Business-critical data is everywhere. Structured and unstructured. Internal and external. Flowing in from partners, enterprise applications, and customer interactions, and resulting in **organizations handling more data than ever before.**

In theory, they should already be in a good position to convert the huge volumes of information they have available into actionable intelligence.

But, in reality, many enterprises face a chaotic and fragmented data environment that limits agility, increases risk, and slows innovation.

And the push to scale AI adoption is causing the cracks in their data foundation to widen.

CXOs understand this well. They know a well-architected and governed data foundation is important. And they've already made investments to address the problem as part of broader digital, AI, and cloud transformation efforts.

But critical data remains scattered across a mix of on-prem databases, cloud environments,

and legacy applications. And traditional centralized data architectures are unable to handle the scale and complexity this creates — which means achieving faster and more cost-effective outcomes at scale depends on a fundamentally different approach.

80% of enterprise data remains untapped as businesses lack the means to process it at scale.[1] This contributes to **87%** of organizations recognizing that poor data quality is a major barrier to digital transformation success.[2]

Putting the Data House in Order

Driving lasting transformation and sustainable growth depends on a new data foundation. One built on a framework that can simultaneously keep up with shifting operational demands while allowing them to quickly seize opportunities as they arise.

This means aligning data engineering with data management — bringing together cloud-based data lakes, intelligent orchestration, unified access, and robust governance to embed AI-ready, cloud-native

data pipelines. It's a shift that will help enterprises both maximize their return on new technology investments and drive smarter, more customer-centric decisions. balance between efficiency, flexibility, and security.

This POV will outline a strategic framework for data — one of the six essential steps in **digitalizing core capabilities**. We will explore:

The long-term impact of failing to modernize, and the technical hurdles that stand in the way.

The hallmarks of a modern data strategy that can accelerate speed to data insight.

Practical steps for creating a robust foundation that will underpin cloud-scalable, AI-ready data pipelines to ensure accuracy, scalability, and accessibility.

By connecting the business imperative to a clear data engineering and management strategy, this framework provides leaders with a roadmap to unlock the full potential of their core capabilities and sustain a competitive advantage in a fast-changing digital economy.

---

[1] Businesses Have Invested Deeply In Data, But They're Still Just Scratching The Surface
[2] Why tackling data issues is the key to success with digital transformation

The Long-Term Impact of an Outdated Data Infrastructure

Poor data quality is estimated to cost enterprises an average of $12.9 million a year.[3] But beyond the financial impact, **outdated data strategies also introduce deeper risks** that undermine performance and limit business agility.

ISSUE #1

An Incomplete View
Fragmented Data Ecosystems

Data often sits in isolated departmental silos and across various fragmented formats and legacy systems. Even when technically accessible, poor data discoverability means critical insights stay hidden, making it nearly impossible to get a cross-functional view of the business.

To put this in perspective, Forrester estimates employees lose an average of 2.4 hours daily trying to find the right information due to data silos, contributing to a 19% drop in productivity across the board.[4]

A modern data strategy takes advantage of unified data platforms, data virtualization, cloud scalability, and AI-enabled data pipelines to break down data silos. These systems extract, standardize, and harmonize data from different formats and protocols in real time, unlocking a holistic view of the enterprise that can enhance experiences and avoid decision paralysis or misguided strategies.

ISSUE #2

Operational Disruptions
Inefficient Processes and Missed Opportunities

Modern operations demand a steady flow of real-time, actionable insights. But inefficient, fragmented data management – across ingestion, processing, and delivery – slows workflows, increases manual effort, and introduces (avoidable) risk.

These inefficiencies come at a heavy cost, and are compounded by data decay. Half of enterprises report that data loses value within an hour, and 75% within days.[5] Poor data management limits their ability to respond in a timely manner, making it hard to capitalize on emerging opportunities.

A modern data foundation supports real-time processing with intelligent lifecycle management – spanning ingestion through to cleansing and archiving. This ensures accurate insights can be made available quickly to facilitate better business decisions, enable intelligent automation, and deliver stronger ROI on data infrastructure.

[4] The Crisis Of Fractured Organizations
[5] Enterprise Intelligence: Digital Differentiation with Decision Velocity

Inefficient Investments
A Roadblock to AI and Advanced Analytics

Layering new technology on top of a legacy data foundation simply won't work. Poor data quality, inconsistent formats, and limited scalability will undermine the reliability of insights, causing AI models to deliver bad recommendations or biased outputs.

The impact of this is twofold: delayed ROI as AI and analytics investments won't deliver on their full potential, and higher costs to retrofit the necessary data infrastructure. In both cases, this will contribute to slow innovation and increased technical debt.

Scalable, AI-ready data pipelines — built on quality data architecture frameworks and management processes — transform data into a trusted, high-value asset. This enables reliable AI outputs, accelerates speed to insight, and supports innovation at scale without increasing technical debt.

Undermined Trust
Governance and Compliance Risks

Data is both a competitive asset and a potential liability. Poorly governed data contributes to inconsistent usage, unclear ownership, and limited accountability.

Without having strong governance, compliance, and security measures in place, enterprises become exposed to more than just reputational and regulatory risk — they compromise overall agility and restrict the trusted use of data across the business.

Embedding governance, security, and compliance throughout modern data infrastructure will help align data usage with business goals and regulations. In turn, this will increase stakeholder confidence while setting the stage for success with planned digital, AI, and cloud initiatives.

Data Engineering and Management Best Practices

Every enterprise has unique operational needs, constraints, and strategic priorities. That's why it's critical to **design a data strategy that aligns with current workflows and long-term business objectives.** And one that's able to support the new digital, AI, and cloud initiatives enterprises are looking to introduce as part of their transformation roadmap.

A robust, modern architecture **lies at the heart of any successful data strategy.**
Let's look at the top options for modern enterprises:

ARCHITECTURE	Data Mesh	Data Fabric	Lakehouse	Lambda	Kappa
ORGANIZATION SIZE	**Large**, complex enterprises with multiple domains and teams	**Medium to large organizations** with complex data landscapes	**Small to medium-sized organizations** needing to manage diverse data types and various formats	**Medium to large organizations** with massive datasets requiring both real-time and batch processing	**Medium to large organizations** with massive datasets requiring both real-time and batch processing
DATA OWNERSHIP MODEL	**Decentralized,** domain-led ownership	**Centralized ownership** in a unified environment	**Centralized** data storage and management but can support a decentralized approach to processing and governance	**Centralized batch layer,** distributed speed layer	**Centralized stream** processing pipeline and a unified data store
SUITABILITY (BASED ON MATURITY AND COMPLEXITY)	**Complex**, requires organization data maturity and culture shift	**Complex**, requires integration of various data sources including legacy systems, databases, and cloud platforms	**Complex** and multi-faceted, requiring specialized technologies and orchestration of multiple layers and components	**Complex**, requires managing two separate data processing systems (batch and speed layers)	**Simple**, with a single stream processing pipeline for all data
IDEAL FOR	Organizations with complex needs that **require autonomous control**	Organizations that **rely on fast insights** for operational and strategic decision-making	Organizations dealing with **large amounts of diverse data types,** particularly those needing both data warehousing and data lake capabilities, and those leveraging AI and machine learning	Organizations that **require comprehensive analytics** and immediate insights	Organizations that need **real-time data processing and analysis**, and those handling large volumes of streaming data
BENEFIT	Faster data access and processing for vast amounts of complex data, scales analytics adoption, and eliminates data management bottlenecks	Eliminates data silos, drives data integration, and improves governance	Unifies data storage and processing, improves data management and governance, and enables faster analytics and machine learning	Handles real-time and historical data, enables complex data analysis, and offers flexibility in data processing	Simple to build and maintain, lower maintenance costs, and faster development cycles

Whichever architecture enterprises choose based on their needs, **the following best practices** will help set the stage for effective data modernization.

Implement a Blueprint for
a Strong Data Foundation

An outcome-led data strategy starts with a scalable, cloud-optimized foundation. As the volume of data and range of sources it stems from continue to grow, enterprises will need infrastructure that can not only keep up with demand but also enable the agility and long-term scalability needed to drive innovation.

Without this, rapid and informed decision making will be impossible, and any AI-led capabilities or future digital initiatives will struggle to deliver the expected results. This requires specific architectural patterns and practices that support the unique requirements of machine learning workloads at scale. **Real-time ML pipelines, MLOps integration, and feature engineering infrastructure will provide the necessary base for reproducible model training, testing, improvement, and scaling.**

By adopting a modern AI-ready architecture, enterprises can better enhance performance, enable faster time to market for data-driven initiatives, and optimize costs, driving measurable bottom-line impact and ROI.

Practical steps

✔ **Conduct a full evaluation of existing data architectures and capabilities,** and use this to form a strategic roadmap for data modernization that aligns to business objectives.

✔ **Break down data silos** by integrating data across systems, platforms, and departments using API-driven design principles and standardized protocols. A unified data ecosystem will enable seamless data flows and cross-functional collaboration that fuels innovation, acting as a foundation for creative problem-solving and competitive advantage.

✔ **Leverage multi-cloud data lakes** using object storage with compute engines for processing and cloud data warehouses for analytics workloads. Harness real-time streaming architectures feeding into transactional data lake formats for ACID compliance.

Build high-performance, cloud-native data pipelines

designed for interoperability and intelligent orchestration. Invest in automation and workflow engines with SQL-based transformation frameworks for maintainable pipelines and to streamline data ingestion, transformation, and accessibility to reduce manual effort and accelerate time to insight.

Architect with resources and outcomes in mind.

Plan data infrastructure for AI, ML, and predictive analytics to leverage advanced tooling that can extract deeper insights, drive hyper-personalization efforts, and power new services and capabilities at scale.

Unlocking the true power of your data means investing in a strong data ecosystem. Sutherland's **data engineering services**, powered by proprietary accelerators and frameworks can help remove data silos to create a single source of truth, transforming legacy data landscapes into an intelligent and unified data platform for faster, smarter decision-making and quicker implementation.
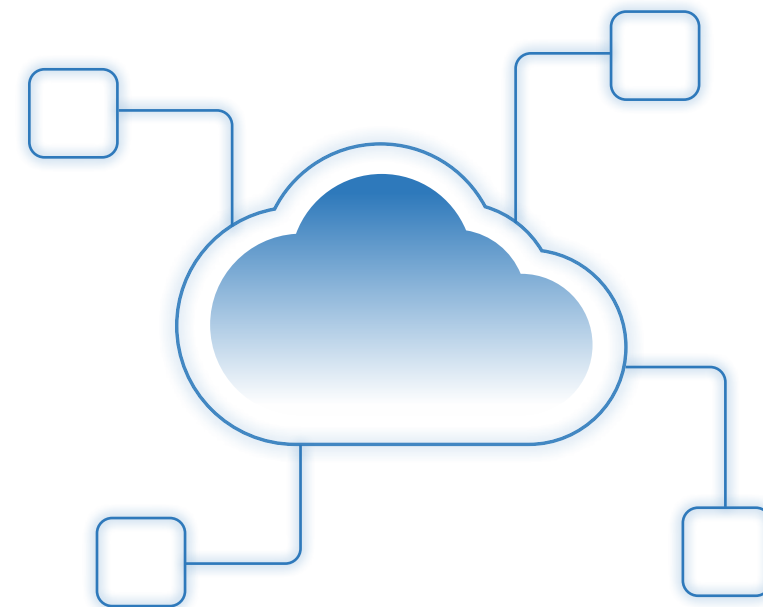
## Prioritize **Data Quality** **From the Start**

With a strong foundation in place, the next priority is ensuring quality data is fueling the systems, models, and applications that are driving business strategies and customer experiences for operational excellence.

Maintaining accuracy, consistency, and completeness of data in fast-moving, high-volume environments can be tough. That's why embedding quality controls throughout the lifecycle – from collection through to usage – is so important. Understanding what "quality" data looks like in the context of business metrics and the regulatory environment is key to this, but so is putting in place the right mechanisms to keep it that way as data moves across the enterprise.

Getting this right requires standardized management processes that will introduce uniformity in terms of how data is handled. It'll also mean ensuring the real-time data pipelines created in the foundation stage incorporate automated quality checks to catch issues early and maintain data integrity.

Adopting DataOps principles can also help improve data management to unlock its full value. Implementing DataOps practices from the beginning simplifies workflows across the entire data lifecycle, using automation for enhanced efficiencies. DataOps CI/CD (Continuous Integration/Continuous Delivery), for instance, **leverages automation and MLOps integration to streamline the creation, testing, monitoring, deployment, and orchestration of data pipelines**, ensuring rapid iteration and quality delivery through dependency management and error handling.

DataOps also focuses on breaking down silos and fostering collaboration between data scientists, engineers, analysts, and business teams, making it easier to respond rapidly to changing needs. All of this serves to reduce time to value.
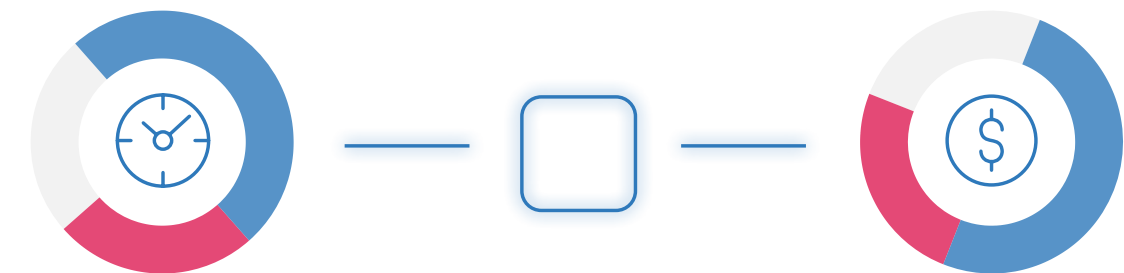
Practical steps

☑ **Standardize data entry and collection** using clear guidelines and policies to enforce consistent formats and processes across all data sources.

☑ **Deploy a data virtualization platform** to establish a single data access layer for a unified view of data from various sources.

☑ **Establish and track data quality metrics.** These will include accuracy, completeness, timeliness, and consistency to support ongoing quality measurement and remediation efforts.

☑ **Implement automated quality checks**, using advanced tools with proactive alerting and remediation to continuously profile, cleanse, and validate data throughout the pipeline and at every stage, ensuring issues can be resolved before they impact downstream systems or decision-making.

Ensuring data accuracy and consistency delivers reliable analytics and intelligence. Sutherland's **data management services** are underpinned by industry-specific data models and transformation templates to eliminate trial and error, helping you ensure the highest possible data quality to improve decision-making, cut costs, and enable new revenue streams.
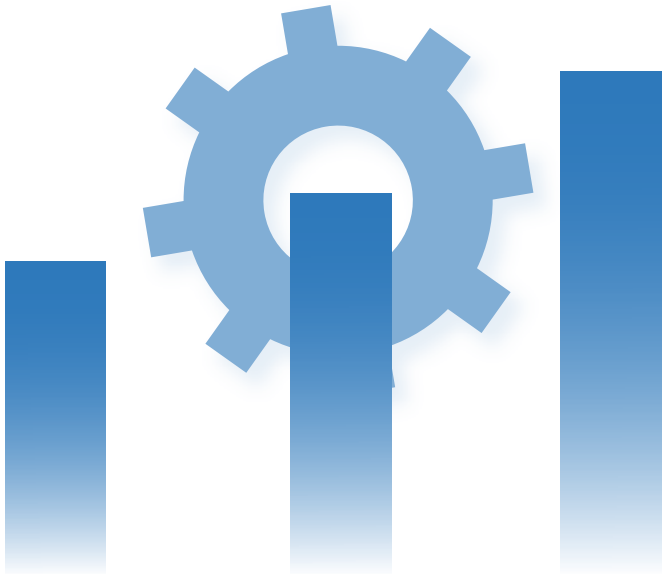
Establish a **Data Governance Framework**

Strong data governance is a critical pillar of an outcome-led data strategy. Introducing this in the early stages will help maintain data integrity and compliance, underpinning a quality data foundation to ensure consistent, reliable analytics and insights.

Traditional governance efforts need to evolve and change. This is particularly important as enterprises adopt decentralized models such as data mesh where rigid, centralized control would slow innovation and decision-making. Instead, modern governance frameworks must strike a balance: maintaining enterprise-wide standards with a level of centralized oversight while still giving domain-specific teams the autonomy to manage their own data effectively.
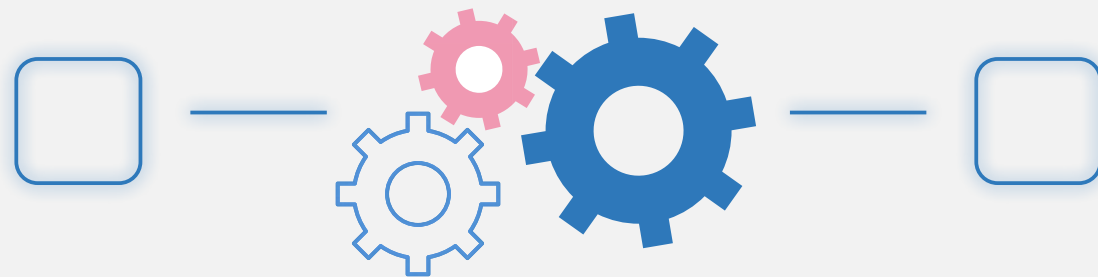
Getting this right is about embedding accountability and transparency across the entire data lifecycle. In turn, this will help maintain data integrity as it'll ensure that as data moves and gets used throughout the enterprise it does so in a controlled and auditable way.

No matter the architecture or governance framework chosen, the goal remains the same: **to create a flexible, agile, and well-governed foundation that turns raw data into a competitive advantage.**
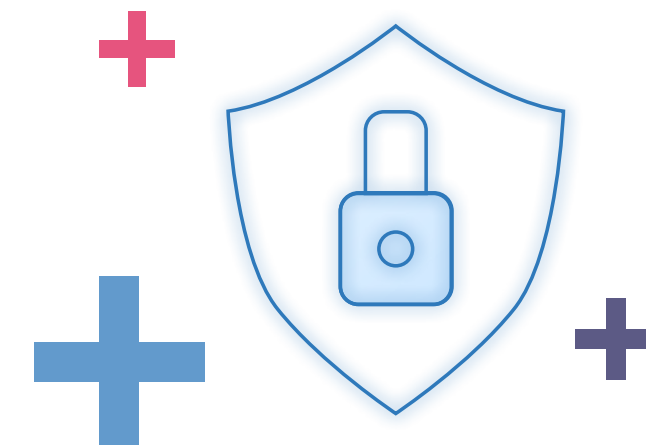
Practical steps

✔ **Adopt a governance model that blends centralized oversight with decentralized ownership.** This works particularly well as part of a data mesh model, enabling agile decision-making within different teams and domains while ensuring compliance.

✔ **Establish data ownership roles and define cross-functional committees** to oversee governance practices, maintain alignment, support standardization, and ensure consistent adoption across different business units.

✔ **Develop comprehensive policies** for data retention, usage, and lifecycle management mapped to business needs, factoring in regulatory requirements and evolving privacy expectations.

✔ **Leverage AI-powered data governance** through end-to-end testing automation that spans ETL validation, BI validation, and DQ monitoring.

Implementing a robust governance framework is imperative for ensuring data integrity and compliance. Sutherland's **data and analytics managed services** with proven methodologies can help you drive measurable, more sustainable business outcomes through the enterprise-wide use of actionable insights while minimizing compliance risk.
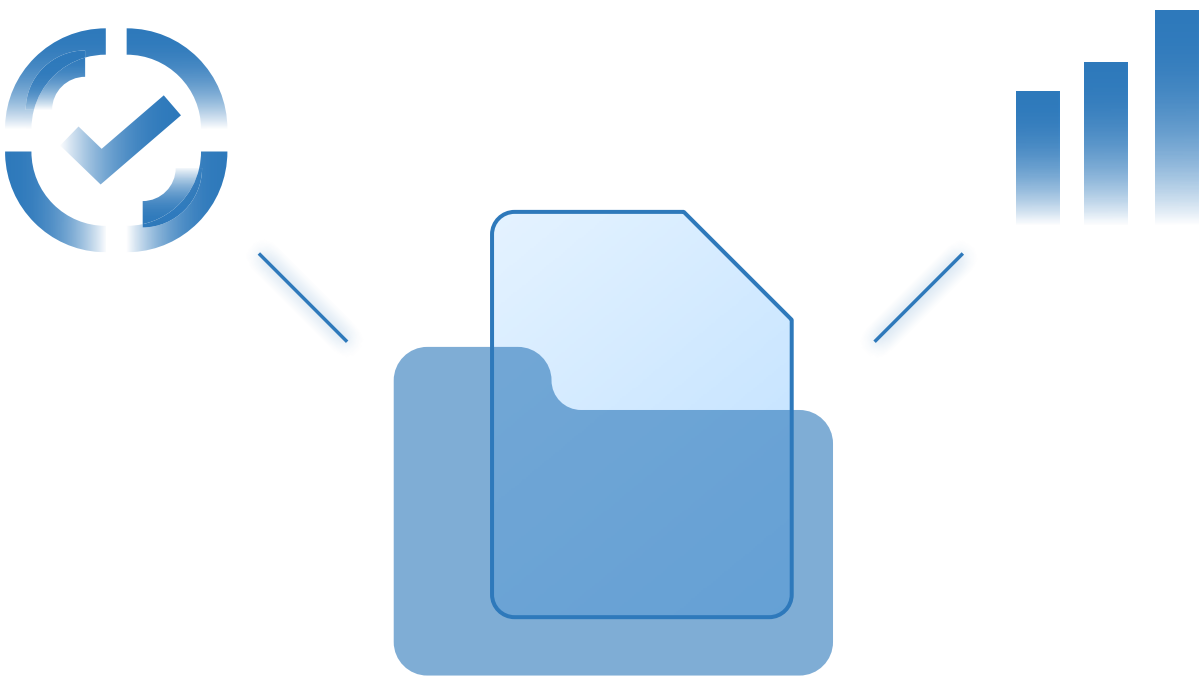
Embed Security **Throughout the Data Lifecycle**

Data security isn't just an IT concern — it's a business-critical consideration. And as enterprises scale up their use of AI and hyper-personalized customer experiences, embedding robust data security across the entire data lifecycle becomes increasingly important.

Modern security strategies that form part of data modernization efforts require a layered approach that includes encryption, role-based access controls, data anonymization techniques, and proactive AI-driven threat monitoring.

Embedding this from the ground up will allow data to be deployed strategically across the business while minimizing the risk of compromise or breach, providing the safeguards needed to meet regulatory requirements without impacting accessibility or usability.

Key to getting this right is **adopting a Zero Trust security architecture, operating on the principle of "never trust, always verify"**. This means verification happens for every data interaction, ensuring data is protected even when working in a multi-cloud or hybrid environment with automated policy enforcement across clouds and data stores.

Practical steps

Implement a Zero Trust strategy and architecture to authenticate and authorize every user, device, and application.

Introduce encryption and advanced access controls to secure data in transit and at rest. Adopt least privilege principles through an enterprise identity management platform so business users only access the data they need and for the duration that they need it.

Adopt data anonymization and masking techniques to protect sensitive information while maintaining usablility. This is critical where data is used for training AI models or advanced analytics to preserve user privacy and maintain compliance.

Continuously assess and update security measures, leveraging AI-driven threat detection tools to help identify vulnerabilities, mitigate threats, and navigate regulatory hurdles.

In the era of AI and cloud-driven transformation, ensuring the security of both data and AI systems has become critical to success. Sutherland's **comprehensive data security solutions**, designed and deployed through certified expertise and a trusted strategic partner ecosystem, safeguard your assets and AI systems against evolving threats.

Build a
Data-Driven Culture

Technology alone can't uncover the full value of a modern, outcome-led data strategy. After all, there's no point in having all this data if no one will use it. Navigating this means forging the right mindsets, skillsets, and behaviors across all levels of the enterprise so that data becomes a shared, trusted, and strategic asset.

Effective change management is critical to any digital, AI, and cloud transformation effort, where data is a core enabler of innovation at scale. Leadership must spearhead the transformation and actively empower and train employees to use data in their everyday decision-making. This will require highlighting opportunities for data-driven efficiencies and new capabilities, as well as demonstrating best practice and promoting data literacy from the boardroom to the mail room.

Taking steps to ensure data is being used consistently and correctly across the enterprise in this way will maximize the value of everything that has come before now: it'll power better business outcomes, drive the success of new digital, AI, and cloud initiatives, and help deliver customer-centric innovation.

Practical steps

Champion data ownership at every level, encouraging both business leaders and teams to fully embrace data-driven decision-making as part of day-to-day operations.

Invest in continuous training to encourage data literacy and help employees understand the value of data concepts and tools, ensuring that all teams know how to use it effectively.

Encourage self-service data collaboration, empowering teams to access, analyze, and act on data independently, helping to further break down traditional data silos.

Creating a culture of any description can be a challenge. Becoming a modern, data-driven enterprise is no different. Sutherland's **comprehensive suite of data advisory services** help integrate advanced analytics throughout your business. This includes conversational dashboards under our Insights as a Service model, delivering streamlined accessibility and scenario planning to each business unit.

Conclusion

Data engineering and data management are pivotal to any successful transformation strategy. As enterprises embrace AI, cloud, and new digital capabilities, the quality, accessibility, and governance of their data will determine how fast they can go – and how competitive they remain.

By investing in a robust, well-architected, outcome-led data foundation today, enterprises can remove barriers to innovation, enable operational agility, and set the stage for continuous transformation, while making it easier to seize new opportunities as they arise and deliver measurable business outcomes.

Ready to break down data silos, democratize insights, and solve your most complex business problems?

Discover Sutherland's Data Engineering Solutions

Artificial Intelligence. Automation. Cloud Engineering. Advanced Analytics. For Enterprises, these are key factors of success. For us, they're our core expertise.

We work with global iconic brands. We bring them a unique value proposition through market-leading technologies and business process excellence. At the heart of it all is Digital Engineering – the foundation that powers rapid innovation and scalable business transformation.

We've created over 200 unique inventions under several patents across AI and other emerging technologies. Leveraging our advanced products and platforms, we drive digital transformation at scale, optimize critical business operations, reinvent experiences and pioneer new solutions, all provided through a seamless "as-a-service" model.

For each company, we provide new keys for their businesses, the people they work with, and the customers they serve. With proven strategies and agile execution, we don't just enable change – we engineer digital outcomes.

SUTHERLAND
digital outcomes.

sutherlandglobal.com