

An executive's guide to delivering value from data and AI

5 steps for maximizing
AI success in your
organization



AI success relies on data

74%

of enterprises using gen AI report ROI within the first year, with 86% of those reporting increased revenue noting a 6% or more increase¹



70%

of organizations adopting gen AI have experienced difficulties with data governance, integrating data into AI models, and insufficient amount of training data²



80%

of organizational data is untapped in the form of unstructured data—customer contracts, call center recordings, employee handbooks, and product specs³



62%

of organizations find data security and privacy are the greatest challenges for gen AI adoption⁴



86%

of organizations recognize that delivering contextual and relevant user experiences through gen AI integrated databases has a substantial positive impact⁵



Across the board, organizations are eager to embrace gen AI, but they're facing challenges in scaling from pilot projects to full production deployments. What's holding them back?

Their data.

As a data leader, you are at the forefront of fueling your organization's AI transformation. Gen AI thrives on comprehensive, accurate, and high-quality organizational data. Unfortunately, many organizations struggle with fragmented data silos, poor data quality, lack of proper governance, and other challenges.

To unlock the true potential of gen AI for transformation, organizations need to rebuild their data strategies—focusing on building modern data platforms that serve as a robust foundation for gen AI initiatives.

Based on our work with tens of thousands of organizations on their data and AI platform strategy, this guide shows you five steps to building a solid foundation for AI success.

5 steps for building a data foundation for AI success



01

Create an AI-first data strategy



02

Connect all your data to AI to drive insights



03

Boost productivity with agents in every data task



04

Ensure data security and governance



05

Improve efficiency to reduce data cost and scale



Create an AI-first data strategy

01

Data is everybody's business, and it's the fuel for AI.

Combining a strong data strategy with a thriving data culture ensures that every team member understands, uses, and contributes to the value of your data—and leads to better utilization of AI.



How do you build a strong data strategy?

One that aligns with your organizational culture, fosters collaboration across business units, and creates a data ecosystem that delivers accessible data experiences. And most importantly, a data strategy that's tailored to meet the demands of the AI era.

Here's how. An AI-first data strategy should:

💡 Treat data as a strategic asset.

Shift the organizational mindset to recognize data's intrinsic value and align people, processes, and technology to drive business goals.

Ensure that your data and AI strategy supports your business goals around revenue growth, marketing objectives, customer service aspirations, margin enhancement.

Weigh the implications of inaction in the context of the organization's competitive position.

Implement strong data governance policies and invest in modern cloud data management tools that address the data to AI lifecycle.

Consider whether your proposed data strategy is able to support your future AI strategy at scale—or whether it will be a lift to organize your data for each model.

Provide employees with the skills, tools, and incentives to utilize data effectively in their decision-making processes.

Create data products and build a monetization strategy that unlocks new revenue streams. Differentiate your data products with embedded analytics and AI infusion.



Create product owners and define roadmap ownership.

Consider leveraging a data mesh strategy. Appoint a central data team that establishes governance requirements, supports the data platform, and creates the data strategy framework for the organization. This team should periodically collect feedback from stakeholders to understand if the data stack is working well and continually meets emerging organizational needs.

Organize the rest of the team, including stakeholders and data practitioners, by data products and data domains—resulting in stewardship, continuous management, and improvement in their respective domains. Stakeholders should be responsible for their data products from sourcing to offering insights.

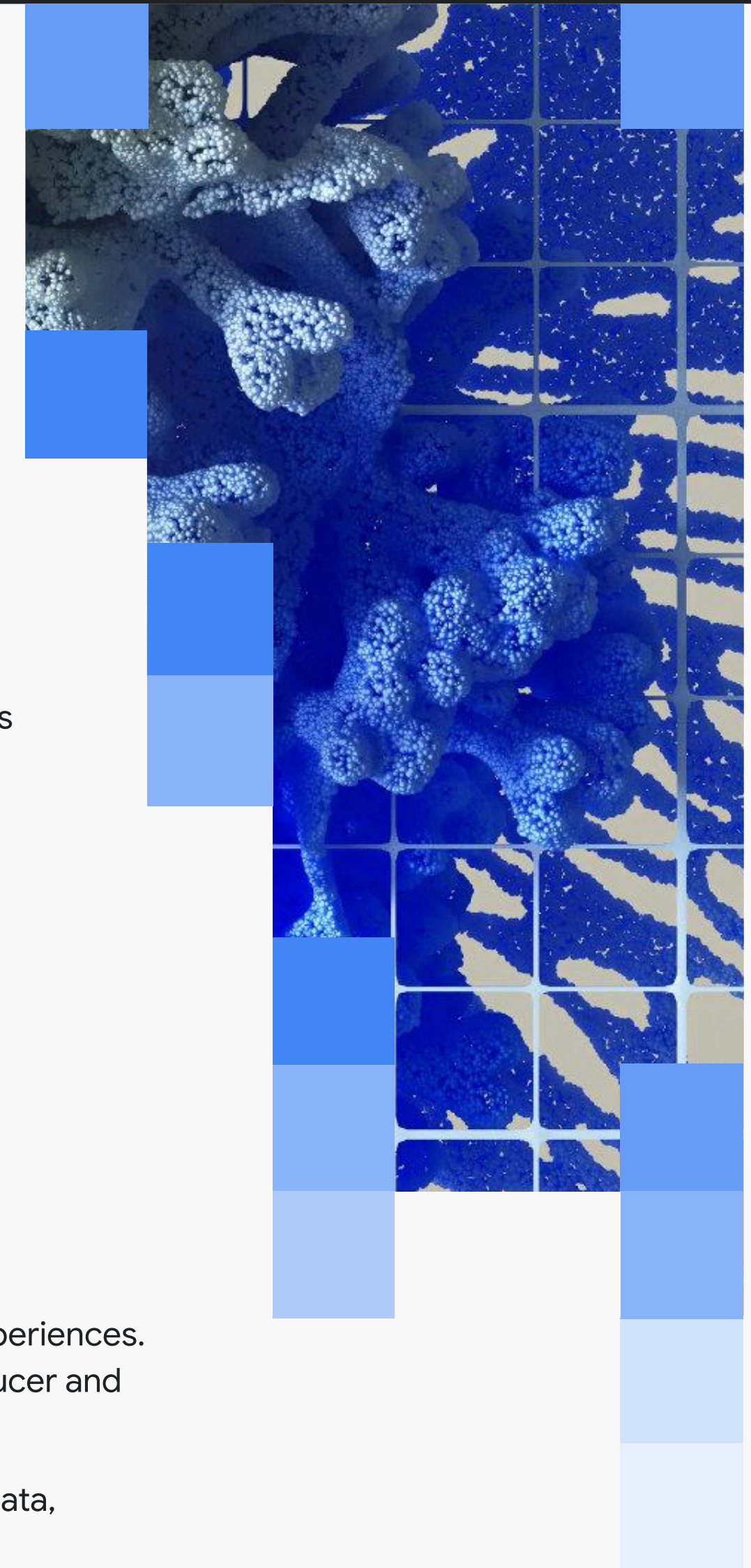


Build a data ecosystem.

Empower data producers and consumers with streamlined, unified, and productive experiences. Also, ensure your data ecosystem is able to seamlessly support the blurring of data producer and consumer roles as models are trained and retrained.

Establish principles and practices for publishing, discovering, building on, and trusting data, enabling efficient collaboration and innovation.

Provide a roadmap to address inefficiencies in the “usability” of data. This roadmap helps you realize the true value of data by maximizing its accessibility, discoverability, and quality.



An effective data culture creates a shared understanding of using, owning, and creating value from data.

This shared set of values, expectations, and capabilities is the foundation for successfully implementing a data strategy.

What will yours look like? Your organization is unique, and so your data culture and strategy will be, too. But know that the more clearly you establish and internally communicate your data culture and strategy, the easier it is to drive internal adoption and develop the required skills internally to unlock the value of AI.

Establishing and sharing an AI-first data vision is crucial to removing the central bottlenecks and silos in your data.



Connect all your
data to AI to
drive insights



02

Organizations thrive on real-time insights.

And they're powered by today's AI technologies.

But in order for gen AI to deliver, your data needs to be available as input. All of it. Not sitting in silos. Not inaccessible to a unified view.

The need to unify your data has not only become more urgent in the AI era, it's also become more challenging. Simply putting all your data in one place is often not practical or cost effective. That's why it's important to create a cohesive cost optimized data foundation that breaks down silos as well as works across clouds—establishing a single source of truth that acts as the bedrock for AI initiatives.

Here's how to do it.



An open data approach to fuel your lakehouse for AI

Multimodal data often becomes fragmented across different platforms, and potentially even stored in diverse table formats, making it difficult to unify your data and incorporate unified security and governance policies. To take full advantage of AI, it's imperative to create a single source of truth with a unified platform across all forms of data, including open formats.

In addition, a highly scalable architecture is crucial, enabling the unification of transactional and analytical systems without impacting performance. This allows for seamless data access between these systems, even with demanding transactional workloads.



Here are best practices when developing an open data approach:



Map capabilities you need in the data platform.

When selecting a platform for your data foundation, prioritize scalability, flexibility, and integration across data management, processing, governance, ML, AI, and BI capabilities. Map all your use cases and workflows to the capabilities of your data platform to identify any gaps.



Connect structured, semi-structured, and unstructured data.

Your data platform should accommodate all types of data, including structured data (e.g., relational databases), semi-structured data (e.g., key-value, JSON, XML), and unstructured data (e.g., text, images, video). An increasing amount of data—including customer information, employee handbooks, and more—is raw, unstructured, and diverse. In fact, unstructured data currently makes up 90% of business data, and so it's important to make sure your platform of choice is suited for extracting insights from a variety of formats.



Define your strategy for multimodal data for enhanced search and AI applications.

Build a single-copy architecture for all your structured and unstructured data across clouds and open file formats. Accomplish this by converting multimodal data into numerical representations, or vectors, for AI models including LLMs to generate more contextually relevant and meaningful responses when interacting with users or processing information.



Simplify your data platform with a single pane of glass across open formats.

This includes a single UX and collaborative workflows. Ensure compatibility with popular open data formats like Apache Iceberg, Apache Hudi, and Delta Lake, each offering unique benefits for data management and analytics. Unification also alludes to interoperability across engines (i.e., SQL, Spark, Python).



Implement data cataloging and metadata management.

Create a comprehensive data catalog to enable easy access, and understanding of your assets through AI-enabled data discovery, profiling, quality, lineage, and lifecycle management. Use a unified metastore and governance model to ensure consistency, flexibility, risk mitigation, compliance, and trust across all assets.

Harness all data types and formats.

Once you've integrated disparate datasets, you empower your AI models to leverage the full spectrum of enterprise information—resulting in more accurate, actionable insights. To effectively harness the power of your data and AI in this complex landscape, a robust and adaptable strategy for analytics is essential.

Here are some tips:



Create a unified storage engine.

Build a unified interface for analytics and AI engines to query data in a secure, governed, and performant manner. Maintain a single copy of data and make it uniformly accessible across engines, while centrally managing security policies in one place, and have it consistently enforced across the engines.



Harness your operational and transactional data.

Models trained on generic information aren't enough to deliver the personalized experiences people expect. Make it easy for your application developers to build contextual, relevant, and accurate gen AI experiences with the right databases storing your operational data. Learn how in our recent whitepaper on [gen AI and operational databases](#).



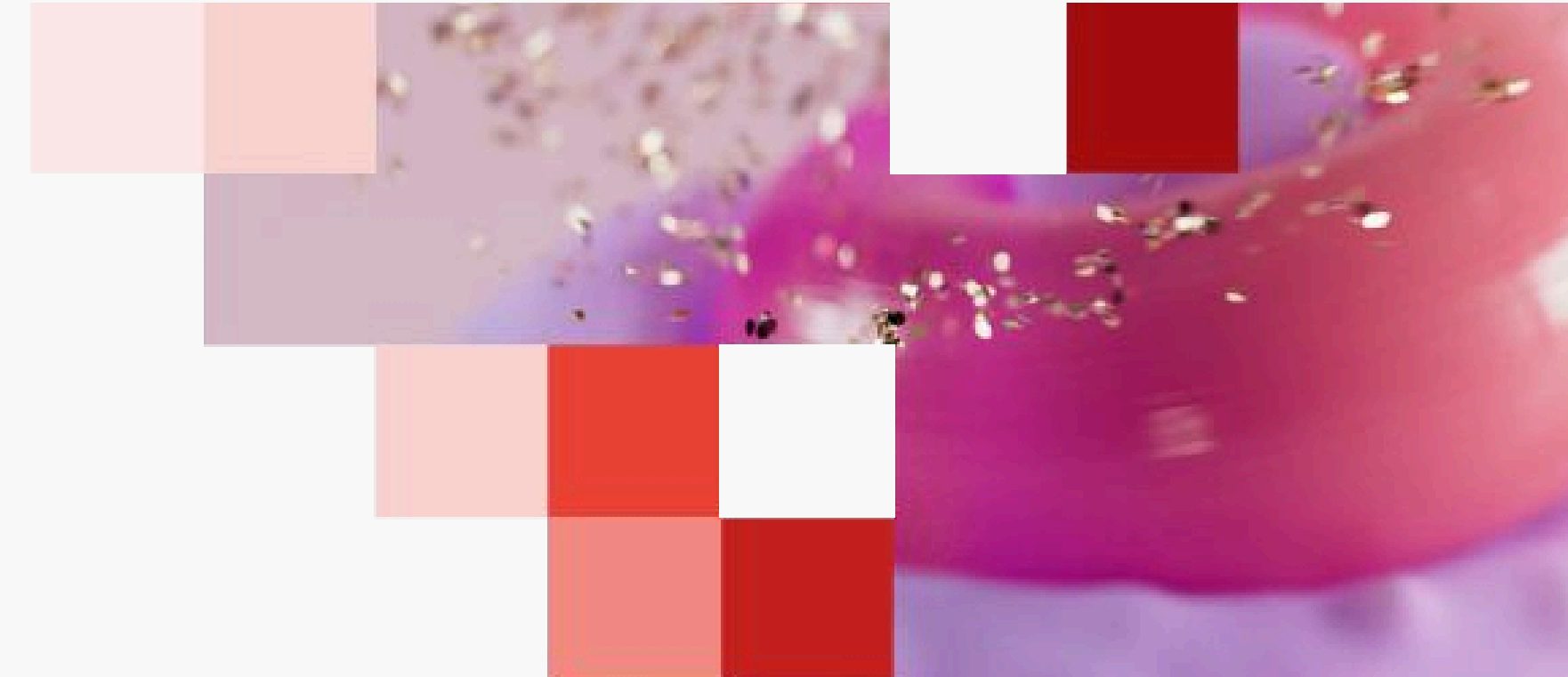
Enable cross-cloud data access and analysis.

Implement solutions that allow you to access, process, and analyze data stored in different cloud environments seamlessly. This approach reduces risk, minimizes data duplication thus reducing cost, and improves security by centralizing data governance and access controls.



Support diverse analytics engines.

Your architecture should support different types of analytics engines. For example, Spark for large-scale data processing, graph processing to find hidden connections in your data, a BI engine for interactive dashboards, and TensorFlow for ML model development.



Activate your data for real-time insights.

To remain competitive, organizations need to make decisions and respond to events with lightning speed. Retailers, for example, should be able detect sudden surges in demand and adjust inventory, optimize logistics, and even personalize marketing messages to capitalize on current trends.

Real-time data streaming has emerged as a critical technology enabling organizations to process and analyze data as it's generated. Incorporating streaming analytics into your data and AI strategy requires careful consideration of several architectural aspects to ensure efficient and effective processing of real-time data.

Real-time data processing requires a platform that offers:



High-throughput, low-latency data ingestion.

Your platform must efficiently handle diverse, high-volume data sources with minimal delay. It should use message queues (e.g., Kafka, Pub/Sub) and ingestion frameworks (e.g., Kafka, Kinesis, Pub/Sub) to support various data formats (e.g., JSON, Avro) while ensuring data consistency through a schema registry.



Scalable and fault-tolerant processing.

Horizontal scaling handles increasing data volumes and maintains availability in the event of failures. This requires distributed processing frameworks (e.g., Spark Streaming, Flink, Dataflow), microservices architecture, and containerization (e.g., Docker, Kubernetes). Stateful processing is essential for computations that span multiple events.



Flexible data transformation and enrichment.

Real-time data transformation, enrichment, and aggregation using stream processing APIs (e.g., Beam, Kafka Streams), integrated with external data sources and machine learning models is crucial for advanced analytics. Data quality checks should be implemented to ensure data integrity.



Real-time analytics and visualization.

Integrating with the appropriate tools enables the creation of real-time dashboards, alerts, and ad-hoc querying. Deeper analysis can be achieved through connectivity with business intelligence tools.



Operational monitoring and management.

Robust logging, monitoring, and alerting mechanisms ensure operational visibility and proactive issue resolution. Auto-scaling and load balancing are essential for optimal performance under varying workloads.



Bayer created a modern data solution to efficiently collect and analyze billions of data points.

A division of Bayer dedicated to agricultural advancements, Bayer Crop Science supports hundreds of teams and applications with maps, satellite imagery, and environmental data for research and development. Its modern data solution, Field Answers, stores and analyzes vast amounts of observational data, but extensive load testing revealed that a new database solution was necessary to handle the dramatic increase in traffic. The team turned to AlloyDB for PostgreSQL, drawn by its compatibility with existing systems and low replication lag. The upgrade has helped streamline operations, centralize solutions, and improve collaboration with data scientists across the company.

They experience reduced response times by over 50% on average and 5 times more throughput with AlloyDB.

[Read the story →](#)



**Boost productivity
with agents in
every data task**

03



It's a virtuous cycle.

Unified data enables you to harness the power of AI and gain real-time insights. And it's through the use of AI that you can unlock the full potential of your data.

Since AI is a suite of tools that can be used at many different levels and for different purposes, the potential for positive gains amplifies with each integration.

Here's how to get started using AI to accomplish your larger AI initiatives.



Accelerate your data to AI journey.

Want it to be simple for users with various levels of technical abilities to build, deploy, and manage AI models and applications? Have a look at your platform. By selecting one that brings AI closer to your data and integrates AI agent experiences into every user journey, you can improve productivity and outcomes. **Here are the best practices to bring AI agents to your data:**

Drive productivity with AI agents.

Gen AI is more than just content creation—it can transform how users interact with your data platform and accelerate their work. Integrate gen AI capabilities into your data platform to provide always-on, AI-powered assistants for your users. AI agents can assist with writing queries, migrating and managing data, creating data transformations, and building data visualizations. They can also simplify analysis by interpreting natural language questions and translating them into appropriate queries. Additionally, AI agents can provide intelligent recommendations to optimize costs and performance, and help ensure security, governance, and compliance.

Improve enterprise knowledge discovery.

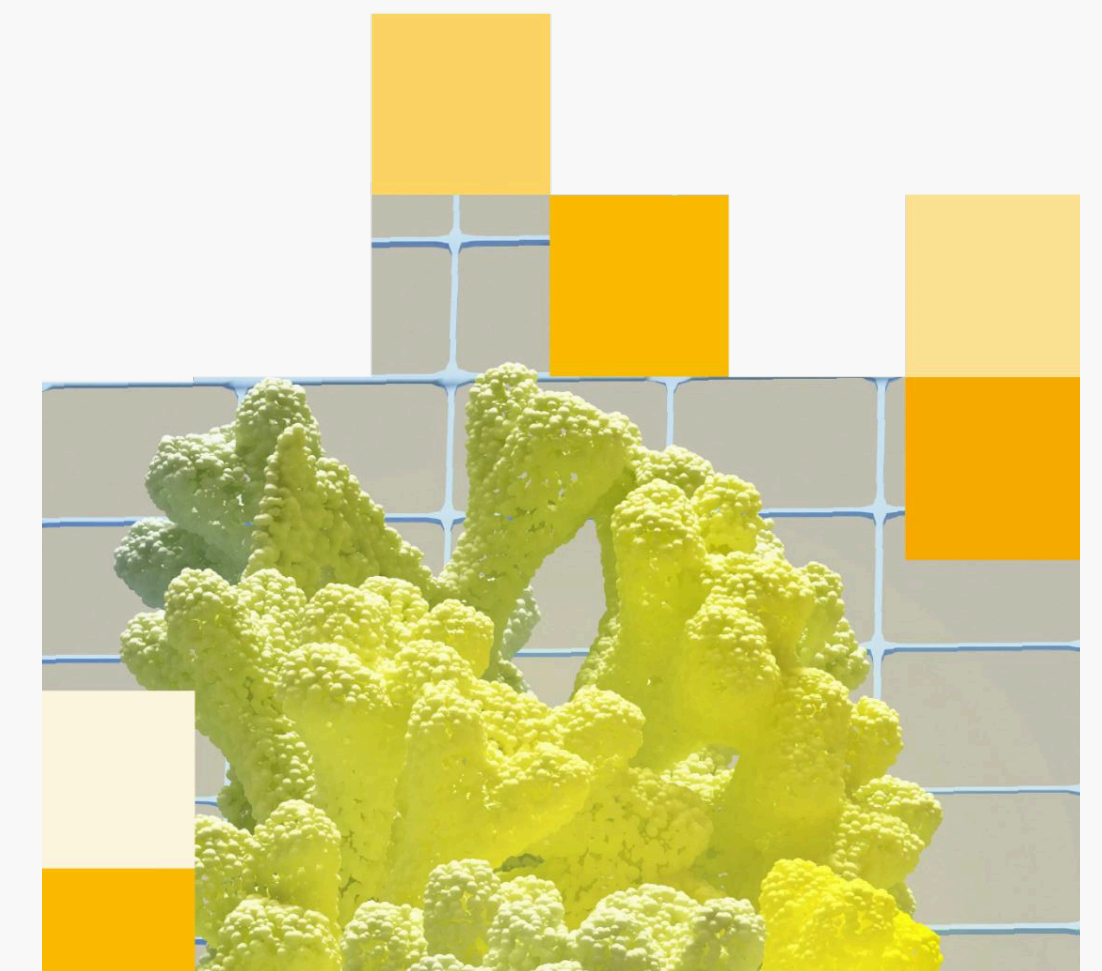
Integrate model development, tuning, and inference capabilities directly into your data platform. This eliminates the need to move data to external AI management tools, which can be time-consuming, costly, and introduce risks. Built-in gen AI capabilities such as vector embeddings and vector search allow for a deeper, more meaningful understanding of your data directly within your data platform, whether the data resides in an operational database or a data warehouse.

Simplify model development and deployment.

Building, deploying, and managing ML models can be complex and require specialized expertise. A modern data platform should simplify this process by providing easy-to-use tools and frameworks for model development and deployment. These tools can abstract away the complexities of coding languages and infrastructure management, allowing data scientists and analysts to focus on building and deploying effective models without needing deep knowledge of underlying technologies.

To empower users further, consider seeking out a data platform that offers:

- **Pre-trained models** for a variety of tasks such as speech-to-text that users can adapt to their specific needs instead of building them from scratch.
- **Tools for fine-tuning models** to specific needs and data characteristics, resulting in more accurate and relevant outcomes. Not all models are one-size-fits-all.
- **Capabilities to analyze and connect** insights from unstructured data with structured data, which opens up opportunities for semantic search and model fine-tuning. Learn more about [multimodal AI](#).







Enable conversational analytics

Conversational analytics delivers a new approach to data analysis that combines natural language processing with gen AI, enabling users to ask questions of their data in plain language and receive clear, concise answers. Fueled by semantic modeling and modern knowledge graph capabilities, analytics empower organizations to unlock deeper insights, make better decisions, and drive innovation. Conversational analytics eliminates the need for users to understand complex data structures or query languages, making data analysis accessible to everyone.

A data platform built for AI needs a semantic model built on shared definitions and figures, enabling anyone in the organization to ask detailed questions and gain accurate insights. Multi-turn capabilities allow for follow-up questions and greater specificity.

Conversational analytics offer several benefits:

-  **Increased data accessibility** empowers everyone to access and analyze data, regardless of technical expertise, leading to better decision-making.
-  **Reduced time to insights** streamlines the data analysis process, allowing users to quickly get the information they need.
-  **More agile and accurate decision-making** provides real-time access to data, enabling users to quickly identify trends and anomalies.
-  **Reduced reliance on centralized IT and data analysts** frees up their time to focus on other strategic initiatives by enabling business users to self-serve their data needs.



The Home Depot uses Google Cloud's AI and scalable infrastructure to revolutionize their business and create a new approach to home improvement.

Data is the cornerstone of The Home Depot's AI initiatives. With Vertex AI, they can leverage massive datasets to power tools like Sidekick—an app that uses computer vision to find a product's location efficiently. This dramatically reduces search times and enhances associate productivity. Using BigQuery as a robust repository for this data alongside Looker's analytical capabilities powers a comprehensive view of business performance. This data-driven approach enables The Home Depot to better understand customer needs and optimize operations across stores and logistics. The result is improved operational efficiency, better customer service, and a streamlined shopping experience.

Watch the video →



A decorative graphic on the left side of the slide. It features a cluster of test tubes with green and yellow liquid inside, set against a white background. This central white area is surrounded by a grid of green squares of varying shades, creating a stepped, staircase-like effect.

**Ensure data
security and
governance**

04

You have more data than ever before.

And keeping it secure and compliant—while extracting valuable insights—is an increasingly complex task.

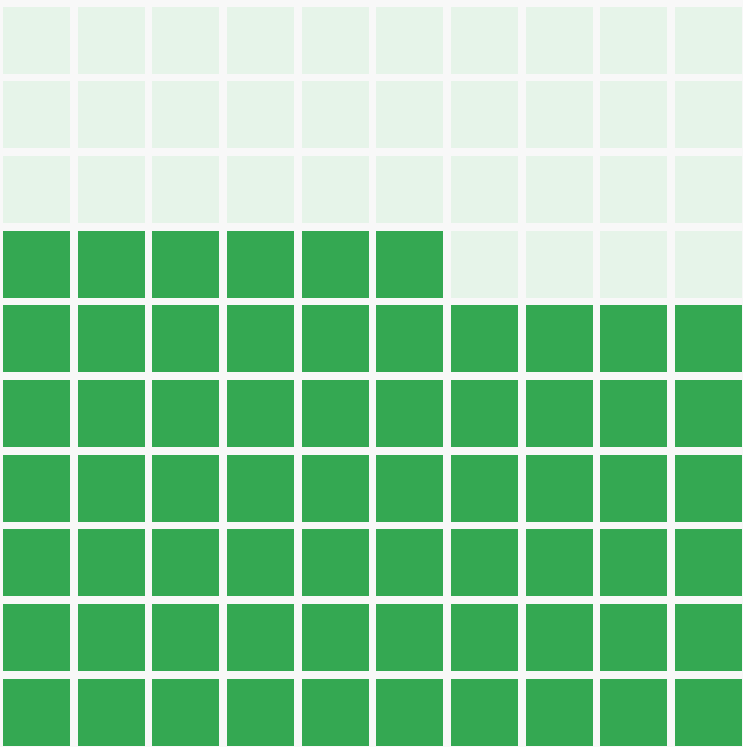
The more data you collect and analyze, the more security and governance matter.

We'll cover some best practices for governing and securing your data.



Manage and govern data, end-to-end.

Data governance has always been a necessity, but the exponential growth of data in the AI era has made achieving it more complicated. It's how you can connect high-quality data to successful AI initiatives, ensuring compliance with growing regulations, empowering data-driven decision-making, and maximizing ROI by addressing data-related challenges like dark, duplicate, and dirty data. The goal is to establish end-to-end, centralized governance and organization-wide search. Start by creating a data to AI catalog across distributed data and implementing AI-powered data profiling for end-to-end data lineage.




66%


of organizations have at least half of their data dark, posing significant risk⁶

[Learn more about data governance](#) →


Here's a detailed look at how to build a robust data and AI governance framework:

- 


Prioritize data quality to drive ROI.

To maximize ROI from AI initiatives, organizations must prioritize data quality by ensuring data is extracted from trustworthy sources, regularly updated, and complete. Implementing data observability tools can help monitor data lineage, detect anomalies, and address data quality issues proactively. High-quality data builds trust, empowers decision-making, and leads to improved user experiences and operational efficiency.
- 

Rise above architecture for compliance and efficiency.

As data regulations evolve, organizations need a proactive, end-to-end data governance strategy that transcends specific architectures. This involves implementing umbrella policies that ensure compliance across jurisdictions and technological environments.
- 

Banish dark, duplicate, and dirty data to maximize ROI.

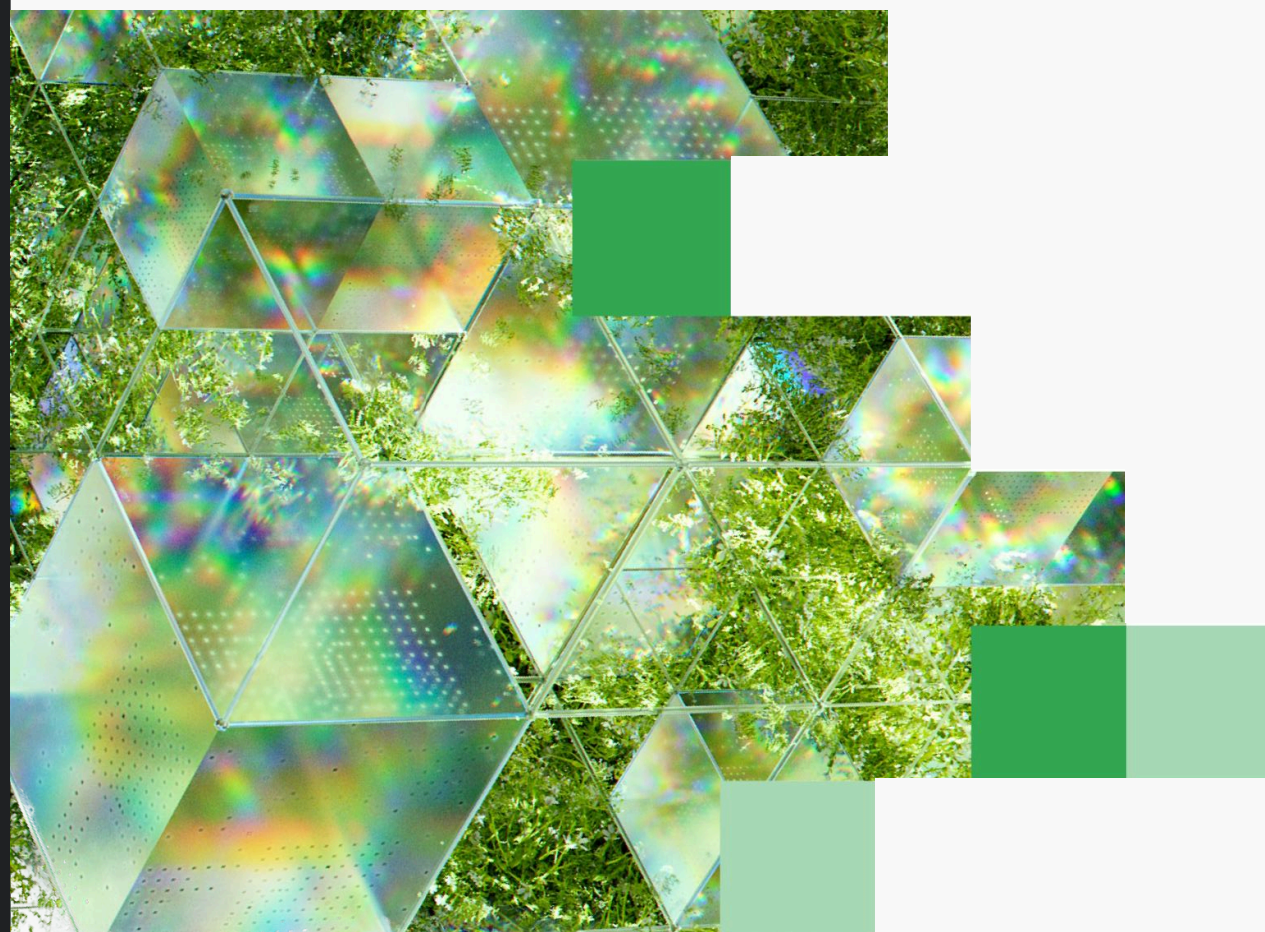
Data that has been collected but isn't in use; data that appears more than once; and data that is inaccurate, incomplete, or inconsistent must be addressed. Tools and processes are needed to identify and classify dark data, including unstructured data from various sources. Data deduplication techniques are essential for ensuring data accuracy and consistency. Implementing data quality management processes helps identify and correct errors, ensuring AI systems are trained on reliable data.
- 

Govern AI models in production.

As new regulations for AI (e.g., [EU Artificial Intelligence Act](#)) are being established, it is crucial to maintain transparency and accountability in your AI initiatives. By tracking the training data and retraining schedules, and better model explainability frameworks, you can better ensure the model's reliability, fairness, and compliance with regulations. This fosters trust and allows for continuous improvement and responsible AI deployment.

Incorporate multi-layer security.

The evolving security threat landscape presents several challenges. Mandiant, in the [2024 M-Trends Report](#), notes that attackers have followed the growing adoption of cloud, pivoting to target cloud-hosted data and leverage cloud computing resources in their operations. Building a secure data platform is of paramount importance and requires a comprehensive security strategy that encompasses multiple layers of defense.



Here are some first steps you can take to secure your data:



Implement strong access controls.

Establish a robust identity and access management system with granular control over data access. Utilize multi-factor authentication with multiple verification methods, integrate with third-party identity providers, and define governance rules for resource usage and configurations.



Protect the perimeter of your data platform.

Create a protective perimeter around your valuable data and define precisely which users or services can reach the data within your protected zone. Deploy a scalable cloud-native firewall with threat detection and prevention capabilities. Implement a web application firewall (WAF) to protect against DDoS, XSS, and SQLi attacks.



Protect sensitive data.

Scan and classify data to understand its composition and associated risks. Employ data obfuscation, de-identification, redaction, masking, tokenization, and transformation techniques. Utilize a secrets management system for secure credential storage and management. Consider customer-managed encryption keys for greater control over data encryption.



Enforce continuous monitoring and auditing.

Maintain data integrity and availability through continuous security monitoring, anomaly detection systems, and data backup and recovery mechanisms. Implement data validation and reconciliation processes. Collect and analyze logs in a SIEM system for real-time insights into system activities and data access patterns. Ensure compliance with data privacy regulations through appropriate security controls and regular audits. Utilize tools that simplify compliance and security management for sensitive data and workloads in the cloud.

Learn more in our recent eBook about building a secure data platform →

Box addresses data lineage, observability, governance, and security challenges with Dataplex.

Dataplex, Google Cloud's data governance solution, enables the leading intelligent cloud content management company to streamline data discovery using metadata tags and tag templates. Additionally, Box leveraged table-level operational metadata tags and Data Lineage APIs to achieve end-to-end data observability for all BigQuery ingestion pipelines, facilitating faster impact analysis and issue resolution. Implementing a comprehensive data classification framework and integrating with Google Cloud's Data Loss Prevention (DLP) enabled Box to reliably identify, classify, and protect sensitive data fields, preventing data loss and unintended exposure.

[Read the full story →](#)



Google Cloud





Improve efficiency
to reduce data cost
and scale

05

Your data is both an asset and a liability.

While your data can power tremendous insights and ROI, it also is costing increasingly more to store.

Extracting the most value from your data requires capitalizing on its worth while simultaneously keeping a lid on its growing demands.



Secure, scalable, data and AI monetization

Data powered by AI can drive significant monetization opportunities. Fortune Business Insights projects the global data monetization market to reach \$12.62B by 2032, exhibiting a compound annual growth rate of 17.5% during the forecast period (2024-2032). Further, Gartner predicts that by 2026, more than a quarter of Fortune 500 Chief Data Analytics Officers will have become responsible for at least one data and analytics-based product that becomes a top earner. AI in the context of data products will help to unlock deeper insights, create new product offerings, and enable data accessibility and usability.



Here are the steps to monetize your data at scale:

- 01

Identify your data products.

Determine the target audience, the value proposition, and the pricing model.
- 02

Develop a data monetization strategy.

Define your goals (new revenue, product enhancement, competitive advantage) and choose your approach.

Indirect Monetization: Enhance core offerings with analytics for better engagement and retention.

Tiered SaaS Offering or Direct Monetization: Offer varying levels of data access and insights at different price points.

Bespoke Direct Monetization: Provide highly customized solutions at a premium price.
- 03

Build data products or services.

Choose the right approach based on the data's state and customer needs.

Raw data through data exchanges/data clean rooms: Facilitate secure data sharing and collaboration.

Curated data through insights enabled by AI. Offer standard and customized insights via dashboards, leveraging AI for deeper analysis

APIs: Enable seamless integration of datasets, insights, and AI models with third-party applications
- 04

Launch, evaluate, and iterate.

Continuously assess the relevance of your offerings and adapt to market demands.

The key is to identify valuable data, package it into user-friendly products, and deliver it through secure and accessible channels. By aligning your data strategy with market needs, you can unlock significant revenue potential and drive business growth.

Minimize costs with thoughtful data platform design.

Organizations are collecting and processing unprecedented volumes of data, driving the need for scalable and powerful analytics data platforms. However, the exponential growth of data also brings the challenge of escalating costs. Unchecked, these costs can strain budgets, hinder innovation, and limit the potential of AI initiatives.

Here are some best practices of optimizing data and analytics costs, so your AI initiatives deliver a higher ROI. And freeing your resources up for more innovation, too.



Automate workloads with managed services

Migrate your data and AI workloads to a cloud environment to eliminate the need to monitor, troubleshoot, update, tune, and plan your underlying infrastructure as you scale.

Minimize overhead and management costs with serverless architectures, which eliminate patching and maintenance—and are optimized for cost.

Scale up or down as needed to meet changing demands with cloud architectures, enabling you to quickly act on new opportunities without the need to plan configuration requirements, pause databases, or spin up dedicated warehouses. This eliminates or reduces the time spent on database administration, ETL management, and new schema modification.



Data ingestion and extraction optimization

Consider the right format and compression techniques for your data. Formats like Parquet and ORC are ideal choices as they offer efficient compression and are optimized for analytical queries by storing data in columns instead of rows.

Leverage native connectors to simplify the data ingestion process and reduce maintenance efforts.

Load raw data into the warehouse and perform transformations within the warehouse itself to maximize flexibility and cost-effectiveness.

Use real-time streaming only when necessary for your specific use case, as batch processing can often be a more cost-effective solution.



Storage optimization

Achieve cost-effective storage by implementing a tiered storage system that uses a combination of hot and cold storage options.

Delete or archive unnecessary data according to established data retention policies to further reduce storage costs.

Use data compression techniques to reduce your overall storage footprint.

Partition tables into smaller segments and organize data within partitions using clustering techniques to improve query performance.



Compute optimization

Analyze query patterns to understand your workloads. This will help you identify areas for improvement.

Right-size resources by selecting appropriate compute resources for different workloads. Also, consider autoscaling to enhance performance.

Use caching (which stores frequently accessed results in memory) to improve query response times and reduce compute resource usage.

Pre-compute and store results of complex queries with materialized views for faster response times.

Adhere to best practices for writing efficient SQL queries to significantly contribute to overall optimization.



FinOps

Implement cost monitoring and alerts to track spending and identify unexpected increases, allowing for proactive cost control.

Use chargeback and showback mechanisms to enable the allocation of data and analytics costs to departments based on their usage, promoting accountability and cost awareness.

Visualize costs through a dashboard to gain insights into usage patterns, helping to identify costly queries and facilitate optimization efforts.

Ford achieves reductions in routine database management tasks using Google Cloud database services.

Ford Motor Company, a global automotive leader, knew they needed to modernize their database infrastructure to keep pace with the demands of the digital age. Their existing system required extensive manual effort for routine tasks, hindering their ability to innovate and respond quickly to market changes. To overcome these challenges, Ford migrated a significant portion of their database estate to Google Cloud's managed services, including Cloud SQL and Spanner. This strategic move yielded impressive results. Not only were Ford's DBAs freed from time-consuming maintenance, allowing them to focus on higher-value projects, but they also saw a 30% improvement in database performance.

[Read the full story →](#)



Google Cloud



Inside our unique approach



Unparalleled infrastructure, data, and AI integration

Extract maximum value from your data with unparalleled speed and efficiency. With Google's Data Cloud, data teams can take advantage of end-to-end integration from infrastructure (TPUs and GPUs) to data management (AlloyDB, Spanner and BigQuery) to AI (Gemini and Vertex AI). And with AI capabilities and assistive agents integrated into our platform, you can deliver new use cases faster—from building and deploying models using simple SQL to analyzing complex data using natural language.



A unified data foundation

Google Cloud's unified data foundation is built on BigQuery and database solutions, integrating your structured and unstructured data to deliver insights and power AI-driven applications. This unified data foundation enables you to manage your entire data lifecycle—making data management, analysis and AI development easier for different types of users. It integrates end-to-end data governance with complete lineage tracking. That includes everything from data ingestion and storage to model training and deployment, so that you can have a single, governed, source of truth for your data, across transaction systems, analytics and AI.



The most open data platform for modernization

Google Cloud is committed to being the most open cloud provider, letting you build modern, data-driven applications wherever your workloads are. We support open source and open standards, and offer managed database services, such as Cloud SQL for MySQL, PostgreSQL and SQL Server, that are fully compatible with popular open-source engines and AI models. With AlloyDB Omni and BigQuery Omni, you can utilize data and modernize your applications across Google Cloud, AWS, Azure, and Google Distributed Cloud, without incurring the costs, security risks, and governance concerns associated with data migration.



Enterprise-grade AI ecosystem

Google Data Cloud is an industry leader in efficiency, security, and scale—enabling you to use advanced techniques for grounding your models in your business data. Build, deploy, customize, and manage models at scale with retrieval augmented generation (RAG), automated prompt engineering, and fine-tuning. Query business data using natural language with greater accuracy and nuance with semantic layer integration across services. Power your applications with capabilities like semantic search, recommendation systems, and anomaly detection with highly efficient vector search technology that Google users for services such as Google Search and YouTube Search.

Priceline develops an AI-enabled pipeline to power customer experiences.

Priceline, an online travel agency, ensures its customers never miss a moment by curating unique travel experiences while offering the best deals.

And they use gen AI, grounded in their organizational data, to make it happen—bolstering their analytic capabilities and setting the stage for enhanced content creation and enriched experiences.

Once Priceline grounded their large language models (LLMs) in their business data, teams could use analytics about their most popular travel destinations and products to rapidly prototype transformative experiences. Now, within hours of launching a new test, they can see the effect of AI-enabled features on customer experiences and critical business outcomes—observing in real time how customers use the new content to make decisions and find travel deals.

This end-to-end pipeline, from development to measurement, enables Priceline to pivot quickly, enhance existing content, and expedite new AI features into their pipelines.

Driving these results required Priceline to define a data driven culture, backed by a robust customer data platform developed within Google Cloud BigQuery. Leveraging their existing AI/ML platform built on Vertex AI, Priceline can now connect millions of customer data points, from hotel information and best-selling destinations to customer preferences, to enhance and personalize content.

Priceline's AI transformation began with a solid data foundation. Yours can, too.

[Read the full story →](#)

priceline

Google Cloud



Make your data cloud AI-ready

Now you know that having unified data is the foundation to successful AI initiatives.

¹ Google Cloud and National Research Group (2024). [ROI of Gen AI](#).

² McKinsey & Company (2024). [The state of AI in early 2024: Gen AI adoption spikes and starts to generate value](#).

³ MIT Management (2021). [Tapping the power of unstructured data](#).

⁴ Recent Google Cloud survey of IT Decision Makers and Influencers from organizations 500+ with AI workload use or interest

⁵ Google Cloud (2024) Customer Intelligence Data & AI Trends Research.

⁶ Forbes (2023). [Five Factors For Planning A Data Governance Strategy](#).

Are you ready to get started?

01

Take our Data & AI Strategy Assessment to discover how ready your organization is for AI-powered digital transformation and receive expert recommendations to get you there faster.

[Take the assessment →](#)

02

Migrate your data to build apps fast, leverage generative AI, and analyze data in seconds—all with Google-grade security. And get credits to get started.

[Learn more →](#)

03

Our Database Modernization Program is a new and stress-free approach to modernizing your databases to Google Cloud. Get tooling, resources, and expertise to accelerate your modernization initiatives.

[Learn more →](#)

