Business Information Systems I

Christian Rossi

Academic Year 2024-2025

Abstract

This course offers a methodology to align IT design decisions with business goals. It introduces the concept of IT architecture, classifies key IT design choices, and examines how these choices impact IT architecture from both a software and infrastructure perspective. The course explores IT architecture within the manufacturing, utilities, and financial services sectors, focusing on both internal and external organizational processes along the industry value chain (e-business). It also equips students with the tools to analyze organizational requirements, with a particular emphasis on executive information systems, including the use of Key Performance Indicators.

Building on the concept of IT architecture, the course outlines a functional map of Enterprise Resource Planning systems, distinguishing between core and extended functionalities. It traces the evolution of information systems over time and highlights how ERPs have emerged through an ongoing process of functional integration. The course begins with a review of organizational theory from an information perspective, providing a framework to understand the organizational changes driven by ERP implementations. It then delves into the core functional areas of ERP systems, such as accounting and finance, operations, and management and control. The course will feature lectures and case study discussions to reinforce these concepts.

Contents

CHAPTER 1

Introduction

1.1 Definitions

Definition (*Technology*)**.** A technology represents a process that a given organization can perform, together with all the resources needed to perform the process.

Definition (*Techincal system*)**.** A technical system represents a set of machines supporting a given technology.

Definition (*Information system*)**.** An information system is a set of coordinated processes producing an information output and executing information processing activities.

Definition (*Information technology architecture*)**.** An information system is a technology, an IT architecture is a techncial system supporting a given information system.

For a long time, there has been an ongoing debate about how technical innovation influences organizations. A well-established set of beliefs links technological advancements to organizational change, shaping how companies adapt and evolve:

1. *Efficiency over effectiveness*: technological innovation primarily enhances efficiency rather than improving overall effectiveness. It streamlines processes but doesn't necessarily guarantee better decision-making or outcomes.

2. *Economies of scale*: as technology advances, businesses can scale operations more efficiently, reducing costs per unit as production increases.

3. *Larger optimal size*: the minimum viable size of an organization tends to grow with technological progress, as larger entities can better leverage new systems.

4. *Increased specialization*: automation and sophisticated systems often lead to a workforce that is more specialized, with employees focusing on narrower, highly technical roles.

5. *Tayloristic perspective*: the traditional view, inspired by Taylorism, assumes that an optimal organizational structure exists.

6. *Limited focus on group work*: early studies largely ignored the impact of technology on teamwork and collaboration, focusing instead on individual efficiency.

7. *Greater bureaucracy and formalization*: as technical systems evolve, so do organizational rules, procedures, and levels of bureaucracy, making work more structured but also more rigid.

8. *More complex management*: with increased technology comes greater managerial complexity, requiring leaders to navigate intricate systems, regulations, and workflows.

1.1.1 Information processing

Emerging in the 1960s and 1970s, the information processing perspective transformed how organizations viewed technology. As IT became widespread within businesses, it led to a fundamental shift in traditional beliefs about the impact of technical innovation. Key changes included:

- A radical shift in management principles, as technology was no longer just a tool for efficiency but a driver of decision-making and strategy.

- Unlike earlier views, IT wasn't just about automation (it processed information, the most critical resource for managerial processes). Since managerial processes shape decision, it processed information, the most critical resource for managerial processes.

- This shift created both virtuous and vicious cycles: when information systems were well-integrated, they improved decision-making, coordination, and adaptability. However, poor implementation or information overload could lead to inefficiencies, miscommunication, and bureaucratic bottlenecks.

As organizations embraced IT and information processing became central to management, three major theoretical approaches emerged: decision theory, transaction cost economics, and agency theory.

1.2 Decision theory

Galbraith's Decision Theory (1973-1977) is based on the idea that organizations function as open systems, constantly interacting with their environment. A key challenge they face is uncertainty, which defines the conditions in which they operate and reflects their ability to predict market demand. Several factors contribute to uncertainty, including market dynamism, the number of suppliers, variations in market requirements, and the level of innovation.

1.2.1 Bounded rationality

Bounded rationality refers to the cognitive limitations of individuals in processing information. Since no single person can handle all the necessary data for decision-making, cooperation becomes essential. Through cooperation, individuals and organizational units develop specialized roles, which, in turn, create interdependencies in information flow. To function effectively, organizations must manage these interdependencies, as coordination is crucial for overcoming individual cognitive constraints. This need for coordination is the fundamental reason organizations exist. Information technology plays a vital role in this process, serving as a tool for organizing and managing information beyond individual capabilities.

1.2.2 Hierarchy

Hierarchy is a coordination mechanism based on command and control, where decision-making authority is centralized rather than delegated. It forms the foundation of many companies and institutions, ensuring the structured flow of information within an organization. To manage uncertainty effectively, hierarchies rely on two main types of information systems: vertical and horizontal.

Vertical information systems Vertical systems manage the flow of information along hierarchical lines, reinforcing structured decision-making. However, they have limitations when dealing with environmental uncertainty. As uncertainty increases, exceptions arise, creating the need for more planning and control mechanisms. These exceptions lead to additional information processing demands, often requiring information to flow upward toward higher hierarchical levels for resolution.

Horizontal systems In contrast, horizontal (or lateral) information systems facilitate direct communication between units at the same hierarchical level. These systems improve coordination by enabling decision-making at lower levels, reducing the reliance on top-down control. With a higher degree of delegation, horizontal systems enhance flexibility and responsiveness in dynamic environments.

1.2.3 Summary

Organizations can address environmental uncertainty in two main ways:

1. They can increase their information processing capacity by implementing vertical and horizontal information systems.

2. They can increase slack resources, such as maintaining warehouses or creating independent organizational units based on the divide et impera (divide and rule) approach, as seen in divisional structures.

However, the decision theory framework has its limitations. It assumes that hierarchies are the only coordination mechanism, overlooking market-based coordination as a viable alternative when hierarchies become inefficient. Additionally, it considers environmental uncertainty as the primary challenge, ignoring behavioral uncertainty caused by opportunistic individual behavior, which can also undermine hierarchical effectiveness. Transaction cost economics seeks to address these shortcomings by providing a broader perspective on coordination and uncertainty management.

1.3 Transaction cost economics

Williamson (1975) introduced the concept of transaction cost economics, which examines the costs associated with coordinating economic exchanges. In its simplest form, a transaction occurs when a customer receives a product or service from a supplier in exchange for payment. Transactions represent one of the oldest and most fundamental ways for individuals and organizations to cooperate, as they enable objectives that go beyond individual or organizational rationality.

Market systems A key function of transactions is to reduce behavioral uncertainty by mitigating opportunism. In market systems, individuals produce goods and services for themselves and maximize the benefits of their own efficiency. However, achieving coordination often requires executing transactions, which come with an associated transaction cost.

The total cost of a coordination mechanism is the sum of production costs and transaction costs. Market systems tend to have low production costs because individuals and firms operate efficiently. However, transaction costs remain low only under conditions of perfect competition, where market frictions such as information asymmetry, bargaining difficulties, and enforcement issues are minimized.

Economic transaction An economic transaction typically unfolds in four key phases:

1. *Matchmaking*: this stage involves identifying potential suppliers based on initial requirements. The outcome is a list of candidates that meet the specified criteria.
2. *Negotiation*: from the set of potential suppliers, one is selected through discussions that refine the requirements. The result is a formal agreement, often documented in a contract with defined service-level agreements.
3. *Execution*: the transaction is carried out according to the contract. The expected output includes the delivery of the product or service, along with any deviations or exceptions from the agreed SLAs.
4. *Post settlement*: if exceptions or issues arise, this phase involves managing them through established procedures to resolve disputes, enforce agreements, or make necessary adjustments.

1.3.1 Price system

The price system serves as the market's primary coordination mechanism, conveying crucial information about supply and demand. Prices are influenced not only by production costs but also by market dynamics. When the market functions efficiently, prices remain close to production costs and serve as a reliable indicator of product quality.

Market systems According to Williamson (1975), several factors can disrupt market efficiency:

1. *Shortages*: when supply fails to meet demand.
2. *Complexity*: when goods or services are too intricate for standard pricing.
3. *Customization*: when products require personalization, limiting standardization.
4. *Uncertainty and information asymmetry*: when buyers and sellers have unequal access to relevant information.
5. *Negotiation power imbalance*: when either buyers or sellers dominate price-setting.
6. *Transaction frequency*: when repeated transactions influence cost-efficiency.

When markets fail, businesses often resort to hierarchical coordination, such as in-house production, rather than relying on external suppliers. The decision between market-based transactions and hierarchical structures is primarily driven by cost considerations

1.3.2 Information technology role

Information technology (IT) acts as an organizational tool that reduces coordination costs. By improving information flow and transaction efficiency, IT strengthens market systems, leading to smaller, more numerous companies and reducing reliance on hierarchical structures.

1.3.3 Limitations

Viewing markets and hierarchies as mutually exclusive coordination mechanisms overlooks hybrid models. Traditional theories ignore behavioral uncertainty within organizations. Agency theory addresses these gaps by considering the complexities of decision-making and incentives within firms.

1.4 Agency theory

Agency theory challenges the traditional view that markets and hierarchies are entirely separate coordination mechanisms. Instead, it suggests a continuum between the two, recognizing that market-like coordination mechanisms exist even within organizations. By applying these mechanisms effectively, organizations can improve efficiency.

The key concepts of agency theory are:

- Organizations function as networks of contracts between individuals.
- Internal coordination is not solely based on command and control but also involves transactional exchanges.
- Just like external markets, organizations incur transaction costs, referred to as agency costs.
- Agency costs arise whenever decision-making responsibilities are delegated to lower levels of the hierarchy.

1.4.1 Agency cost

Delegation within an organization mirrors market transactions, creating an internal market with its own coordination expenses, known as agency costs. These costs include:

- *Control costs*: expenses related to monitoring and ensuring compliance.
- *Warranty costs*: costs associated with guaranteeing performance.
- *Residual loss*: inefficiencies that arise despite control measures.

1.4.2 Hierarchical control

In a perfectly competitive market, customers have no direct control over their suppliers (transactions are based entirely on trust and delegation). However, in imperfect markets, customers (or suppliers) may exert some level of control over their counterparts. This control can take the form of visibility into production processes or even hierarchical oversight, where suppliers operate under certain constraints imposed by their customers.

As a result, the distinction between internal markets and hierarchical coordination is not always clear-cut; instead, there is a spectrum of overlap between the two.

1.4.3 Limitations

The main limitations of agency theory are:

1. Hierarchical mechanisms exist within market transactions, blurring the boundaries between markets and organizations.

2. Agency theory overlooks task-related uncertainty, which affects the efficiency of coordination mechanisms.

3. The role of technology is task-dependent: technical innovation influences organizational structures and can shift the cost balance between market-based and hierarchical coordination.

To address these gaps, information systems theory explores how technology can enhance coordination and reshape organizational structures.

CHAPTER 2

Operational portfolio manufactoring

2.1 Enterprise Resource Planning

ERP systems often feature vertical solutions specifically designed to meet the needs of different industries. These solutions are tailored to be highly specialized. However, we can make a broader distinction between manufacturing and service companies:

- Manufacturing companies produce tangible products.
- Service companies provide intangible products.

2.1.1 Manufactoring companies

Manufacturing companies typically rely on three main functional portfolios within their ERP systems:

1. Administrative portfolio.
2. Operational portfolio.
3. Executive portfolio.

These portfolios, while initially developed separately, are now integrated within modern ERP systems, forming the core functionalities of these systems.

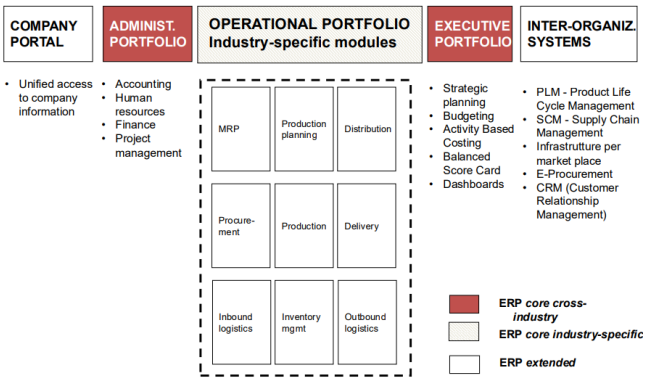


Figure 2.1: ERP functional architecture

2.1.2 Administrative portfolio

The administrative portfolio focuses on automating organizational activities that are often administrative and bureaucratic in nature, including: accounting and tax management, finance, human resources, project management (from an accounting perspective), and governmental procedures.

This portfolio is largely industry-agnostic, although it is country-specific. It represents an early stage of automation, alongside office automation, by streamlining tasks that involve number crunching. It typically involves minimal decision-making, as its processes are procedural and repetitive. Although traditionally viewed as stand-alone, it often links to other processes, such as activity-based costing.

Despite its simplicity in design, the administrative portfolio can be functionally complex.

2.2 Production and information processes

Porter's concept of information intensity and IT diversity helps explain how different industries rely on information and technology:

1. Information intensity refers to the amount and complexity of information required in an organization's processes. Generally, service industries require higher information intensity than manufacturing.

2. IT Intensity measures how well IT systems meet an organization's information processing needs. IT intensity is higher in banking than in insurance due to the industry's reliance on real-time transactions and data analysis. However, IT intensity can sometimes be greater in manufacturing than in services, depending on automation and digital integration.

3. Management inclination reflects how much a company's leadership views IT as a strategic asset. This varies based on factors like digital literacy, organizational culture, and company history. Historically, manufacturing companies have adopted IT earlier, while service industries experienced a lag of around ten years.

Information technology drivers Several factors determine how IT-intensive a company or industry can be:

1. *Structure of information processes*: the more structured and rule-based an activity is, the easier it is to automate using IT.

2. *Data volume*: the sheer amount of information that needs to be processed influences IT requirements.

3. *Operational frequency*: tasks that are repeated frequently benefit more from IT automation.

4. *Computational complexity*: simpler processes are easier to digitize and automate efficiently.

Porter's value chain Porter's value chain concept highlights how IT supports various business activities to create competitive advantages.

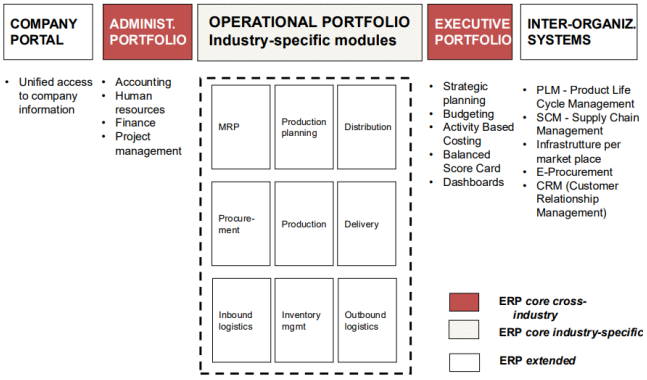


Figure 2.2: Porter value chain

Activity cycles Manufacturing involves continuous, iterative cycles that ensure efficiency and product quality. These cycles include:

1. *Development cycle*: focuses on designing and industrializing both products and production processes.
2. *Logistics cycle*: manages customer orders through:
   - *Procurement*: acquiring and handling materials, including reception, warehousing, and distribution to production plants.
   - *Production*: the physical transformation of raw materials into finished goods.
   - *Sales and distribution*: managing orders, external logistics, and post-sale services such as maintenance and customer support.

CHAPTER 3

Data and analytics in consulting

3.1 Consulting

Reply is a global network of over 150 specialized companies dedicated to helping organizations leverage cutting-edge technologies. Our mission is to drive innovation by enabling businesses to adapt to economic shifts and technological advancements, particularly those driven by the internet.

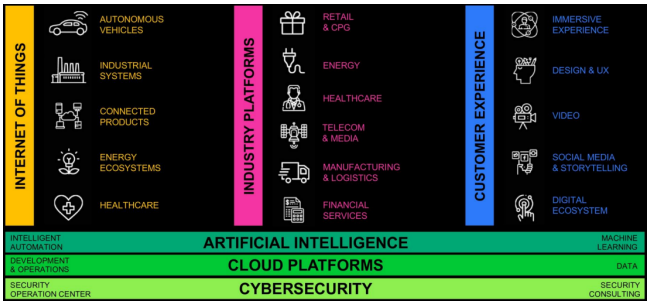


Figure 3.1: Reply services

Reply provides end-to-end solutions for businesses looking to transform raw data into valuable insights. From data collection to advanced analytics, our approach ensures that companies can harness the full potential of their data.

In a consulting firm, professionals progress through several key roles, each with increasing responsibility and expertise:

1. *Consultant or analyst*: this is the entry-level position, where individuals focus on data analysis, research, and supporting senior consultants. It's a foundational role that helps develop problem-solving, analytical, and research skills.

2. *Senior consultant or associate*: at this stage, professionals take on greater responsibility, leading specific project components and engaging more directly with clients. They begin to build deep expertise in a particular industry or domain while honing their client management skills.

3. *Manager*: managers oversee project teams, ensuring smooth project delivery while maintaining client relationships. They are also involved in business development and sales. This role emphasizes leadership, project management, and strategic client engagement.

4. *Senior manager*: with a broader scope, senior managers oversee multiple clients, contribute to business strategy, and play a key role in client acquisition. Their focus shifts towards strategic thinking, high-level client management, and business development.

5. *Partner*: as part of the firm's leadership, partners are responsible for setting strategic direction, expanding business opportunities, and managing client relationships at the highest level. They bring extensive experience in business strategy, leadership, and client management.

3.1.1 Working models

Consulting firms typically engage with clients through different contract models, depending on the project's nature and requirements.

- *Time and material*: in this model, the client pays based on the actual time spent and materials used, with an agreed hourly or daily rate for the resources employed. This approach offers flexibility to adjust project requirements as work progresses and is well-suited for projects where specifications may evolve or are not fully defined at the outset. However, budgeting can be challenging since the final cost depends on actual time and resources used.

- *Turnkey*: the consulting provider takes full responsibility for delivering a complete and functional product or service. The client receives the final result without managing the development process. Costs and timelines are clearly defined, as the provider oversees all project phases. This model requires minimal client involvement in day-to-day operations but offers limited flexibility to make changes once the project is underway. It is typically more expensive, as the provider factors in risks and contingencies.

- *Service*: the client pays for a predefined set of consulting services, such as technical support, strategic guidance, or other specialized expertise. This model allows clients to access specific skills without committing to a full project and is ideal for ongoing consultation or long-term support needs. However, it may require clear agreements on service scope and expectations and is less suitable for projects with well-defined and temporary objectives.

The choice between these models depends on the project's complexity, flexibility needs, and budget considerations. While time and material offers adaptability, turnkey ensures an end-to-end solution, and service contracts provide specialized expertise on demand.

3.1.2 Key Performance Indicators

Definition (*Revenue*)**.** The revenue is the total income generated by the company from its consulting services before any expenses are deducted.

It is a key indicator of overall sales performance and business growth. In a consulting firm, revenue typically comes from client contracts and project fees.

Definition (*Earning before tax*)**.** Earning before tax is a financial metric that measures a company's profitability before accounting for income tax expenses.

Earning before tax reflects the profit generated from core operations and other activities, such as investments or interest income, before taxes are deducted.

Definition (*Cost on revenues*)**.** cost on revenues is the ratio of costs directly associated with generating revenue, expressed as a percentage of total revenue.

This metric helps assess how much of the revenue is consumed by costs such as consultant salaries, software tools, and travel expenses. A high cost-to-revenue ratio may indicate inefficiencies in service delivery or pricing strategies.

Definition (*Unallocation*)**.** Unallocation refers to staff who are not directly assigned to revenue-generating activities.

Monitoring unallocated costs is crucial for identifying inefficiencies and ensuring that expenses are properly distributed across projects and services.

3.1.3 Project development

The two main approaches used in project development are waterfall and agile, each with its own strengths and limitations.

Waterfall The Waterfall model follows a sequential process, where each phase—analysis, design, development, testing, and implementation. Once a phase begins, changes to requirements are difficult to implement. This approach is best suited for projects with well-defined requirements from the start. In consulting, waterfall is ideal for projects with stable and predetermined requirements, particularly in industries where compliance, documentation, and structured processes are essential.

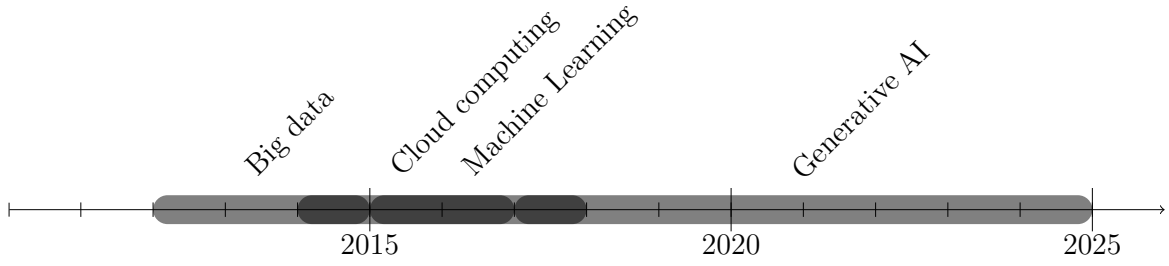
Advantages	
Clear requirements	A well-defined project scope ensures a structured development process
Predictability	Fixed timelines and structured phases make planning and resource allocation more manageable
Comprehensive documentation	Each phase includes detailed documentation, providing a thorough project record
Disadvantages	
Limited flexibility	Adapting to changes mid-project is challenging, making it less suitable for evolving requirements
Delayed feedback	Since testing happens at the end, user feedback may come too late, requiring costly revisions
Minimal client involvement	Limited collaboration during development can lead to misaligned expectations

Agile Agile follows an iterative and incremental approach, allowing for greater flexibility. Work is organized into sprints, each delivering a working product increment. Agile encourages continuous client collaboration and adapts easily to changing requirements. In consulting, agile is well-suited for projects where requirements may evolve, or when quick, tangible results are needed.

Advantages	
Adaptability	Changes can be accommodated at any stage, making Agile ideal for dynamic projects
Continuous feedback	Regular iterations ensure alignment with user needs and expectations
Client collaboration	Ongoing client involvement fosters a more interactive and responsive development process
Disadvantages	
Uncertain timeline	Iterative cycles can introduce unpredictability, making planning and resource management more complex
Minimal documentation	Agile prioritizes working software over documentation, which may be a drawback in highly regulated industries
Scope creep	Frequent changes and added features can lead to uncontrolled project expansion if not properly managed

3.2 Data

Data has become the foundation of decision-making, shaping industries and redefining business strategies. The journey of data-driven innovation can be divided into several key phases:



Big data Big data refers to the exponential growth of structured and unstructured data generated daily. It brought new challenges in storage, management, and analysis but also unlocked vast opportunities for business intelligence.

Technology	Hadoop, Hive, Impala, Cloudera
Key impact	Data-driven decision-making, process optimization, competitive advantage

Machine Learning Machine Learning marked the next stage of data evolution, enabling computers to learn patterns and make decisions without explicit programming. Machine Learning applications expanded rapidly, offering predictive insights and automation capabilities.

Technology	Neural Networks, Deep Learning, Reinforcement Learning, Clustering
Key impact	Advanced automation, improved predictions, enhanced decision-making

Cloud Computing Cloud computing revolutionized data storage and processing by providing scalable, cost-effective solutions over the internet. Businesses gained access to flexible computing power, reducing infrastructure constraints.

Technology	AWS, Google Cloud Platform, Microsoft Azure
Key impact	Scalability, cost reduction, innovation acceleration

Generative Artificial Inteligence Generative AI represents the latest frontier, where AI systems exhibit human-like understanding, learning, and application of knowledge across diverse domains. Its potential is reshaping industries and redefining human-technology interaction.

Technology	Large Language Models, synthetic data, Retrieval-Augmented Generation
Key impact	Creative automation, enhanced productivity, AI-driven decision making

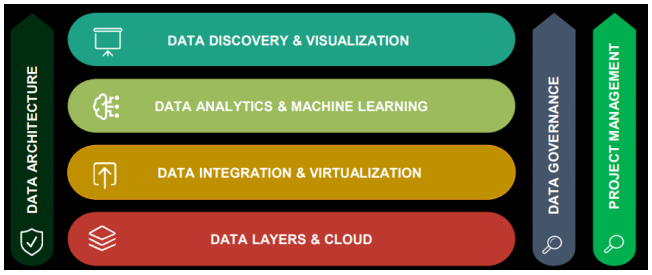


Figure 3.2: Data framework

3.2.1 Data job roles

Data and Cloud architect Data architects design comprehensive data infrastructures based on business objectives, ensuring seamless data integration and optimal storage solutions. Cloud Architects specialize in designing scalable, cloud-based architectures to support modern data needs.

Data engineer Data engineers build and maintain the systems required for collecting, processing, and storing data. They design ETL pipelines, work with big data technologies, and ensure that raw data is transformed into a usable format.

Data analyst Data analysts interpret and analyze data to provide actionable insights. They clean datasets, perform statistical analysis, and create visualizations to support business strategies and decision-making.

Data scientist Data scientists develop machine learning models and apply advanced analytics to uncover patterns and predictions from complex datasets. They work with programming languages like Python and utilize AI-driven techniques for deeper insights.

Data privacy and security specialist Data privacy officers ensure compliance with data protection regulations (e.g., GDPR), while security officers implement safeguards to protect organizational data from cyber threats and breaches.

3.2.2 Data storage

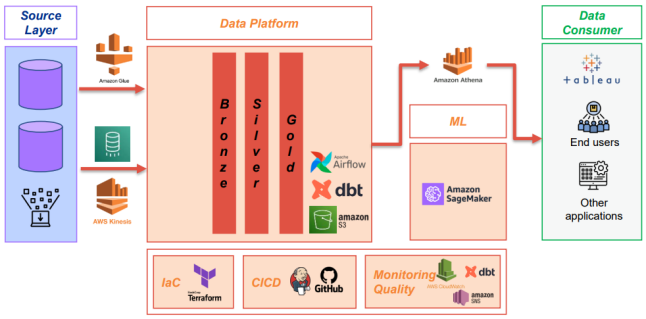


Figure 3.3: Data architecture

The data architecture is structured as follows:

1. *Source layer*: this is where raw data originates, coming from operational systems, websites, e-commerce platforms, and other sources.

2. *Data ingestion*: as data is collected from the source, it undergoes transformations to make it usable within the data platform. This step ensures that the data is properly formatted for analysis, enabling it to generate value for businesses.

3. *Data platform*: this is where the core data analysis happens. Given the large volume of data, robust systems are needed to process it effectively. The data refinement process follows a tiered approach:

   - *Bronze layer*: raw ingested data.
   - *Silver layer*: data undergoes Extract, Transform, Load processing.
   - *Gold layer*: fully analyzed and user-friendly data, optimized for reporting and decision-making.

4. *Machine Learning* (optional): in this stage, advanced analytics and machine learning techniques extract valuable insights and patterns from the processed data.

5. *Data consumer*: the final processed data and insights are presented to end users or integrated into other applications, often in a visualization tool like Tableau.

Definition (*Data warehouse*)**.** A data warehouse is a centralized repository that stores large volumes of structured data from various sources within an organization.

Unlike transactional databases, which prioritize real-time operations, data warehouses are designed for analytical queries and historical data analysis. They facilitate reporting, business intelligence, and decision-making by providing a consolidated and structured view of organizational data.

Definition (*Data lake*)**.** A data lake is a scalable repository that allows organizations to store vast amounts of raw, structured, semi-structured, and unstructured data.

Unlike a data warehouse, a data lake does not impose a schema on the data upon ingestion. Instead, it retains data in its original format until it is needed for processing or analysis. While data lakes provide flexibility for advanced analytics and machine learning, they require careful governance and security measures to maintain data quality.

Data lakehouse Organizations often integrate data warehouses and data lakes as part of a comprehensive data strategy, leveraging the strengths of both approaches:

- *Data integration*: raw data is ingested into the data lake, preserving its original format and serving as a staging area before further processing.
- *Data transformation*: Extract, Transform, Load processes can take place in both the data lake and data warehouse. Structured data required for immediate reporting is transformed and stored in the warehouse, while raw and unstructured data remains in the lake for exploratory analysis.

Definition (*Data lakehouse*)**.** A data lakehouse is a hybrid approach that combines data warehouses and data lakes.

It provides a unified architecture that supports structured, semi-structured, and unstructured data, enabling both traditional business intelligence and modern machine learning workflows.

Definition (*Polyglot persistence*)**.** Polyglot persistence refers to the practice of using multiple types of data storage technologies and databases within a single system.

Instead of relying on a single storage solution, organizations select the best-suited technology for each specific data type and use case, optimizing performance and scalability across different applications.

3.2.3 Data mesh

Definition (*Data mesh*)**.** Data Mesh is a modern approach to analytical data architecture that treats data as a product.

It is domain-driven, meaning that data ownership is distributed among teams that have the best understanding of their data and how it is used.

The key idea behind Data Mesh is that no one understands data better than its owner. Instead of relying on a centralized data team, Data Mesh distributes data responsibilities to domain-specific teams. Each domain aligns with a business function rather than specific applications or systems. Each domain manages its own data pipelines within a shared infrastructure and provides access to domain-specific data and functionalities via APIs.

Data Mesh aims to improve the way organizations handle data by focusing on three key objectives:

- *Business enablement*: democratizing data with a self-service approach, reducing dependence on IT.
- *Data management*: simplifying data processing, organization, and governance.
- *Organizational efficiency*: facilitating seamless exchange of data products between producers and consumers.

Feature	**Data warehouse**	**Data lake**	**Data mesh**
Centralization	Centralized	Centralized	Decentralized
Data structure	Structured	Unstructured	Both
Use cases	Reporting	Advanced analytics	Data product
Integration	Data lake	Data warehouse	Autonomous
Flexibility	No	No	Yes

3.2.4 Data governance

Data governance is the framework that defines how an organization manages its data to ensure accuracy, security, and compliance. It encompasses processes, roles, policies, standards, and metrics to optimize data usage, enabling a company to become truly data-driven.

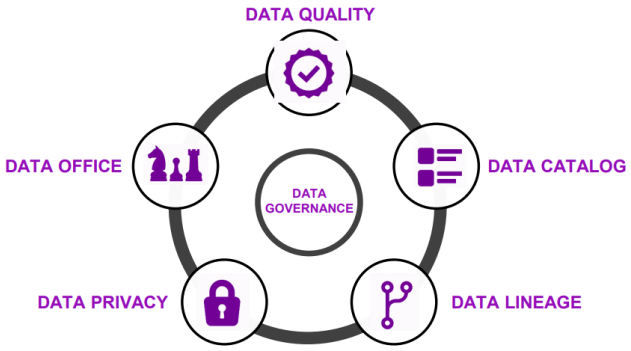


Figure 3.4: Data governance pillars

Data governance relies on several foundational elements:

- *Data quality*: ensures that information is accurate, reliable, and well-maintained. It involves defining validation rules, monitoring quality at various stages, and using analytics to assess reliability. AI-powered techniques, such as machine learning-driven validation, further enhance the integrity of data.
- *Data catalog*: helps organizations manage and understand their data assets. It integrates metadata from diverse sources such as databases, cloud platforms, ETL processes, and business intelligence tools. By maintaining a structured data dictionary for IT teams and a data glossary for business users, companies can enhance collaboration and ensure consistency in data interpretation. AI-driven search capabilities enable efficient metadata discovery and classification, supporting better knowledge sharing across the organization.

- *Data lineage*: tracks the entire lifecycle of data, from its origin to its final destination. By mapping transformations and interdependencies, it provides insights into how data moves through systems. This capability is essential for assessing the impact of changes, optimizing service requests, and ensuring regulatory compliance. Understanding data lineage also plays a crucial role in data protection, allowing organizations to pinpoint storage locations and enforce appropriate security measures.

- *Data privacy*: focuses on the identification and safeguarding of sensitive information. Effective privacy solutions assess risks, monitor data movement, and implement preventive measures to mitigate breaches. Encryption, data masking, and access control policies ensure compliance with regulatory requirements while maintaining confidentiality. Continuous analysis of data risks and proactive remediation strategies strengthen an organization's ability to protect its assets.

- *Data office*: plays a key role in overseeing data governance initiatives. Beyond technological solutions, governance efforts require clearly defined roles, responsibilities, and standardized management processes. The Data Office establishes documentation templates, development guidelines, and monitoring frameworks to ensure governance policies are consistently applied.

Organizations rely on specialized technology platforms such as Informatica and Collibra to implement robust data governance strategies. These tools provide comprehensive capabilities for managing data quality, cataloging metadata, tracking lineage, and ensuring privacy. However, technology alone is not enough: effective governance requires a combination of tools, processes, and dedicated personnel within a structured Data Office.

Beyond compliance and data management, data governance serves as a strategic enabler for organizations. By establishing strong governance practices, companies can leverage their data assets more effectively, leading to innovations in areas such as data as a service and data monetization.

Data monetization Data monetization transforms data into a valuable business asset by extracting insights that can be sold or leveraged for competitive advantage. High-quality, well-governed data enables businesses to identify new revenue opportunities, enhance market positioning, and strengthen partnerships. With the right governance framework in place, organizations can shift from merely being data-driven to becoming proactive data providers, delivering insights that generate tangible business value.

3.3 Data generation

Definition (*Synthetic data*)**.** Synthetic data refers to data that is artificially created to replicate the characteristics, patterns, and statistical properties of real-world data.

Synthetic data can be generated in large quantities and is designed to closely resemble the original dataset, making it indistinguishable from real data. Unlike real data, synthetic data does not contain any actual observations but is constructed to mimic the same structure and trends.

The primary purpose of synthetic data generation is to address situations where real data is limited, insufficient, or unavailable, especially when privacy concerns restrict access to actual

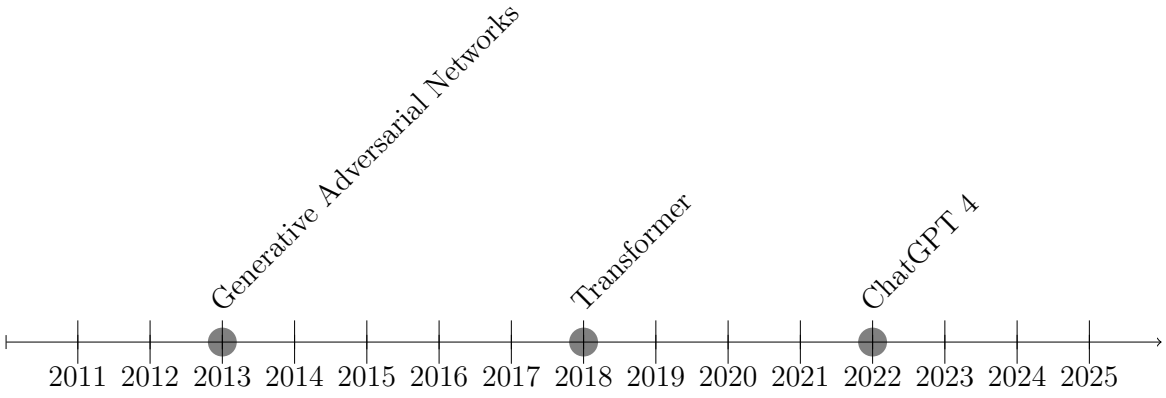
data. It can be used to supplement or replace real-world data, particularly when the use of the latter is challenging due to various reasons.
The advantages are:

- *Handling missing data*: it can help address gaps in data collection by providing a way to fill in missing information.
- *Reducing bias and balancing classes*: synthetic data can be used to reduce model bias and ensure more balanced datasets, improving model performance.
- *Facilitating data sharing*: since synthetic data does not contain real personal information, it can be shared more freely while remaining compliant with privacy regulations.
- *Cost reduction*: by reducing the need to collect large volumes of real data, synthetic data helps lower costs for companies and organizations.

Simulation Synthetic data can take advantage of the underlying mechanisms that generate real-world events, enabling the exploration of new, previously unencountered conditions. It allows for the simulation of scenarios that might be rare or extreme.

3.3.1 Data synthetization

Data synthetization started in 1928 with sample bootstrapping and improved in more recent times.



The possible methods for synthesizing data include:

- *Generative Adversarial Networks*: a type of neural network that consists of two components (a generator and a discriminator) that work together to create realistic synthetic data.
- *Variational Autoencoder*: a model that compresses data and then reconstructs it, allowing the generation of new samples that closely resemble the original dataset.

These data synthesis techniques can be applied across a wide range of fields, including tabular data, text generation, relational tables, time series, 3D models and computer vision.

3.3.2 Applications

The synthetic data framework can be used for:

- *Data monetization*: enables data sharing and monetization, fostering data-driven competitiveness and innovation for businesses of all sizes. Companies owning valuable data face difficulty sharing it due to privacy concerns and the need for third-party data for various products. Synthesizing sensitive data allows businesses to share it with third parties without violating GDPR policies.

- *Fraud detection*: creates a balanced dataset to improve fraud detection models, automates fake ID recognition, and removes sensitive customer data from the process. Identifying fake IDs used to obtain loans is difficult due to sensitive data concerns, lack of counterfeit documents, and the complexity of analyzing image-based data. Synthesizing document features allows data augmentation and balancing the dataset to better identify fake IDs.

- *Data masking for development*: synthetic data can be easily generated in large quantities for testing and development, allowing for reproducible scenarios and better control over data customization for specific needs. Sensitive data cannot be moved to non-production environments, and random or masked data often lacks representativeness, limiting its usefulness in testing. A synthetic model can be trained on production data to replicate sensitive information while maintaining privacy.

According to Gartner, by 2030, synthetic data will completely overshadow real data in Artificial Intelligence models.

Challenges The main challenges with synthetic data are:

- *Representativeness*: ensuring that synthetic data accurately mirrors the statistical properties and patterns of real data is challenging. The generated data must capture the complexity and diversity of real-world scenarios to be useful.

- *Validation*: it's essential to validate synthetic data by assessing whether models trained on it generalize well to real-world data. This ensures that the synthetic data truly reflects underlying patterns.

- *Bias and fairness*: the process of generating synthetic data must avoid introducing biases.

- *Transparency*: maintaining transparency and accountability in the use of synthetic data is crucial. Understanding the origin, characteristics, and limitations of synthetic data ensures its proper application and responsible use.

3.3.3 Explainable Artificial Intelligence

Explainable Artificial Intelligence is a research field focused on machine learning interpretability techniques designed to understand Machine Learning model predictions and explain these predictions in human-understandable terms.

As models grow more accurate and effective, their complexity increases, often transforming them into black boxes where decision-making processes are not easily understood. This raises concerns about the transparency of model predictions and the need to explain the rationale behind decisions made by AI systems. AI models, particularly complex ones, are often seen as black boxes, meaning their decision-making processes are opaque. This lack of transparency can hinder trust and adoption. AI systems making decisions that benefit only a few individuals, particularly in areas with high responsibility, can lead to significant issues. XAI ensures that decisions are made fairly and transparently, addressing biases in decision-making processes.

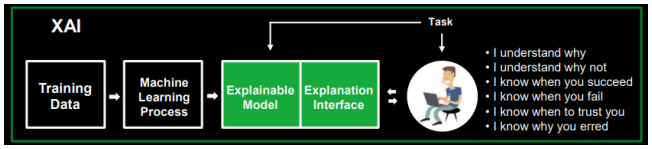


Figure 3.5: Explainable AI

3.3.3.1 Post-hoc methods

Post-hoc explainability methods aim to make predictions from existing machine learning models interpretable. These can be categorized as:

- *Model specific*: supports explainability constraints based on the learning algorithm and the internal structure of the model.
- *Model agnostic*: nnvolves applying pairwise analysis of model inputs and predictions, making the explanation applicable regardless of the model type.
- *Global*: provides an explanation that covers all units in the dataset, offering insights into the entire model's behavior.
- *Local*: 0ffers an explanation for a specific subset of data, focusing on a particular kind of dataset or input.

Common post-hoc XAI approaches includes:

- *Feature importance*: assigns an importance value to each variable based on how much it influences the prediction.
- *Surrogate models*: simplifies a complex model by approximating it with a more explainable, interpretable model.
- *Rule-based*: uses if-then rules to express combinations of input features and their activation values.
- *Saliency maps*: highlights the important input pixels or features that have the most significant impact on the prediction.

Benefits The benefits of XAI are:

- *Justify*: helps explain and justify the system's behavior, allowing users to detect unknown vulnerabilities and flaws.
- *Improve*: facilitates the improvement of model performance, debugging, and verification by applying structured engineering approaches.
- *Control*: ensures better control and supervision of the algorithm, guaranteeing its proper usage and compliance with standards.
- *Discover*: enables the discovery of potential biases or faults in the model by comparing its reasoning, helping assess its validity and reliability. It also identifies the most influential factors behind predictions.

3.3.3.2 Reponsible AI

Responsible AI encompasses various sub-disciplines, including: explainable AI, human-centered AI, compliance, ethical AI, secure AI, and interpretable AI.

The goal of responsible AI is to develop AI systems that are not only effective but also transparent, ethical, secure, and compliant with established standards, ensuring they are used responsibly in real-world applications.

3.4 Generative Artificial Intelligence

Definition (*Generative Artificial Intelligence*)**.** Generative AI is a branch of artificial intelligence focused on creating systems that can generate new data samples, which closely resemble those in the training dataset.

Unlike discriminative models, which are designed to classify input data into predefined categories or make specific predictions, generative models aim to learn the underlying structure of the data and generate new instances that reflect the distribution of that data. In essence, generative AI seeks to replicate the creativity and generative abilities of humans by learning from examples and producing new content with similar characteristics. These models are trained on large datasets and can generate realistic data across a variety of domains.

Applications Generative AI has a wide range of applications across different domains, including:

- *Image generation and editing*: Generative Adversarial Networks are commonly used to generate realistic images. They are also effective for image-to-image translation tasks, such as style transfer and colorization.
- *Text Generation and summarization*: language models can generate coherent, contextually relevant text.
- *Content creation and augmentation*: generative AI can be used to create entirely new forms of content, including music, videos, and artwork. Additionally, these models can enhance existing content by generating variations or filling in missing parts.
- *Data synthesis and simulation*: generative models can create synthetic data that closely resembles real-world data, making them useful for data augmentation, training machine learning models, and simulating realistic scenarios.

Definition (*Large Language Models*)**.** A Large Language Model is a type of artificial intelligence designed to understand and generate human-like text based on the input it receives.

These models are trained on vast amounts of text data, typically sourced from the internet or other extensive corpora, in order to learn the complex patterns and structures of human language.

Large language models are typically built using deep learning architectures, often based on transformer architectures. These models consist of multiple layers of neural networks that process input text in a hierarchical manner, capturing both local and global dependencies in the text. The result is a model capable of producing highly coherent and contextually relevant text.

Once trained, LLMs are capable of performing a wide variety of natural language processing tasks, including text generation, text completion, translation, summarization, and question answering. The ability of LLMs to generate text that is contextually appropriate and fluent has made them widely used in many applications, including virtual assistants, content creation, and more.

Prompt engineering

Definition (*Prompt engineering*)**.** Prompt engineering is the practice of designing and crafting prompts or inputs that are used to interact with language models or AI systems in order to achieve specific, desired outputs.

The goal is to optimize how questions or requests are framed to ensure the AI generates accurate, useful, and relevant responses. Key tips for effective prompting include being clear and specific, providing relevant context, asking open-ended questions, using important keywords, avoiding ambiguity, engaging in conversation, offering feedback, and experimenting with different approaches.

3.4.1 Multimodal generative AI

Multimodal generative AI refers to systems that can create content across multiple modalities. These advanced AI systems use machine learning techniques to process and generate content from various data types.

Unlike traditional generative models that focus on a single modality, multimodal models can handle multiple data types at once, enabling more expressive and comprehensive content generation. These models understand the relationships between different forms of data. This cross-modal understanding results in more contextually relevant and dynamic content.

Training Training multimodal AI models requires large, diverse datasets that include paired examples of different data types These datasets help the model learn the correlations and connections between various modalities, enhancing its ability to generate accurate and coherent content across different forms of media.

3.4.2 Ethic

Generative AI raises several ethical concerns that need careful consideration.

3.4.3 Enterprise applications

Definition (*Retrieval Augmented Generation*)**.** Retrieval Augmented Generation is a technique in natural language processing that combines retrieval-based methods with generative models to improve the quality and relevance of generated text.

In RAG, a retriever component searches a large database or text corpus to find information relevant to the input query or context. This information is then used to enhance the generative model's process producing more accurate, coherent, and contextually relevant responses. With RAG, company data can be used to enrich prompts.

Ethical concern	**Description**
Misinformation	Generative AI can create convincing fake content, leading to misinformation and manipulation
Privacy Concerns	Models may inadvertently reproduce sensitive information, risking privacy violations
Bias and Fairness	AI models can perpetuate biases, leading to unfair or discriminatory outcomes
Intellectual Property	AI-generated content raises issues about ownership and copyright infringement
Security Risks	AI can be misused for malicious activities, such as phishing or creating deepfakes
Identity Theft	Generative AI may be used for identity theft, fraud, and other criminal activities
Regulatory Challenges	Current laws may not fully address the complexities of generative AI technology
Creative Industries	AI could disrupt creative fields, leading to job displacement and economic challenges

Table 3.1: Ethical concerns in generative AI

Architectures In the enterprise architectures with generative AI, we have the following components:

- *Vectorial database layer*: a storage layer that organizes data in a format easily understood by a large language model.
- *Feedback*: a system to collect feedback and contextual information, allowing for continuous improvement and evolution of the model.
- *RAG agent*: the core component of the system responsible for managing inputs and outputs.
- *Guard eail*: a safeguard to prevent model leakage or malicious attacks, ensuring the quality and security of outputs. Limiting the LLM's ability can result in performance degradation and increased latency as all inputs and outputs must be checked.
- *LLM gateway and catalog*: this component is responsible for routing prompts to the most suitable LLM for processing.

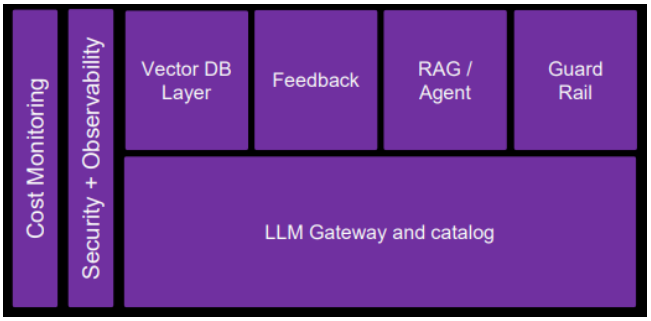


Figure 3.6: Explainable AI