Yalantis

DATA MODERNIZATION:

THE KEY TO GOING FROM DATA MESS TO BUSINESS GROWTH

CONTENTS

QUICK INTRO:

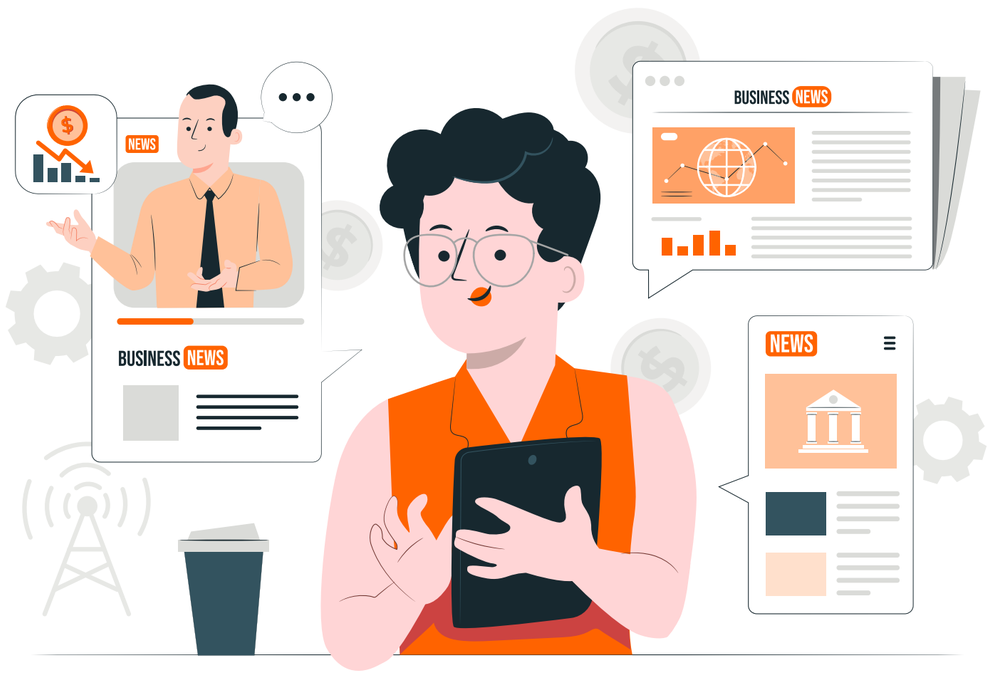
WHAT YOU'LL LEARN IN THIS WHITE PAPER
[+ BRIEF DATA MODERNIZATION Q&A]

Business data is pouring in from every corner of your operations. New customers are signing up, new sales are coming in, new shipments are going out. You know you need a modern system that can efficiently handle growing data volumes. One that makes it easy to access data, analyze it, and put it to work. But updating old processes seems like it will take ages. Where do you start? What makes a business data system "modern," anyways? How do you implement one without draining your budget? It's a lot to figure out, but this white paper is here to guide you.

Diving into data modernization is a smart move. Businesses actively invested in data modernization initiatives and reaped significant results in 2023. This trend continues in 2024 and is likely to stick around (we explain why later).

You don't want to lag behind your competitors, right? No need to worry — Yalantis market researchers and data experts have created this white paper to help you understand what data modernization entails and determine whether it will be worthwhile for your business.

BY READING THIS WHITE PAPER, YOU'LL GET INSIGHTS INTO:



- **Modernizing operations:** See how and why companies are doing it.
- **Organizational change:** Understand the connection between data modernization and company culture.
- **Data challenges:** Find out which challenge 77% of companies face that data modernization solves.
- **Efficiency gains:** Look at life with and without modern data systems.
- **Requirements:** Learn what makes a business data system modern.
- **Navigation:** Get a six-week roadmap to data modernization.

SHORT ON TIME? ANSWERS TO FOUR KEY QUESTIONS ABOUT DATA MODERNIZATION [+ LINKS TO RELEVANT SECTIONS FOR MORE INFO]

Q1: WHAT IS DATA MODERNIZATION?

Data modernization is the process of updating older data systems and infrastructures to newer, more advanced technologies. Most often, this means:

1/ moving data from legacy databases to cloud-based storage solutions[1] **(like Amazon Redshift or Microsoft Azure SQL Database)**

2/ using efficient tools (like Apache Spark) to process large datasets **for fast and detailed analysis**

3/ allowing employees to access, analyze, and visualize data on their own **with self-service analytics and business intelligence tools**[2] **(like Tableau) — no IT specialists required**

The goal is to make data more accessible, more secure, and easier to analyze.

See how a data modernization system looks in a modern enterprise in *section 3 of this white paper*.

Q2: WHY SHOULD COMPANIES MODERNIZE THEIR DATA SYSTEMS?

Modernizing your data estate paves the way for novel AI use cases, improved decision-making, and more. Companies in various industries are actively investing in data modernization initiatives (discover more in *section 1 of this white paper*) and are seeing promising results:

- AUTOMATED REPORTING
- QUICK DATA ACCESS
- ADVANCED ANALYTICS
- COST OPTIMIZATION

See benefits, use cases, and real-life case studies of *data modernization across industries*.

[1] *This is the **cloud migration** process — moving your data storage and processing from on-premises servers to a secure online platform (cloud). Think of it as transferring your data library from a physical building to a virtual one that provides easier access and greater scalability*

[2] ***Business intelligence (BI) tools** are software applications that help users visualize, analyze, and understand data. Examples include Microsoft Power BI, Amazon QuickSight, and Tableau.*

Q3: HOW DO YOU KNOW YOUR COMPANY NEEDS DATA MODERNIZATION?

Top signs it's time for a data upgrade:

- You and your teams spend hours looking for data across different departments and systems. Collaboration is weak.
- You're flying blind without a real-time view of the business and market. You miss out on crucial opportunities and don't respond quickly to market changes.
- Costly legacy system Band-Aids and high IT maintenance costs are draining your budget and preventing investment in growth and innovation.

Find the full checklist with seven signs that it's time for data modernization in *section 2*.

Q4: HOW CAN COMPANIES MODERNIZE THEIR DATA SYSTEMS?

1: Define a data strategy
Set clear goals and outline a roadmap for how you will achieve them. What data will you need to collect and analyze?

2: Migrate to the cloud
Move data from on-premises systems to a secure cloud platform like a data warehouse[3] (Amazon Redshift) or a data lake[4] (Amazon S3).

3: Integrate data
Combine databases, manual files, and external sources in one single source of truth. Prepare data for analysis.

With reliable tech support (from your in-house team or a dedicated technology partner), you can upgrade your systems within a few weeks. The Yalantis data and BI team usually **completes such projects in 5–6 weeks.** Find our roadmap *in section 5*.

*[3] A **data warehouse** is a centralized repository for storing structured data (e.g., customer records, financial transactions) to facilitate analysis. Imagine it as a well-organized data archive specifically for querying and reporting.*

*[4] A **data lake** is a central storage location for all data, regardless of its format (structured, unstructured, semi-structured). Think of it as a vast reservoir holding all your data, ready to be explored and analyzed later.*

THE DATA BOOM:

HOW AND WHY COMPANIES ARE MODERNIZING THEIR OPERATIONS

Data is the new race businesses are running, and you can't afford to fall behind. All this talk about "becoming data-driven[5]" and "driving innovation with data-driven decision-making" comes down to one thing — companies need to organize and use their data effectively to stay ahead.

What is your business's position in the data modernization race? If you haven't begun the data modernization process, you risk falling behind your competitors who are unlocking their data's full potential.

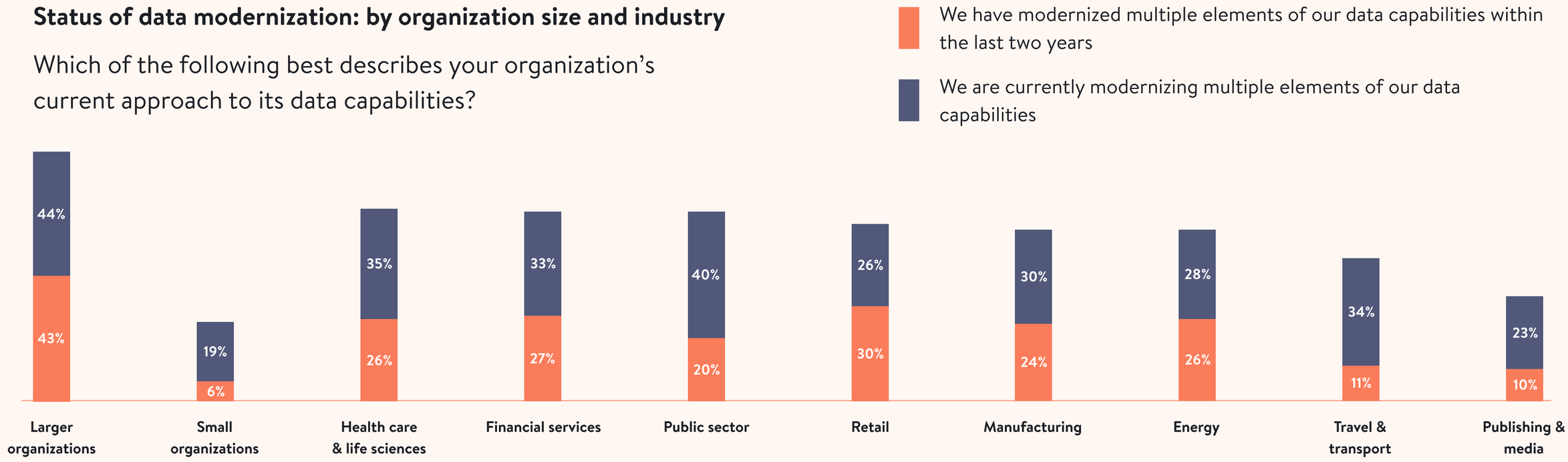
More and more businesses are actively upgrading their data operations: *one in two companies* globally has either modernized their data infrastructure in the past two years or is currently working on it. An additional 25% plan to kick off modernization efforts within the next two years.

If your data operations still rely on legacy systems and outdated processes, **it's time to prioritize modernization** or risk being outperformed by data-driven disruptors. Partnering with experienced tech partners (either on a project basis or through outstaffing) can help you speed up your transformation.

DATA MODERNIZATION INITIATIVES ARE BOOMING ACROSS FINANCIAL SERVICES, MANUFACTURING, RETAIL, AND OTHER INDUSTRIES — ESPECIALLY IN LARGER ORGANIZATIONS:

Status of data modernization: by organization size and industry

Which of the following best describes your organization's current approach to its data capabilities?



[A *survey* of 350 senior data and technology executives; MIT Technology Review Insights report, 2024]

[5] ***Data-driven*** *describes an organizational culture where data isn't just collected but also used to guide decision-making and innovate.* ***Data-driven decisions*** *are based on evidence and insights derived from data analysis, not just intuition or guesswork. Making decisions based on data is like using a map instead of wandering blindly.*

WHY 2023 WAS THE YEAR OF DATA MODERNIZATION (AND WHY THE TREND IS ONLY ACCELERATING)

In 2023, more companies became data-driven. The major reason for this was **AI[6] breakthroughs, particularly with generative AI (GenAI)[7]**. Companies see GenAI as potentially the most transformative technology in a generation — in fact, *89% of businesses* globally are increasing investment in generative AI technologies in 2024.

> ***"The arrival of Generative AI has at the very least stimulated more serious attention to the data and analytics capabilities of organizations, and how data, analytics, and AI can be combined to seriously transform key business processes."***
>
> Extract from a 2024 Wavestone ***report****
>
> *NewVantage Partners (now Wavestone) is a leading global consulting firm with European roots.

AI offers powerful capabilities to effectively handle and get insights from complex and growing datasets:

Natural language processing (NLP)[8]
allows users to ask questions in plain language to analyze data instead of writing code.

Machine learning (ML) models[9]
can be trained to automate tedious data preparation tasks like cleaning, deduplicating, and structuring data.

Generative AI
can dynamically generate data visualizations[10], analyses, and insights based on ML-prepped data and NLP queries.

[6] ***Artificial intelligence (AI)*** *is an umbrella term for technologies that enable machines to mimic human capabilities. It encompasses various techniques like machine learning or natural language processing.*

[7] ***Generative AI (GenAI)*** *is a type of AI that can generate new content like images, text, or video based on existing data patterns.*

[8] ***Natural language processing*** *is a set of techniques for computers to understand and process human language, either written or spoken. It enables tasks like analyzing customer reviews and extracting information from documents.*

[9] ***Machine learning*** *is a subset of AI that uses algorithms to learn from data and improve performance over time. This enables tasks like predicting future trends and making automated decisions.*

[10] ***Data visualization*** *is the process of transforming complex data sets into clear and visually compelling images and charts for effective communication and analysis.*

USE CASE: USING AI TECHNIQUES TO PROCESS AND ANALYZE DATA

Imagine that a retail bank wants to identify its top revenue-generating product categories year over year. They use tools like Google Cloud DataPrep or Microsoft Azure Data Factory to clean their transaction data with machine learning.

Next, they use a natural language processing interface: for example, a chatbot integrated with a GenAI model like GPT.

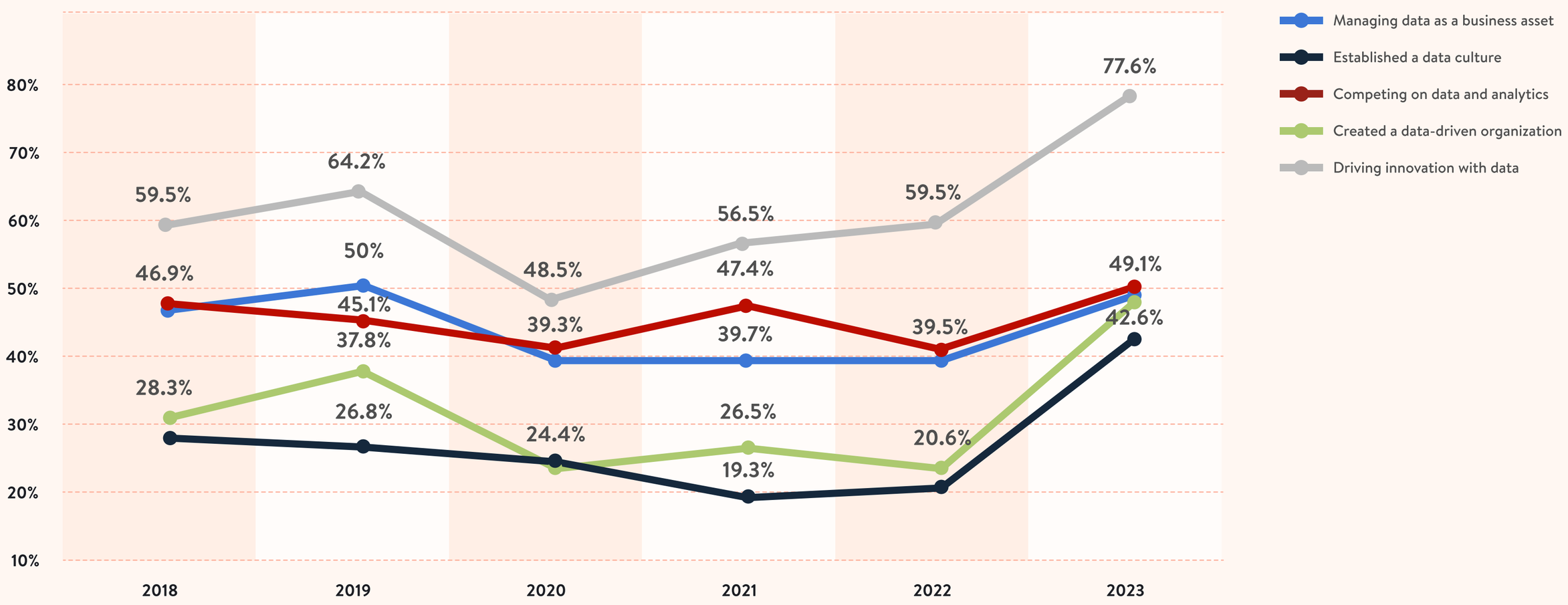
An analyst can ask the chatbot:

What were our five product categories with the highest revenue and total sales year over year?

The NLP interface understands the question, and the AI model writes code, analyzes the cleaned data, and creates easy-to-understand charts using data visualization tools like Tableau or Power BI.

In real life, using AI techniques to process and analyze data allows bank employees to quickly and easily gain insights without needing to be data experts.

Businesses want to tap into that AI power. But to do so, they must make their data infrastructure "AI-ready." This means collecting, integrating, and preparing high-quality data for AI models, and it can explain the dramatic increase in data/AI initiatives from 2018 to 2023: driving innovation with data, establishing a data culture,[11] and more:



[2024 data and AI leadership *executive survey* of over 100 diverse Fortune 1000 organizations]

[11] *Cultivating a **data culture** means making sure everyone in the company, from the C-suite to customer-facing positions, understands the value of data and knows how to use it when making decisions.*

These initiatives inherently involve modernizing how data is handled, from technical infrastructure to organizational processes and mindsets. But what do they mean in real life?

TOP DATA-RELATED INITIATIVES IN PRACTICE (BASED ON A REAL-LIFE PROJECT)

The above statistics show how businesses worked with data in 2023. Let's break down these initiatives and see how they can help you make quick, effective business decisions. We'll use an example of a US food manufacturing company that modernized their data systems and adopted these initiatives:

DRIVING INNOVATION WITH DATA

Meaning: Analyzing data to spot emerging trends, customer needs, and inefficiencies. *To do this, you need modern data platforms and tools to quickly analyze large volumes of data.*

Real-life example: The food manufacturer had been selling a popular dip for five years and wanted to analyze sales trends. They used historical sales data stored in their cloud database to uncover seasonal peaks and customer preferences. This way, they could innovate with product variations and targeted marketing campaigns.

MANAGING DATA AS A BUSINESS ASSET

Meaning: Treating data as a valuable asset by cataloging it, organizing it, and making sure it's easily accessible to the people who need it. Updating data infrastructure[12] and processes enables just that.

Real-life example: The food manufacturer had years of sales data buried in dusty reports. Without clear visibility into campaign performance, they struggled to measure success. But by centralizing all business data into a cloud database, implementing a data catalog[13], and organizing reports, the company can now easily identify successful campaigns and replicate them, leading to higher sales and improved efficiency.

[12] ***Data infrastructure*** *is the foundation of a data ecosystem. It includes the hardware, software, and networks that store, manage, and process data.*

[13] *A* ***Data catalog*** *is an organized list of all your data assets, including descriptions, location, and usage details. Think of it as a detailed library for your data, making it easier to find what you need.*

ESTABLISHING A DATA CULTURE

Meaning: Making sure **everyone in the company, from the C-suite to customer-facing positions, understands the value of data and knows how to use it**. To make this happen, data shouldn't be limited to a few IT experts. Every employee must get access to relevant data and the tools to use it.

Real-life example: The food manufacturer provided self-service dashboards to its sales and marketing teams. These dashboards show real-time sales data, competitor insights, and past performance. Now, teams can analyze past promotion data to decide whether investing $200K into a new campaign is worth it.

COMPETING ON DATA AND ANALYTICS

Meaning: Using data and analytics to gain a competitive edge in the market. Competing on data and analytics involves **using data to provide smarter offerings, operations, and go-to-market strategies that rivals can't easily replicate**.

Real-life example: Using advanced analytics[14], the food manufacturer can forecast demand for a new line of appetizers, optimizing production and distribution. This precision helps them stay ahead of competitors who might face stockouts or overproduction.

CREATING A DATA-DRIVEN ORGANIZATION

Meaning: Data literacy[15] is high, and **stakeholders from the C-suite to customer-facing positions understand how to interpret data, run analyses, build dashboards**, etc. To enable this, there should be a cultural shift to viewing data as a critical asset. Employees need tools to easily access data and visualize it.

Real-life example: The food manufacturer's regional managers can now build custom dashboards that analyze sales performance, demographic data, and social media mentions for their specific markets. This lets them create product offerings, promotions, and marketing strategies tailored to local tastes.

[14] ***Advanced analytics*** *are a collection of sophisticated techniques and tools (machine learning algorithms, statistical modeling, data mining, etc.) used to extract deeper insights from data, going beyond basic reporting and descriptive statistics.*

[15] ***Data literacy*** *is the ability to read, work with, analyze, and understand data. This includes knowing how to find relevant data, interpret its meaning, and use it to make informed decisions.*

WHAT CAN YOU EXPECT FROM DATA MODERNIZATION INITIATIVES IN THE LONG TERM?

By putting these data / AI initiatives into action, leaders in various industries can achieve the following:

STRATEGIC BUSINESS RESULTS	*EXAMPLES*
Confident, data-backed investment decisions	A food manufacturing CEO uses granular demand forecasting to decide if launching a new product line is worth the investment
Accurate budgeting and financial planning	A bank CFO uses AI-driven predictive analytics on revenue trends to set realistic growth targets
Deep understanding of customers to improve products and experiences	A retail company CMO analyzes unified customer data to tailor marketing, promotions, and new offerings based on precise buyer personas
Ability to quickly capitalize on emerging market opportunities	A logistics company COO optimizes delivery routes and scales capacity using up-to-the-minute demand predictions
Agile operations that can quickly respond to changes	A logistics company Operations Manager optimizes delivery routes and scales capacity using real-time demand predictions
Eliminating inefficiencies and streamlining processes	A manufacturing Plant Manager uses production data to identify and remove bottlenecks, avoid excess waste, and lower the defect rate

2023 WAS THE YEAR OF DATA MODERNIZATION. WILL THIS TREND KEEP UP IN 2024 AND BEYOND?

Most likely it will because:

1. AI technology will become more accessible and powerful. Consider this: In his book The Coming Wave, *DeepMind co-founder Mustafa Suleyman argues* that AI will soon "complete a calculation in seconds that would have taken a conventional computer 10,000 years."

2. Data volume and complexity will only grow. Today, *80% of all new data is unstructured*[16], meaning that it is not easily captured or quantified.

3. Timely and accurate data is key to making quick, smart business decisions. Most tech executives globally cite *data quality (41%) and timeliness (33%)* as their weakest data capabilities. This stems from poor data governance[17], which is now a key focus for companies modernizing their data capabilities as the same research cites.

"As we make progress in advanced analytics and AI, the importance of data timeliness will continue to grow and create new opportunities to transform business processes."

Extract from the 2024 MIT Technology Review Insights ***report***

If your data operations still rely on legacy systems and outdated processes, it's time to prioritize modernization or risk being outperformed by data-driven competitors with AI-ready infrastructure.

[16] ***Unstructured data*** *doesn't follow a specific format, such as text documents, images, or videos. This type of data requires additional processing to be analyzed effectively. Examples are customer reviews and social media posts.*

[17] ***Data governance*** *is a set of guidelines to ensure your data is reliable, secure, and used responsibly. This includes aspects like data ownership, access control, quality, and security.*

DATA MODERNIZATION ACROSS INDUSTRIES:

REAL-LIFE BENEFITS, USE CASES, AND CASE STUDIES

TOP FIVE BENEFITS OF DATA MODERNIZATION

This is what companies are currently looking to achieve by modernizing their data estate:

IMPROVE DECISION-MAKING

Modern data platforms provide self-service analytics and visualization tools so teams can quickly access insights rather than waiting on IT reports.

SUPPORT AI USE CASES

Cloud platforms offer scalable compute power and services to train and deploy AI models cost-effectively for use cases like predictive maintenance, demand forecasting, and personalization. (Especially relevant for large organizations.)

IMPROVE REGULATORY COMPLIANCE

With data privacy and governance becoming stricter, having a modern data platform with built-in compliance controls, audit trails, and security capabilities is essential.

ENHANCE OPERATIONAL EFFICIENCY

Automated data processing and real-time analytics reduce manual work and enable more prompt actions based on data insights.

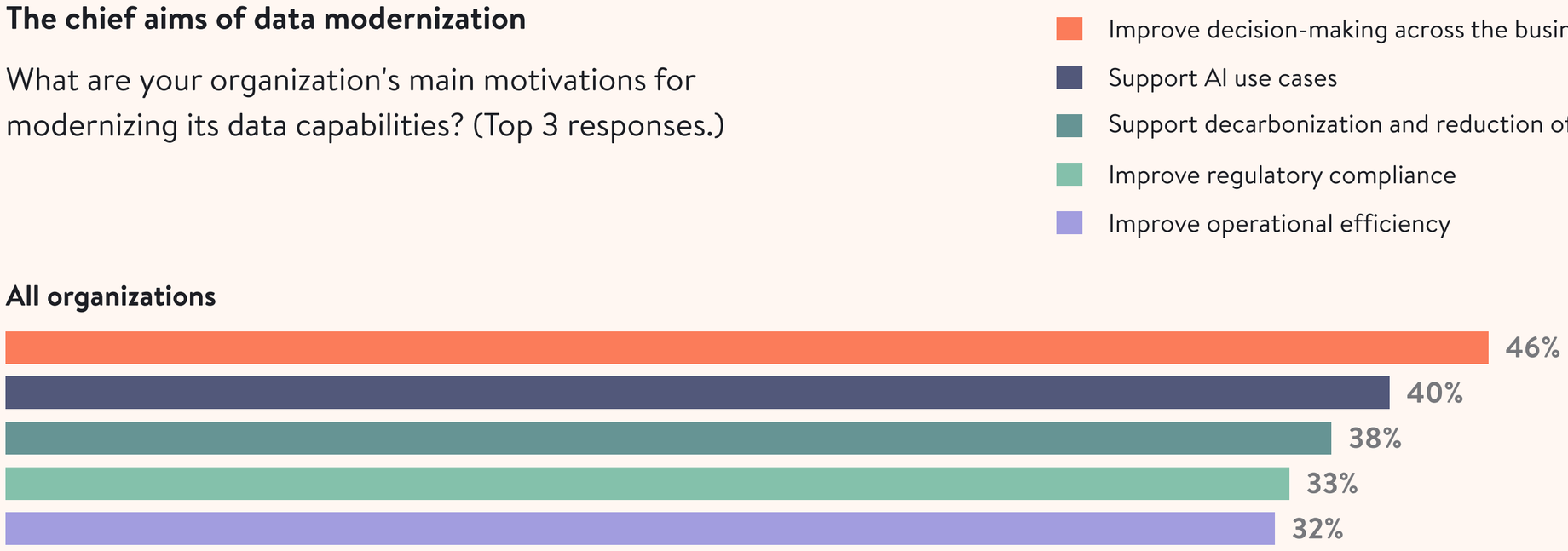
REDUCE THE ENVIRONMENTAL FOOTPRINT

Migrating to more efficient cloud data centers can significantly reduce energy consumption and the carbon footprint compared to running on-premises data centers.

Research by MIT Technology Review Insights highlights these five top goals of data modernization across industries including FinTech, healthcare, and manufacturing:

The chief aims of data modernization

What are your organization's main motivations for modernizing its data capabilities? (Top 3 responses.)



[A *survey* of 350 senior data and technology executives; MIT Technology Review Insights report, 2024]

Besides, the **chief aims of data modernization vary with an organization's size**. Smaller organizations prioritize improved decision-making (55%) over supporting AI use cases (23%). In contrast, larger organizations focus on supporting AI development (58%) and reducing their environmental footprint (54%), with improved decision-making (36%) being less of a priority.

What does this data tell us?

Companies should align data modernization with core goals, regardless of size. Smaller firms can focus on improving decision-making, as their competitive edge depends on swift, data-driven choices. Larger enterprises can prioritize AI adoption and sustainability, using their scale to invest in advanced technologies and lead on environmental responsibility.

Modernize data not for its own sake but to achieve specific objectives that propel your business forward.

TOP DATA MODERNIZATION USE CASES ACROSS FINTECH, RETAIL, AND MORE

A modern data system is a solid foundation to drive digital transformation initiatives across industries, providing plenty of use cases:

FINTECH

Fraud detection and risk management, personalized financial products and services, credit scoring and loan approvals, market analysis and investment strategies

REAL ESTATE

Property valuation and pricing optimization, tenant screening and risk assessment, predictive maintenance and facility management, market analysis and investment decisions

LOGISTICS

Route optimization and delivery efficiency, vehicle predictive maintenance and fleet management, inventory management and demand forecasting, supply chain visibility and risk management

MANUFACTURING

Machinery predictive maintenance and quality control, production optimization and process improvement, demand forecasting and inventory management, product development and innovation

HEALTHCARE

Personalized treatment plans, disease prediction and early intervention, clinical research and drug development, fraud detection and healthcare cost management

RETAIL

Personalized product recommendations and promotions, demand forecasting for optimized inventory planning and supply chain operations, analysis of market basket data, optimization of pricing strategies through consumer demand modeling, customer churn prediction

UTILITIES

Demand forecasting for better resource planning and load balancing, identification of energy consumption patterns, automated meter data analytics for usage monitoring and billing, vegetation management using geospatial data for safety and outage prevention, predictive maintenance of power plants, transmission lines, etc.

READY, SET, ACTION: THREE CASE STUDIES FROM BANKING, MANUFACTURING, AND IT

See how three companies successfully modernized their systems with the help of the Yalantis data and BI team.

A TOP US ONLINE BANK MODERNIZED THEIR DATA SYSTEMS TO EFFICIENTLY MANAGE DATA WORKFLOWS DURING RAPID GROWTH

- ***What Yalantis did:*** Helped the client create and implement a future-proof data strategy. The aim was to upgrade the current data setup for handling more data and to establish strong data governance for effectively managing data activities.
- ***Results:***
  - Increased the user base by 15% due to targeted marketing campaigns and proactive customer engagement
  - Reduced fraud-related operating losses by over 50%, preventing potential compliance risks
  - Minimized risks around data inconsistencies, errors, and inaccessibility

See full case study

A FOOD MANUFACTURER AND RETAILER BUILT A SCALABLE FOUNDATION TO PAVE THE WAY FOR ADVANCED AI/ML TECHNOLOGIES

- ***What Yalantis did:*** Helped the client migrate data from countless manually processed files to a cloud-based data warehouse. Our team established a full process for managing, organizing, and securing data through its entire lifecycle, from collection to deletion.
- ***Results:***
  - Quick insights into business trends and changes from auto-refreshed reports and dashboards
  - Optimized costs due to accurate historical data analysis
  - Advanced analytics enabled by a solid data foundation

See full case study

YALANTIS ACHIEVED 40% FASTER DECISION-MAKING WITH AN ERP & BI SYSTEM

- ***What Yalantis did:*** The Yalantis data and BI team developed an in-house ERP system with business intelligence capabilities that:
  - Automatically retrieves data from internal systems (ERP, CEP, FinOps, HRM)
  - Visualizes this data in a BI tool for easy insights
  - Provides quick data access and cross-departmental visibility into operational and HRM processes
- ***Results:***
  - 40% faster decision-making with quick insights from data
  - Single view of all company operations
  - Reduced data overload and improved data quality
  - Management and authorized teams can quickly access reliable data

See full case study

Data modernization initiatives brought transformative results for these three companies: quick insights into valid business trends and changes, decreased compliance risks, an AI foundation, and more. But is data modernization the right approach for your company? The checklist below will help you decide.

TOP SEVEN SIGNS IT'S TIME FOR A DATA UPGRADE

- ☐ You and your teams spend hours looking for data across different departments and systems. Collaboration is weak.
- ☐ Your current systems are showing their age, creaking under the weight of modern digital demands. Integrating new tools or technologies like AI feels impossible.
- ☐ You rely on inaccurate, inconsistent data full of duplicates and errors. This leads to bad decisions that impact revenue, customer churn, and operational efficiency.
- ☐ Your data comes from multiple streams — Internet of Things (IoT), mobile, web — but aging data pipelines are struggling to process it all, resulting in analytics bottlenecks.
- ☐ Outdated security protocols leave gaping holes that cybercriminals could use. One breach could mean huge compliance fines and lost customer trust.
- ☐ You're flying blind without a real-time view of the business and market. You miss out on crucial opportunities and don't respond quickly to market changes.
- ☐ Costly legacy system Band-Aids and high IT maintenance costs are draining your budget and preventing you from investing in growth and innovation.

The more items you check off this list, the more urgent your company's need for data modernization. How can you achieve the same effective results as the Yalantis partners highlighted above? From a technical standpoint, one main thing unites all three case studies — a **modernized data system** as the foundation for success.

WHAT DOES A DATA MODERNIZATION SYSTEM

LOOK LIKE IN A MODERN ENTERPRISE?
BEFORE AND AFTER

Data modernization means moving away from outdated, isolated data systems towards a unified, central data platform. For example:

Data modernization could mean moving from legacy on-premises infrastructure and systems towards a cloud-based data platform that automatically collects enterprise data from key sources. Such infrastructure is scalable. This means the data system can grow with the company's increasing data needs and user base.

A centralized data platform is the foundation for a modernized data system. Yalantis data and BI experts have helped partners across industries create such systems. Through this experience, we've gained a clear understanding of the essential elements that make a modern data system successful for various departments and deliver long-term benefits to the company.

*A CENTRAL DATA HUB (CLOUD-BASED DATA LAKE/DATA WAREHOUSE/DATA MART)**

acts as a single large storage space for all your data. It can store both structured data (customer records) and unstructured data (social media posts).

*We discuss the differences between a data lake, data warehouse, and data mart later in this white paper.

DATA PIPELINES

automatically collect and clean data from various sources (databases, SaaS apps, sensors, etc.), bringing it all together in your central data hub.

STRONG DATA GOVERNANCE, SECURITY, AND PRIVACY CONTROLS

ensure your valuable data is protected with access controls, security measures, and privacy practices.

SELF-SERVICE ANALYTICS

provide business users with easy access to this centralized, trusted data through business intelligence (BI) tools with automated real-time dashboards and reporting.

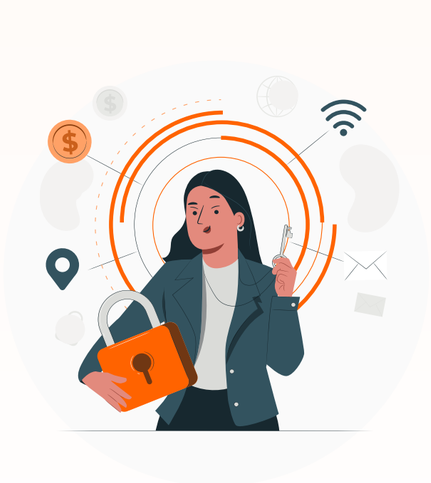
REAL-TIME DATA INTEGRATION

enables up-to-the-minute data updates for accurate and timely decision-making.

COMPREHENSIVE METADATA MANAGEMENT

maintains a detailed catalog of data sources, definitions, and usage to improve data understanding and accessibility.

The outcome? Instead of scattered, inaccessible data, you have a centralized, integrated system where everyone can get the information they need, when they need it:

Life without modernized data systems:		*Life with modernized data systems:*	
	Information is scattered across various systems and departments, making it difficult to access and analyze.		You have a centralized and readily accessible repository for all your structured and unstructured data.
	Business users outside of IT have limited access to needed data. A lack of data sharing across departments obscures the big picture of the company's performance.		A free flow of data between departments enables comprehensive analysis and better cross-functional collaboration.
	Much time is required to create reports and extract insights, leading to delayed decisions.		Key facts and figures are available at any time through web services and BI tools to quickly generate reports.
	Inconsistent data across systems raises concerns about accuracy and reliability.		Robust security measures and data quality controls ensure data integrity and compliance.

CHALLENGES AND CONCERNS

THAT DATA MODERNIZATION SOLVES

Modernizing your data systems sets the stage for a shift to a data-driven culture. It addresses a key challenge — internal organizational factors like culture, people, and processes that are resistant to change.

The primary obstacle companies face on their data-driven journey isn't technology itself. It's internal factors that have persisted as a major pain point for years:

Principal Challenge to Becoming Data-Driven	2018	2019	2020	2021	2022	2023	2024
Culture/People/ Process/Organization	80.9%	92.5%	90.9%	92.2%	91.9%	79.8%	**77.6%**
Technology Limitations	19.1%	7.5%	9.1%	7.8%	8.1%	20.2%	**23.4%**

[2024 data and AI leadership *executive survey* of over 100 diverse Fortune 1000 organizations]

"It's an understandable problem... becoming data-driven is about the ability of people and organizations to adapt to change. Long-established companies, which have been successful over generations or centuries, are unlikely to change overnight."

Randy Bean, an advisor to Fortune 1000 organizations on data and AI leadership, in ***Harvard Business Review***

In the meantime, data modernization helps your organization embrace a data-driven mindset and tackle these challenges:

LACK OF COMPETITIVE ADVANTAGE:

While competitors use modern data platforms for agility and innovation, your outdated systems leave you trailing behind. Say you're a retailer. You may miss market trends and drop the ball on new product launches if you lack advanced data analytics.

OPERATIONAL INEFFICIENCIES:

Working with legacy systems is like sailing while dragging an anchor — disconnected data and manual processes weigh you down. Your analysts could waste days compiling sales reports by manually combining data across multiple systems and spreadsheets.

POOR CUSTOMER EXPERIENCE:

Disconnected data doesn't give you a clear view of customers. You can't personalize marketing and recommendations due to siloed customer data, and no personalization means frustrated, dissatisfied customers who churn.

COMPLIANCE ISSUES:

Outdated systems make it hard to keep up with ever-changing regulations and standards. This leaves you exposed to penalties and reputational damage.

LACK OF SCALABILITY:

As your business and data volume grow, your systems can hit performance bottlenecks and capacity limits. It's like trying to fit a growing family into a tiny studio apartment.

TALENT RETENTION:

Top talent doesn't want to work with prehistoric tech. You can lose your best people to companies with modern data platforms that allow them to do their best work.

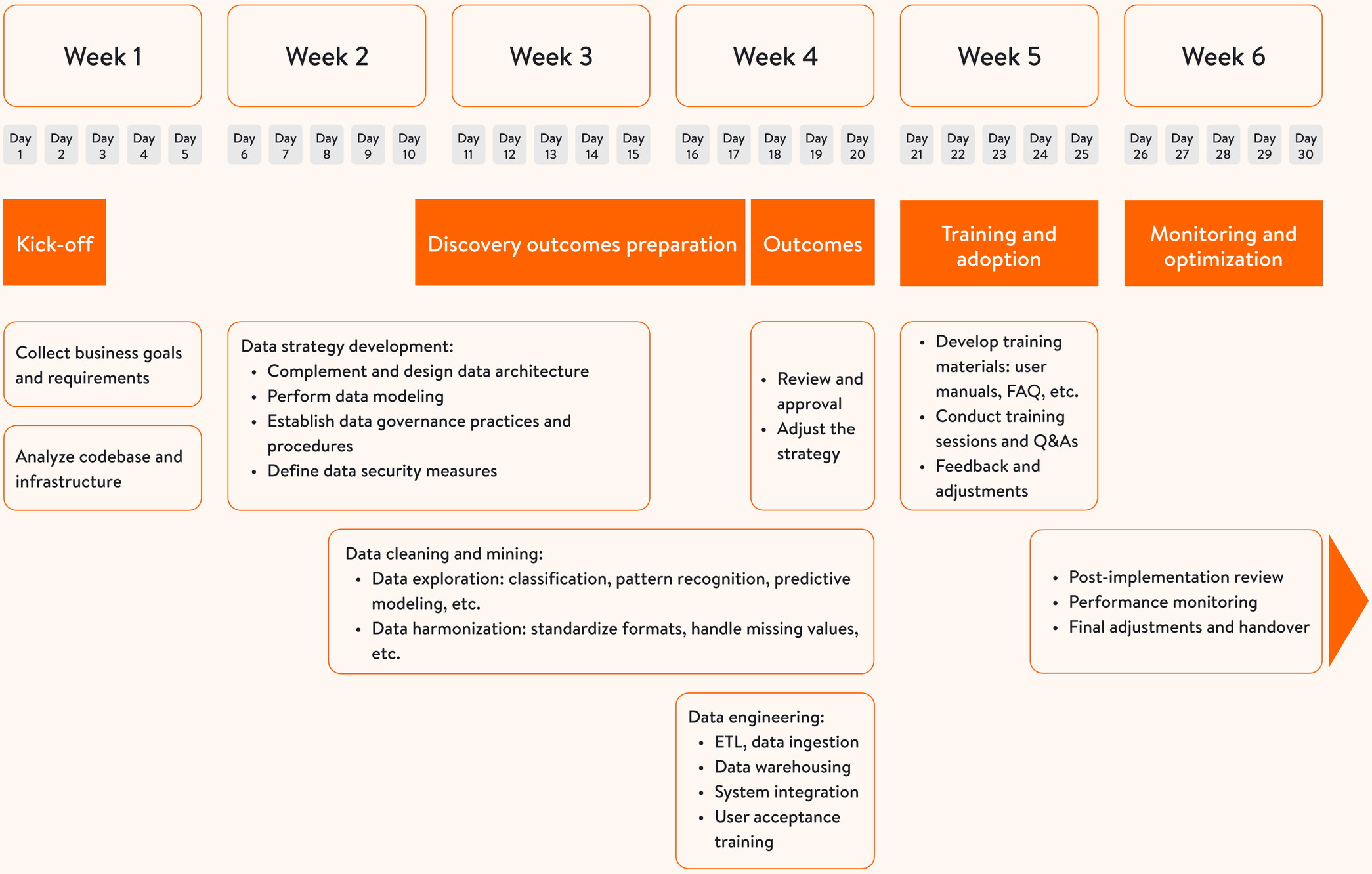
TECHNICAL DEBT:

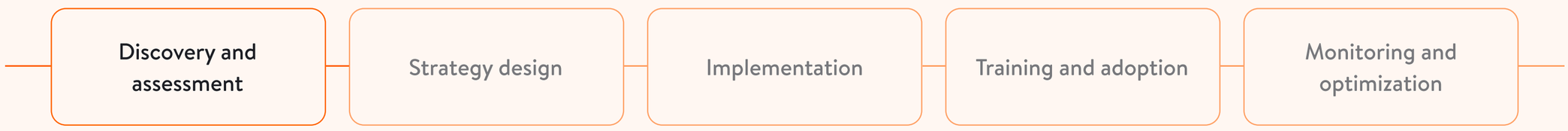
Years of quick fixes become an unstable, tangled mess. It's a ticking time bomb — sooner or later, it will all blow up.

YOUR SIX-WEEK ROADMAP

TO A MODERNIZED DATA SYSTEM

With reliable technological support (either from your in-house team or a dedicated technology partner), you can upgrade your data systems within a few weeks. The Yalantis data team usually completes such projects in 5–6 weeks:





STAGE #1: DISCOVERY AND ASSESSMENT

Stage outcome: Knowing your data inside out, including its strengths, weaknesses, and opportunities for improvement. With this thorough understanding, you can develop a modernization strategy perfectly suited to your specific needs and challenges.

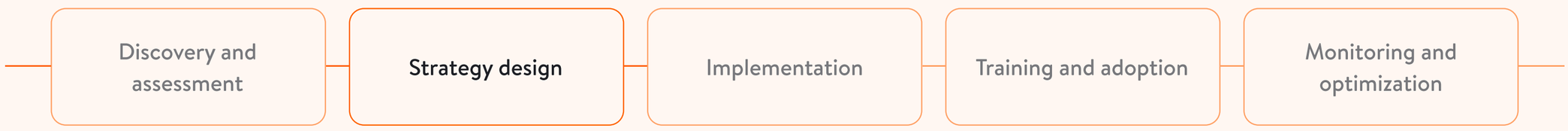
Actions to take:

COLLECT BUSINESS REQUIREMENTS

- Talk to stakeholders from different departments (sales, marketing, finance, etc.) about their current data challenges and desired capabilities.
- Map critical processes (forecasting for sales, campaign planning for marketing) and specify analytics needs such as reports, dashboards, and models.
- Define KPIs tied to business outcomes: improved sales forecast accuracy by X%, reduced regulatory fines from better data controls, etc.

ANALYZE CODEBASE AND INFRASTRUCTURE

- Inventory all data sources: databases, apps, documents, etc.
- Analyze the types of data you have: structured (customer records), semi-structured (emails), unstructured (social media posts).
- If a mix of legacy and modern systems exists, evaluate existing components to retain, integrate, or replace.
- Identify data silos restricting cross-departmental/system access.
- Assess data security, governance, compliance posture, and gaps.



STAGE #2: DATA STRATEGY AND GOVERNANCE

Stage outcome: A well-designed target data architecture blueprint and a robust data governance framework to ensure data quality, security, and compliance.

Actions to take:

DESIGN A DATA ARCHITECTURE

- Map out future state components including data lakes, data warehouses, data pipelines, and analytics tools required to support identified business needs.
- Engage cross-functional teams spanning data engineers[18], analysts, architects and business SMEs to provide input.

PERFORM DATA MODELING

- With the help of *data modeling* tools (ER/Studio, Erwin, LucidChart, Toad Data Modeler), design database schemas and dimensional data models.
- Define logical data layers and data marts that determine how information will be structured and accessed in the new repository.

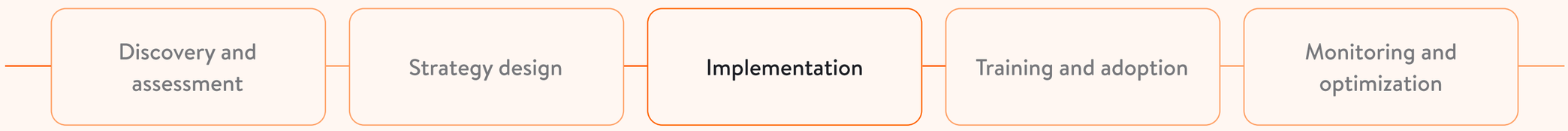
DEFINE MEASURES TO ENSURE DATA GOVERNANCE AND SECURITY

- Develop policies, standards, and processes for ensuring data quality, security, privacy, and compliance across the new platform.
- Define data ownership models with clear roles and responsibilities for data stewardship.
- Set up data quality gates and validation rules to ensure accuracy and integrity as data flows through the platform.

[18] ***Data engineering*** *is the process of designing, building, and maintaining systems and pipelines that collect, transform, store, and analyze data.*

The strategy stage involves choosing the optimal data repository (or combining different repository types — see how in the implementation stage). Here's how the three main types of repositories compare:

DATA REPOSITORY	*DATA LAKE*	*DATA WAREHOUSE*	*DATA MART*
Supported data types	Structured, semi-structured, unstructured, binary	Structured data	Highly structured data
Purpose	Big data analytics, advanced analytics, data discovery, data storage, data archiving	Historical data analysis, BI reporting and data visualization	Department-focused data analysis
Data quality	Raw data (not ready for use)	Curated data (ready for use)	Highly curated data
Data sources	Relational and non-relational databases, IoT devices, wearables, social media, emails, images, etc.	Relational databases, transactional systems, internal and external corporate systems	Data warehouse, relational databases, external and internal corporate systems



STAGE #3: IMPLEMENTATION

Stage outcome: A modern, integrated data platform built on the designed architecture, with data pipelines to ingest, transform, and serve high-quality data optimized for analytical workloads.

Actions to take:

EXPLORE AND HARMONIZE DATA

- Conduct in-depth analysis of existing data sources across databases, apps, documents, etc.
- Identify gaps, duplications, and discrepancies in data that need to be resolved.
- Cleanse, standardize, and map data fields to a unified data model for consistent integration.

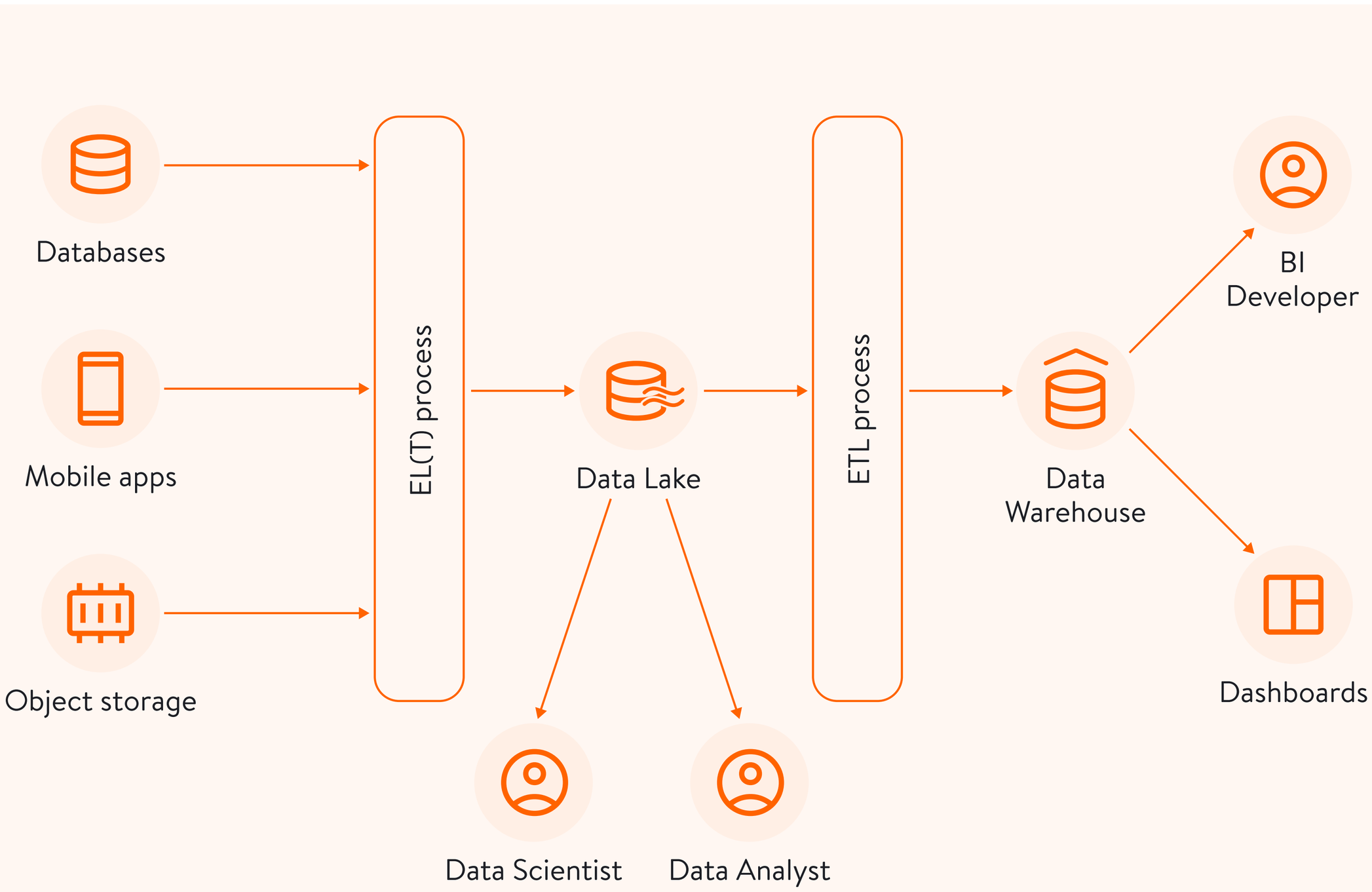
BUILD DATA PIPELINES AND DEPLOY THE DATA REPOSITORY

- Implement ETL/ELT processes and tools (e.g., Apache Airflow) to automate data flows.
- Establish connections and pipelines to ingest data from source systems into the new platform.

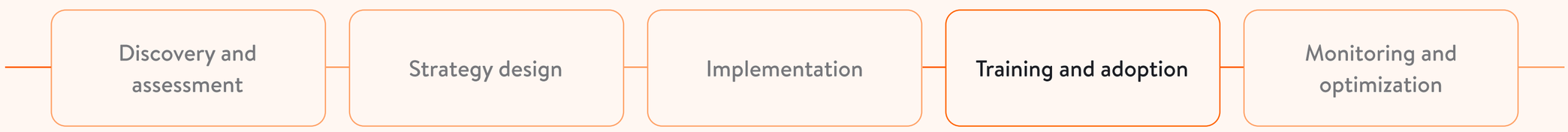
ENABLE ANALYTICS AND BI

- Integrate the data repository with visualization/BI tools for self-service analytics (Microsoft Power BI, Tableau, AWS QuickSight).
- Build reports, dashboards, and data models tailored to different business use cases.
- Implement access controls and security layers for data consumption.

When deploying a data repository, consider a hybrid approach that combines a data lake for raw data storage with a data warehouse for structured data[19] analysis (Amazon Redshift/S3, Snowflake, Google BigQuery, etc.):



[19] ***Structured data*** *is organized in a predefined format with a clear structure, such as rows and columns in a spreadsheet (example: a customer database with columns for name, address, email address, and purchase history).*



STAGE #4: TRAINING AND ADOPTION

Stage outcome: Enabling all employees, from executives to frontline staff, to comfortably use data and analytics in their day-to-day roles to make better informed decisions that improve business performance. Data skills and usage aren't limited to technical teams.

Actions to take:

DRIVE A CULTURAL SHIFT

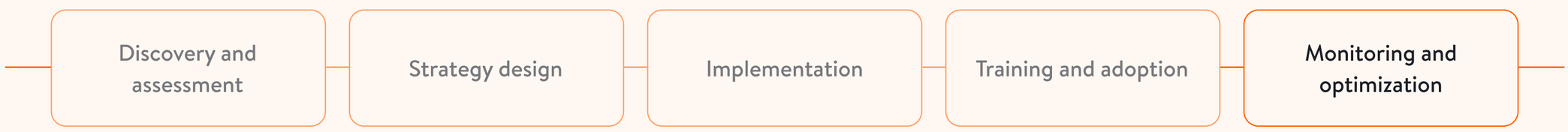
- Make data-backed insights a requirement for major decisions in meetings and planning sessions.
- Share positive impact stories widely to highlight how data-driven choices lead to better outcomes.
- Tie incentives and KPIs to effective use of data and analytics capabilities.

UPSKILL TEAMS

- Assess existing skills and identify gaps across data roles such as data engineer, data analyst, and data scientist.
- Develop comprehensive training programs tailored to varied proficiency levels.
- Consider partnering with external training providers or data consulting firms.
- Explore hiring or outsourcing options to bring in advanced data expertise.

ESTABLISH CENTERS OF EXCELLENCE (COE)

- Create physical or virtual CoEs staffed with data professionals and subject matter experts.
- Define a CoE's charter to enable best practice sharing, training, and support services.
- Implement knowledge repositories, discussion forums, and mentorship programs.
- Promote CoEs as hubs for data-related expertise, resources, and collaboration.



STAGE #5: MONITORING AND OPTIMIZATION

Stage outcome: Ensuring your data infrastructure adapts to changing needs and delivers ongoing value. Constantly measuring, iterating, and exploring new possibilities.

Actions to take:

MONITOR AND OPTIMIZE

- Implement dashboards and reporting to track metrics and KPIs in real time.
- Regularly review defined KPIs and success metrics with stakeholders.
- Gather feedback from users on pain points, areas for improvement, or new requirements.
- Based on metrics and feedback, iterate the data platform, pipelines, and models as needed.

FOSTER DATA INNOVATION

- Encourage ideation sessions to explore innovative use cases for leveraging data assets.
- Evaluate emerging trends, technologies like AI/ML, and their potential data applications.
- Run proof-of-concept pilots to test-drive high-impact data products and solutions.
- Promote a culture of continuous learning and experimentation with data.

By going through each of these stages step by step, businesses can fully tap into the game-changing benefits of data modernization — whether it's speeding up innovation, making operations more agile, cutting costs, or reducing risks.

Is your business ready to uncover the benefits?

WHAT'S NEXT: CHECK OUT THESE RESOURCES

Explore these additional resources to dive deeper into data modernization.

BLOG

Comparison of top data repositories: data mart, data warehouse, and data lake

Pick the perfect data home. Compare top options to store your business data.

BLOG

How to build an enterprise data warehouse

Build your data powerhouse. Dive into this guide for creating an enterprise data warehouse.

BLOG

Major data-related challenges your enterprise has probably faced

Data got you down? Explore common challenges and discover fixes.

RESEARCH

Modernizing data with strategic purpose

MIT reveals: Learn how leading companies are modernizing their data in this 2024 report.

RESEARCH

Data and AI executive leadership survey

How does your data strategy compare? Check out this AI and data leadership survey.

ARTICLE

Survey: GenAI is making companies more data-oriented

Been there, tried that: Is your data-driven culture stuck in neutral? This survey points to an (un)expected hero: generative AI.

EXPERTISE

Yalantis' digital transformation services

Need help modernizing? Explore services from a trusted tech partner.

BLOG

How to choose a reliable and scalable database

Future-proof your product. Find a reliable, scalable database for smooth operation.

DATA MODERNIZATION DICTIONARY: 21 KEY TERMS

1. Advanced analytics: A collection of sophisticated techniques and tools (machine learning algorithms, statistical modeling, data mining, etc.) used to extract deeper insights from data, going beyond basic reporting and descriptive statistics.

2. Artificial intelligence (AI): An umbrella term for technologies that enable machines to mimic human capabilities. AI encompasses various techniques like machine learning and natural language processing.

3. Generative AI: A type of AI that can generate new content like images, text, or video based on existing data patterns.

4. Machine learning (ML): A subset of AI that uses algorithms to learn from data and improve a model's performance over time. This allows for tasks such as predicting future trends and making automated decisions.

5. Natural language processing (NLP): A set of techniques that allows computers to process and understand human language, including speech and written text. This allows for tasks like analyzing customer reviews and extracting information from documents.

6. Business intelligence (BI) tools: Software applications that help users visualize, analyze, and understand data. Examples include Microsoft Power BI, Amazon QuickSight, and Tableau.

7. Cloud migration: Moving data storage and processing from on-premises servers to a secure online platform (cloud). Think of cloud migration as transferring your data library from a physical building to a virtual one for easier access and greater scalability.

8. Data culture: Cultivating a data culture means making sure everyone in the company, from C-suite executives to frontline workers, understands the value of data and knows how to use it when making decisions.

9. Data literacy: The ability to read, work with, analyze, and understand data. This includes knowing how to find relevant data, interpret it, and use it to make informed decisions.

10. Data-driven: This describes an organizational culture where data isn't just collected but is used to guide decision-making and fuel innovation.

11. Data-driven decision: A decision made based on evidence and insights derived from data analysis, not just intuition or guesswork. Making data-driven decisions is like using a map instead of wandering blindly.

12. Data infrastructure: The foundation of a data ecosystem. It includes the hardware, software, and networks to store, manage, and process data.

13. Data engineering: The process of designing, building, and maintaining the systems and pipelines that collect, transform, store, and analyze data.

14. Data warehouse: A centralized repository for storing structured data (customer records, financial transactions) to facilitate analysis. Imagine a data warehouse as a well-organized data archive specifically for querying and reporting.

15. Data lake: A central storage location for all your data, regardless of its format (structured, unstructured, semi-structured). Think of a data lake as a vast reservoir holding all your data, ready to be explored and analyzed later.

16. Data governance: A set of guidelines to ensure your data is reliable, secure, and used responsibly. This includes aspects like data ownership, access control, security, and quality.

17. Data catalog: An organized list of all your data assets, including descriptions, location, and usage details. Think of a data catalog as a detailed library for your data, making it easier to find what you need.

18. Data visualization: The process of transforming complex data sets into clear and visually compelling images and charts for effective communication and analysis.

19. Structured data: Data that is organized in a predefined format with a clear structure, such as rows and columns in a spreadsheet (example: a customer database with columns for name, address, email address, and purchase history).

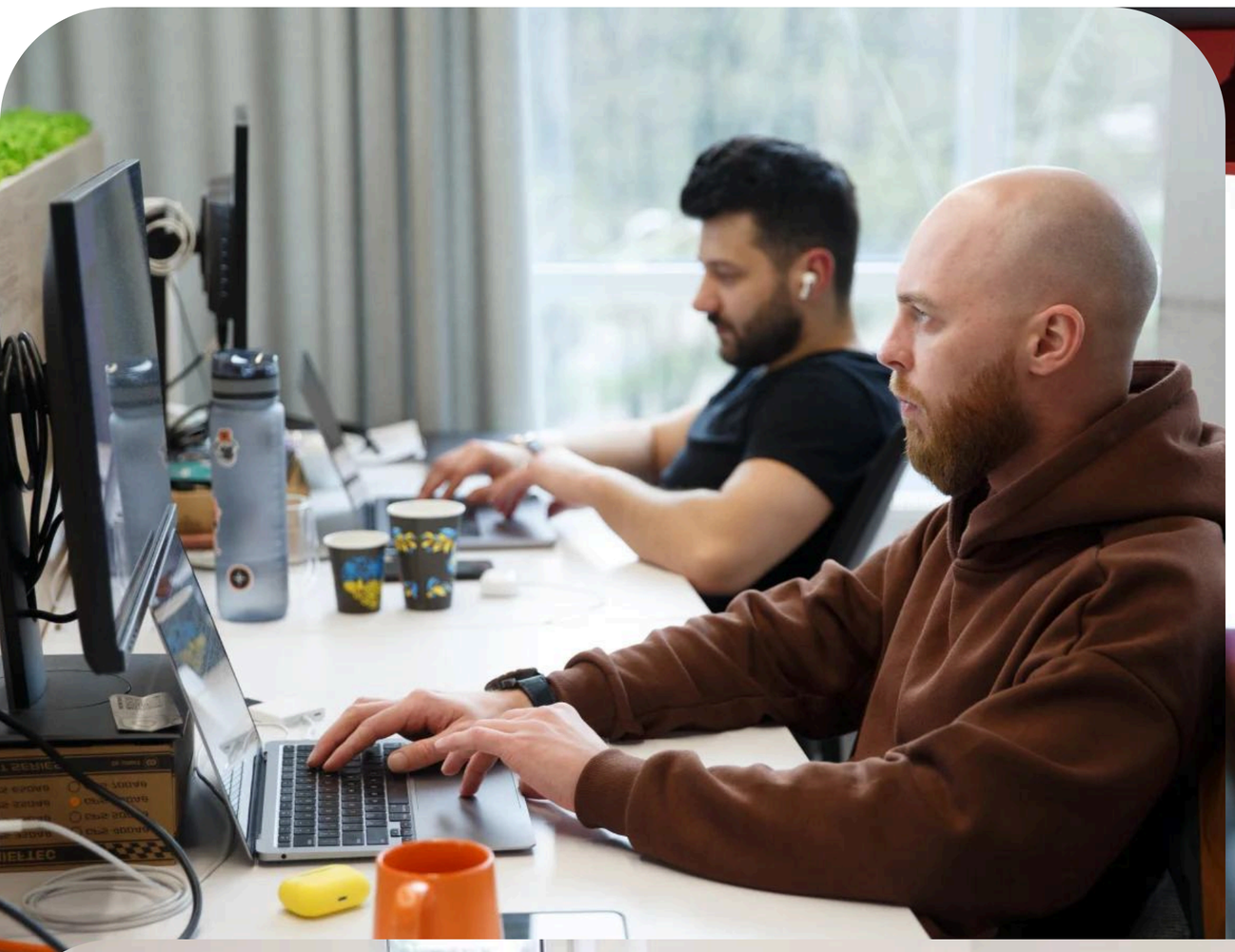
20. Semi-structured data: Data that has some internal organization but doesn’t conform to a strict tabular structure. It often uses tags or markers to define elements within the data. Examples are JSON files, log files, and emails with embedded HTML.

21. Unstructured data: Data that doesn’t follow a specific format, like text documents, images, or videos. This type of data requires additional processing to be analyzed effectively. Examples are customer reviews and social media posts.

ABOUT YALANTIS

Yalantis is a software development and IT consultancy company with over 15 years of software engineering experience. We are your strong partner in building innovative solutions.

- 5 development offices in Ukraine (Kyiv, Dnipro, Lviv), Cyprus, and Poland, plus a representative office in Estonia
- 500+ employees (+55–60% since 2021)
- Recognition as a top custom software development company in Eastern Europe according to Clutch ratings
- 4.8/5 customer satisfaction rating on Clutch
- More than 150 successfully delivered projects, with clients ranging from seed stage startups to companies with billion-dollar revenues
- 10+ partners that have been with us for more than 5 years of ongoing development



Yalantis

www.yalantis.com

hello@yalantis.com