern data strategy is not an easy endeavor. It requires new skills, methodologies and a willingness to bring agility to data curation, management, integration and security.

Despite the urgent call for modernisation, few companies are finding success when trying to move their data analytics projects into production.

Common challenges that conspire to drive data architecture to a new level of complexity include:

- **Data volume:** The explosion of new data from the Internet of Things (IoT), customer and value chain interaction capture, and the ability to store and process large amounts of unstructured data generated by traditional sources
- **Emerging technology:** A wide range of newer technologies and methodologies, including serverless data platforms, data lakes, data pipelines, DevOps processes and tools, AI/machine learning and more
- **Data discovery:** Unknown data sources, poor data quality, data silos and compliance restrictions
- **Cost:** Infrastructure costs, lack of structure for utility, limited and expansive talent and large investments with no guaranteed return
- **Skills gap:** Data analytics skillsets often require specialised education, training and experience, but this is not typically available to IT teams.

In this e-book, we seek to address the barriers and opportunities facing today's modern businesses and provide an outline for how to overcome these challenges through the deployment and ongoing management of a modern, cloud-based data platform on Microsoft® Azure®.

We'll explore what it means to be a data-led organisation, we'll identify the core design principles of a modern data estate, and we'll go beyond the technology to outline what it takes to be truly successful in today's data-driven world.

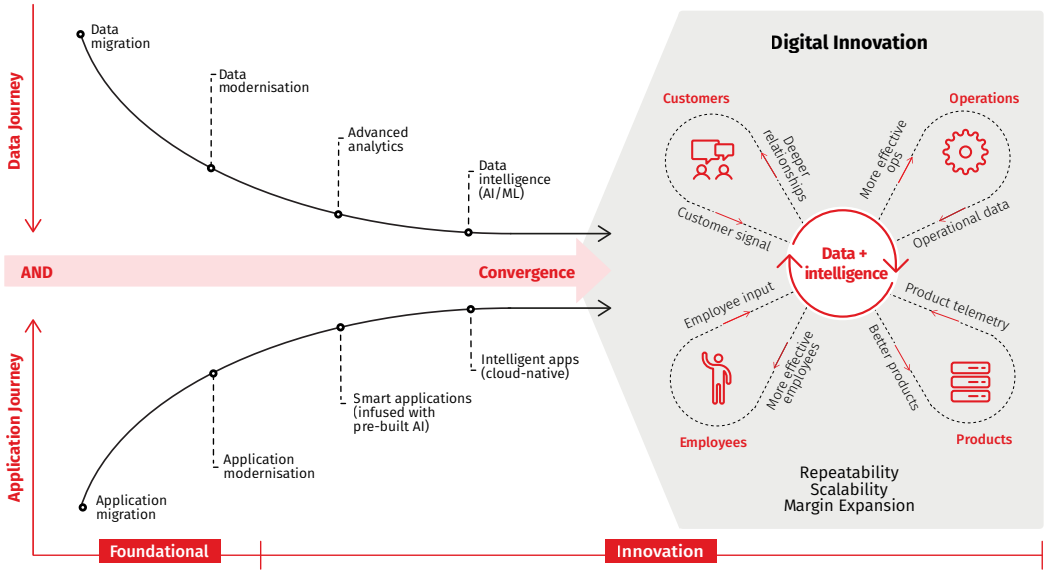
PART I. What does it mean to be a data-led organisation?

What is the data-led approach, and how does it differ from an application-led approach?

Most organisations begin a data modernisation journey with a cloud-envisioning session, a standard methodology for helping make key decisions regarding their move to the cloud. These envisioning sessions are typically followed by a cloud assessment, during which the application landscape is inventoried to identify application dependencies that will later help to shape a cloud migration and modernisation plan.

Typically, these plans are heavily focused on lift-and-shift migrations or application modernisation. In both of these scenarios, data is often a secondary consideration, usually piecemealed into different application and workload silos, without a view of how to ensure trust or connect the data cross the entire application portfolio.

By contrast, a data-led organisation approaches modernisation by realising the value of its data from the beginning, and quickly opening the door to broader discussions on data estate modernisation. For some, the data journey occurs separately from the application journey, thereby accelerating the pace of data estate modernisation. For others, the data and application journey will run in parallel — often converging to build intelligent tools and processes that unleash the possibilities of digital innovation.



When organisations lead with data, smaller data projects that would otherwise be executed in siloes materialise into a connected, holistic approach to modernising the data estate. You can start with relatively small business challenges, including financial reporting, data quality or security incidents. Or you can start with a business need such as point-of-sale data needs, IoT project integrations, improved efficiencies, reduced costs, or business insights and visualisation. Once value and ROI are proven for the initial business challenge, you can then move on to more radical, large-scale data estate transformations.

As you embark on the data estate modernisation journey, there are two best-practice goals to embrace to create maximum value. The first is continuous and iterative design, and the second is the efficiency brought about through repeatable and scalable processes.

Continuous and iterative design

Operating in short, one- to four-week sprints based on agile methodology accelerates the discovery process, proves value early and validates the investment. In the early stages, ROI is best demonstrated by focusing on a single use case and the impact on a set of prioritised success indicators. This helps to achieve stakeholder buy-in early before scaling the project to larger parts of the estate. As the process of discovery, agile deployment and value realisation is executed over additional use cases, the value multiplies. The data builds upon itself, creating new value and encouraging new sources, which, in turn, create more value.

Efficiencies through repeatability and scalability

Execution across multiple use cases enables organisation to catalyse value realisation, and establishing structure and governance, while realising operational efficiencies that increase profitability. Three assets in particular are key to building a repeatable and scalable modern data strategy. The first is a solution library comprised of algorithms, machine learning and analytics models, and data patterns collected across projects to be leveraged for future projects.

The second is a continually refreshed centralised knowledge repository containing reference architectures, solution templates, use cases, DevOps process templates, solution runbooks, deployment guides, governance repositories, analytics practice libraries, data connector repositories, and frameworks for assessment, scripts and security.

The third key asset is a methodology to showcase the “art of possible” to key decision makers. This methodology should help key stakeholders visualise specific use cases and business value that can be unlocked by modernising their data.

Part II. Design principles of a Modern Data Estate

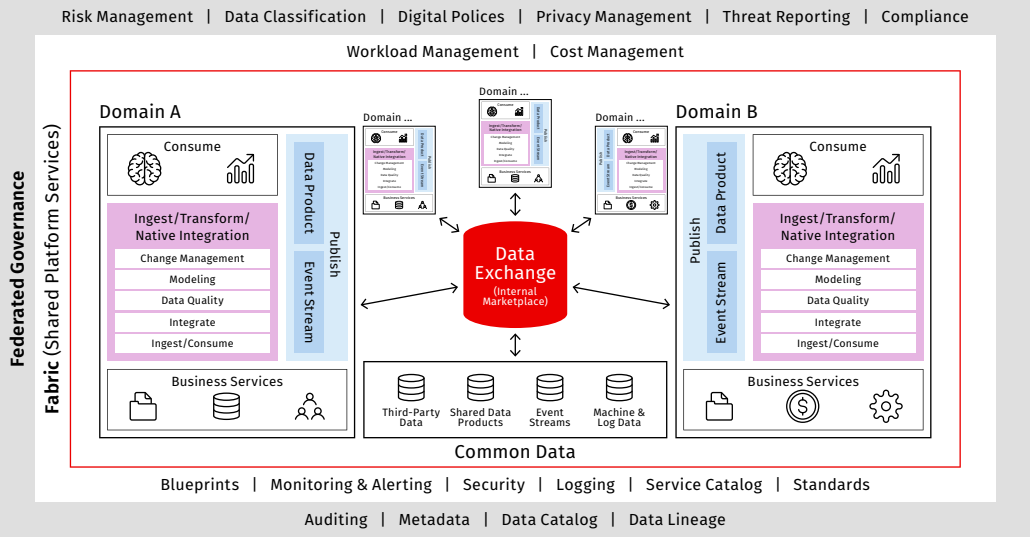
What are the core design principles?

A modern data estate leverages the power of the cloud to ingest, process, store, serve and visualise structured and unstructured data from multiple different sources.

Generally, the litmus test for a modern data estate is whether it delivers on the following five tenants:

- Moves analytics closer to its consumption
- Centralises administration and governance
- Has low cost of experimentation
- Unifies platforms to support a diverse range of tools in a single solution
- Offers scalability and resilience

End State Governed Federated Architecture



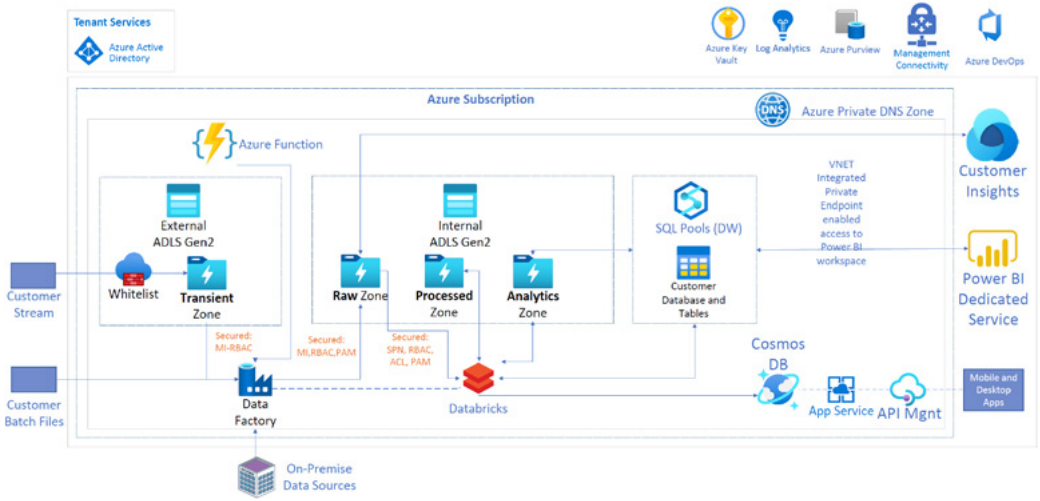
The core design principles that underpin a modern data estate will enable the following capabilities:

1. **Self-service for a diverse range of users:** A modern data platform should be intuitively usable by almost everyone in the organisation. Users should be able to discover and understand not only the data, but also metadata behind it, such as column descriptions and lineage. Users should be able to gain insight from their data with minimal involvement from IT.
2. **Agile data management:** You should be able to establish an open and lake-centric hub that allows data engineers to connect and curate data from different sources while eliminating sprawl and creating custom user views.
3. **Support for powerful AI models:** Accelerate analysis by developing AI models on a single foundation without data movement and reduce the time data scientists need to deliver value.
4. **Flexibility, quick deployment and consumption pricing:** You should be able to deploy your data platform quickly while maintaining cost controls and the flexibility required for easy adjustability as organisational needs change.

Data from a variety of unstructured, structured and semi-structured sources is published to a virtual data lake that users centrally access to generate reports, APIs and dashboards. To eliminate duplication and optimise efficiencies, compute is brought to the data lake, as opposed to data being extracted, copied locally and then compute applied.

Shared managed services and analytics tools are also made available help maintain the integrity of the estate and its users. Finally, the entire system is contained within a single security context to ensure data is protected and secured.

Azure-based Modern Data Architecture



The solution architecture demonstrates the ability of a modern data estate to ingest a variety of complex data from multiple sources, including high-velocity data from IoT devices, sensors and other gadgets, as well as high-variety and high-volume data from images, video, audio, csv, xml, log and more. This data is ingested through streaming or batch processing, and then processed and stored in a centralised, raw storage location before being transformed from its source format to the format required for analysis.

Processes will vary, such as for streaming data, through which near real-time insights can be unlocked. This data can be transformed in the streaming pipelines themselves, as opposed to being stored before being transformed. Finally, the data is ready to model and serve, ultimately powering data-driven applications and supply rich data visualisations to users.

Microsoft Fabric: A SaaS-based solution that unifies data across storage, analytics and operations

While democratising analytics and insights is a necessity, a successful analytics strategy hinges on the quality and completeness of the data it's built on. In other words, it is essential to combine data across all disparate sources with a single platform to leverage a unified, holistic source of truth. Yet, even with centralised data, it can be difficult to build and manage analytics models across the organisation without a unifying service.

Microsoft Fabric combines Data Factory, Synapse Analytics, Data Explorer and Power BI into a single, unified experience in the cloud. The open and governed data lakehouse foundation is a cost-effective and performance-optimised fabric for business intelligence, machine learning and AI workloads at any scale. It is the foundation for migrating and modernising existing analytics solutions, whether they be data appliances or traditional data warehouses.

By establishing connectivity and integration, organisations can transform their unstructured and siloed data into a valuable strategic asset through:

- Data modernisation backed by Microsoft Azure
- Cloud-native applications at any scale
- Responsible, powerful AI to make more informed decisions
- Analytics and insights at a fast rate
- Responsible machine learning and artificial intelligence
- Governance backed by Microsoft Preview

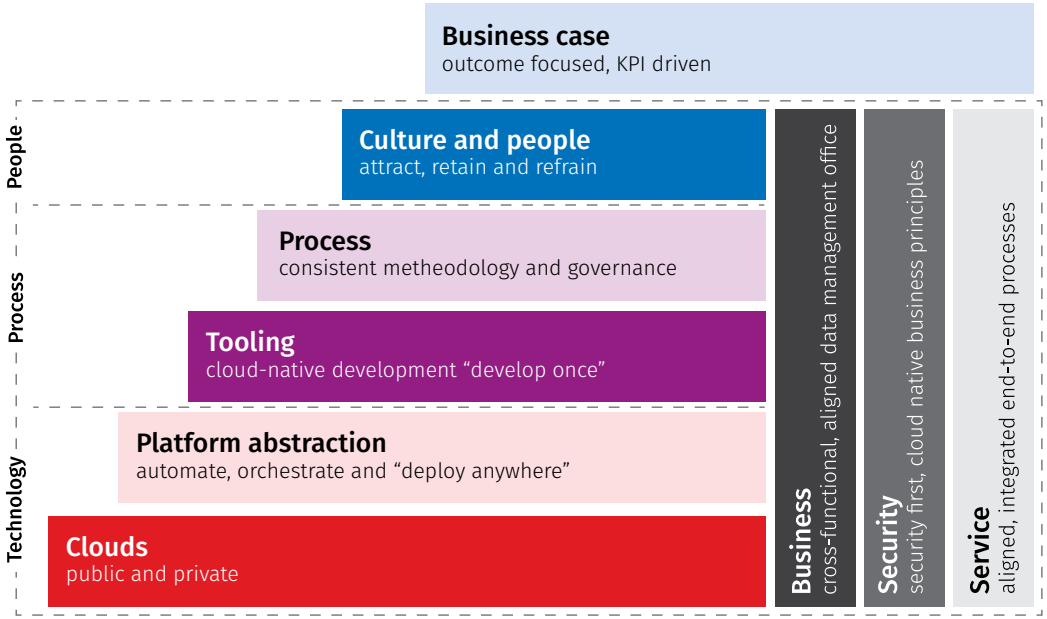
By providing employees with the tools to harness your corporate data, you can help them make better decisions for your business.

Rackspace Technology® can help you modernise your data estate to harness the analytic power of Microsoft Fabric, and accelerate time to value with up to 70% less implementation time and cost. Our extensive team of highly skilled, certified cloud data engineers and solution architects partner with you on every step of your transformation, helping you accelerate access to real-time analytics and intelligence.

PART III. Beyond the technology: ongoing data management

A successful data modernisation strategy requires more than just technology to deliver expected business outcomes. Although the technical aspects of an implementation are critical to its success, organisations too often focus on these aspects at the expense of organisational change management. They fail to adequately consider how best to manage the experience of end-users whose acceptance is critical to successful use of any new application. In addition, organisations often fail to drive the appropriate process changes needed to deliver a smooth transition to a new operating model. This is often exacerbated by the fact that internal resources haven't historically been developed to provide specialised expertise across a wide range of new and emerging technologies.

From a technology, process and people perspective, the diagram below illustrates what a modern data approach should encompass.



A modern approach to data should incorporate the following attributes:

- **Outcome focused:** Begin by building out business use cases and applications focused on customer outcomes with well-articulated metrics of success.
- **Deep expertise:** Adequately account for the expertise required to design, deploy, manage and optimise the data platform.
- **Attention to detail:** Implement processes to ensure data from new sources is properly ingested consistently while minimising risk.
- **Cloud-native:** Build a cloud-based platform with the technologies to leverage automation and orchestration as core tenants as opposed to an afterthought.

How data service providers can encourage transformation

As businesses move toward a modern data estate, the need for professional services has also expanded. The right service provider should be able to function as an extension of your team, filling the skills gaps where needed and providing expertise across the entire data journey: design, build, manage and optimise.

Working with the right partner can help ensure that your data modernisation projects are "stickier," quicker-to-market, and have faster time-to-value realisation compared to heavy enterprise lines of business applications that may take years to modernise and move to the cloud. However, we know that a modern data estate alone does not guarantee long-term success and profitability.

As businesses and their data grow, monitoring and managing large volumes of data ingestion across several data pipelines throughout their entire data analytics platform becomes a labor-intensive task, often depleting development teams of time that is better spent on the application and core engineering efforts.

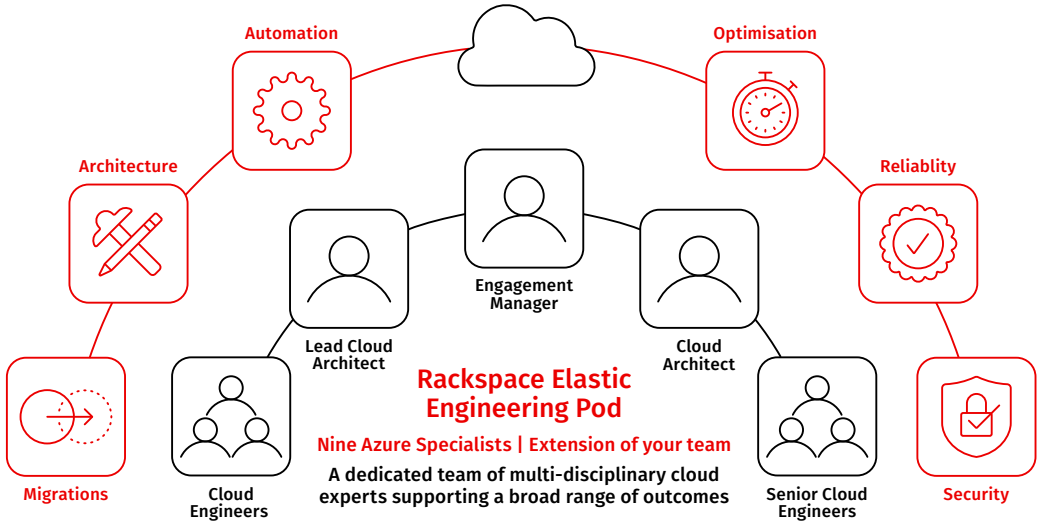
What is DataOps?

To accelerate innovation and remain competitive, companies must adopt a collaborative data management practice focused on improving communication, integration and automation of data flows between data managers and data consumers across the organisation.

Of course, achieving predictable delivery and change management of data, data models and related artifacts doesn't come without challenges.

Common barriers include:

- Complex data landscapes and processes with highly diverse data sources and tools
- A lack of agility, making it hard to respond to rapidly changing business requirements
- The need to manage a growing number of reports and dashboards over the entire lifecycle
- An inability to understand business needs and quickly respond to changing requirements
- The challenge of transforming statistical models from and experimental stage to production
- A lack of alignment among data scientists, product owners and data engineers



In response to these challenges, Rackspace Technology has developed an innovative managed service model called Rackspace DataOps, built to support today's modern data estate. Through the orchestration of people, processes and technology, Rackspace DataOps delivers an organisation-wide data management solution that improves the speed and accuracy of analytics through process automation and cultivating collaboration between data scientists, engineers and technologists.

The right approach to modern DataOps can enable predictable delivery and change management of data, data models and related artifacts, and deliver on three primary initiatives:

- **Data lifecycle process framework:** Incorporates DevOps, agile development and statistical process controls
- **Collaborative data management practice:** Improves communication, integration and automation
- **Cloud platform and tooling:** Orchestrates and automates security, quality and metadata

In addition to providing greater value from data, a modern DataOps approach also emphasises the importance of continuous innovation that delivers needed data to modern applications, including cloud-native and SaaS, and incorporates the IoT as appropriate.

By optimising agility and automation wherever possible, DataOps facilitates a virtual cycle of continuous improvement, keeping your data architecture where it needs to be to deliver on both the immediate and future needs of your business.

Conclusion

The pivot to becoming a data-driven organisation is fundamental for businesses. To fully capture this opportunity, they must rethink the way they approach their cloud journey. They must shift from an application only approach to an application and data-led approach. This means conducting early discovery efforts, and assessing the data estate as a whole, apart from any singular application, to ensure repeatability and scalability across current and future customer use cases that impact business outcomes.

Organisation looking to adopt a data-led approach should focus on these key tenants:

1. A data-led approach should be designed around two best practices: continuous and iterative design to prove value early, and repeatability and scalability to catalyse value realisation and increase profitability.

2. If you want to completely understand your business environment, including your customers and your value chain, you need a modern data architecture and the ability to continuously enhance and optimise it. This new data architecture becomes the foundation for intelligent cloud-native applications that incorporate all data needed when it's needed.

3. This will require new skills, methodologies and a willingness to bring agility to data curation, management, integration and security. To maximise value, look for data service providers who can support you across every stage of your data journey.

These tenets offer a roadmap for organisation looking to innovate with data and, in doing so, they become the linchpin for companies in a time where data may very well determine the winners and losers of the fourth industrial age.

Next steps

Schedule a free **two-hour strategy session** to see how a modern data estate and analytics strategy with Rackspace Technology and Microsoft Azure Synapse can help your business thrive.

rackspace technology.

About Rackspace Technology

Rackspace Technology is a hybrid multicloud solutions expert. We combine our expertise with the world's leading technologies — across AI, applications, data and security — to deliver end-to-end solutions. We have a proven record of advising customers based on their business challenges, designing solutions that scale, building and managing those solutions, and optimising returns into the future.

As a global hybrid multicloud technology services pioneer, we deliver innovative capabilities to help customers build new revenue streams, increase efficiency and create incredible experiences. Named a best place to work, year after year according to Fortune, Forbes and Glassdoor, we attract and develop world-class talent to deliver the best expertise to our customers. Everything we do is wrapped in Fanatical Experience® — our obsession with customer success that drives us to help each customer work faster, smarter and stay ahead of what's next.

Learn more at **www.rackspace.com/microsoft**

March 6, 2024 / Rackspace-Ebook-Data-Modernization-on-Azure-EMEA-MST-TSK-5801