Politecnico di Torino

Master's Degree in Engineering and Management

A.y. 2024/2025

Graduation Session November 2025

Toward Multi-Cloud Ecosystems

An adoption framework for federated multi-provider environments

Supervisors:
Paolo Neirotti
Fulvio Risso

Company Tutor:
Giuseppe Zangari

Candidate:
Carlo Alberto De Angelis

Abstract

In today's digital era, data are considered among the most valuable assets for organizations across every sector of the economy. Within this context, cloud computing enables outsourcing IT services required to store, analyze, and process such data, reducing capital expenditures while improving scalability, flexibility, and collaboration. However, as enterprises increasingly migrate to cloud environments as part of their digital transformation journey, they face critical challenges related to interoperability, governance, and digital sovereignty. To overcome these limitations, multi-cloud has emerged as an innovative approach that allows the distribution of workloads across multiple cloud providers, enhancing flexibility and resilience while mitigating risks of dependency on a single vendor.

This thesis, developed in collaboration with ArubaKube, explores the evolution of cloud architectures from both a strategic and technical perspective. In particular, it focuses on Kubernetes and multi-cluster topologies as enablers of federated, multi-provider ecosystems. The effectiveness of these ecosystems is further enhanced by exploiting cloud-agnostic solutions and open-source tools, such as Liqo, which allows organizations to seamlessly share a pool of computing resources across cloud and edge environments, enabling the computing continuum.

Through a comprehensive analysis of the ecosystem dynamics and innovative federation technologies, the work proposes a multi-dimensional adoption framework that spans the firm's strategy, operations and infrastructure, with ecosystem governance acting as the central coordination mechanism.

Acknowledgements

I would like to express my gratitude to Professor *Paolo Neirotti* for being the first to spark my interest in strategic disciplines, and for his invaluable guidance throughout the development of this project.

I also wish to extend my deep appreciation to Professor *Fulvio Risso* for his remarkable contribution to the open source community, in particular through the development of *Liqo* which inspired both my thesis works, and for enabling the collaboration with *ArubaKube*, an experience that has been extremely formative for my future professional career.

A special acknowledgement goes to *Giuseppe Zangari*, who accompanied me with exceptional expertise and professionalism throughout this long yet inspiring journey. His insights and our many discussions over these months have been an invaluable source of personal and professional growth.

I am also grateful to the entire *ArubaKube* team for their support, an exceptional group with rare competencies in the field, for whom I would like to express my sincere admiration for their work.

Table of Contents

List of Figures

List of Acronyms

ABAC Attribute-Based Access Control

AI Artificial Intelligence

API Application Programming Interface

AWS Amazon Web Services

AZ Availability Zone

CAGR Compound Annual Growth Rate

CapEx Capital Expenditures

CCoE Cloud Center of Excellence

CI/CD Continuous Integration and Continuous Delivery

CIS Cloud Infrastructure and Services

CNCF Cloud Native Computing Foundation

CRM Customer Relationship Management

CSP Cloud Service Provider

DevOps Development and Operations

DR Disaster Recovery

ERP Enterprise Resource Planning

FaaS Function-as-a-Service

FinOps Financial Operations

GCP Google Cloud Platform

GDPR General Data Protection Regulation

GPU Graphics Processing Unit

GTM Go-To-Market

HA High Availability

IaaS Infrastructure-as-a-Service

IaC Infrastructure-as-Code

IAM Identity and Access Management

IDP Internal Developer Platform

IDSA International Data Spaces Association

IoT Internet of Things

IPCEI Important Project of Common European Interest

ISP Identity Service Provider

IT Information Technology

K8s Kubernetes

LLM Large Language Model

ML Machine Learning

MSP Managed Service Provider

NIST National Institute of Standards and Technology

OpEx Operational Expenditures

OS Operating System

PaaS Platform-as-a-Service

PII Personally Identifiable Information

RBAC Role-Based Access Control

RBV Resource-Based View

SaaS Software-as-a-Service

SME Small and Medium-sized Enterprise

TOE Technology-Organization-Environment

TTM Time-to-Market

VM Virtual Machine

VPC Virtual Private Cloud

VPS Virtual Private Server

Chapter 1

Introduction

As digital technologies become increasingly relevant in everyday life, people interact with cloud computing services on a daily basis, even without noticing it. Uploading a file to Google Drive, Dropbox, or iCloud; streaming media on Netflix or YouTube; interacting with Large Language Models (LLMs) such as ChatGPT or Claude; or simply checking emails through Gmail are all everyday actions that rely on cloud computing infrastructures. To provide a numerical perspective, the number of personal cloud storage users has increased from approximately 1.1 billion in 2014 to an estimated 2.3 billion in 2025 [1], and public cloud service end-user spending is constantly growing, as shown in Figure 1.1.

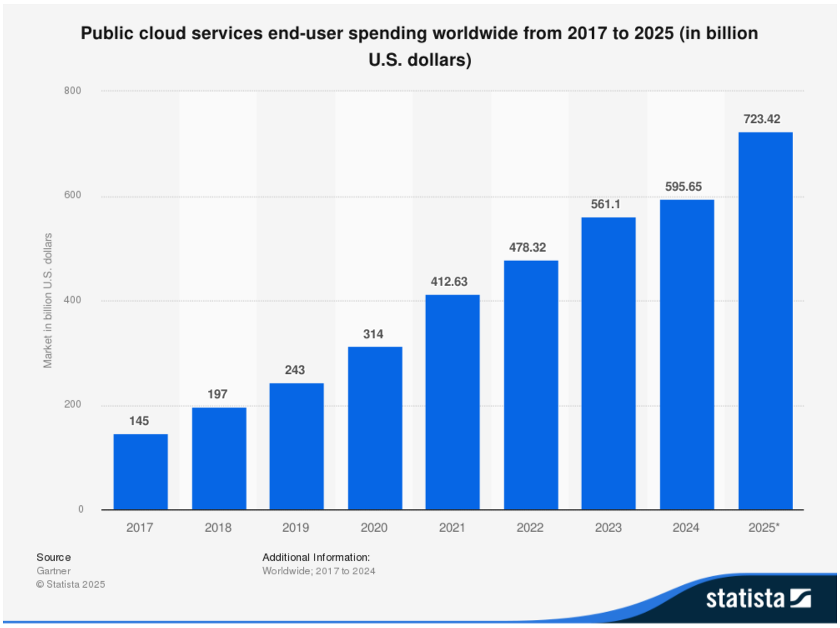


Figure 1.1: Public cloud services end-user spending worldwide [2].

The total amount of data created, captured, copied, and consumed worldwide is projected to exceed 200 zettabytes[1] in 2026 [3]. This growth is being accelerated by the widespread adoption of Artificial Intelligence (AI) and LLMs in particular, which demand large amounts of data and computational resources.

While individual customers interact with cloud services daily, companies derive the most strategic value from the cloud ecosystem. In today's digital era, data are considered some of the firm's most valuable assets. Data not only constitute the backbone of emerging technologies such as AI and Internet of Things (IoT) but are also essential for organizations to make informed decisions, enhance customer engagement, and support process modernization. As of 2025, more than 50% of organizations' data reside in the public cloud, as shown in Figure 1.2.

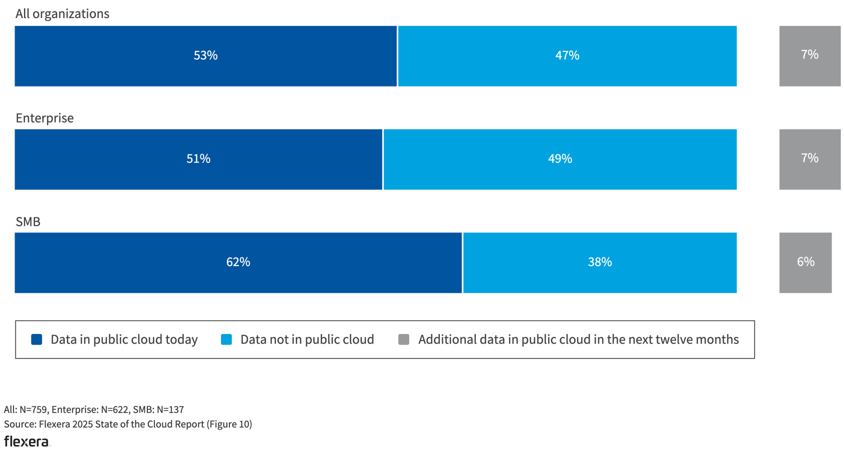


Figure 1.2: Percentage of data stored in the public cloud by organization type [4].

The process of transferring data, applications, and workloads from an on-premises environment to a cloud-based environment is referred to as cloud migration [5]. This process is a fundamental step in broader initiatives of digital transformation, which involve integrating digital technology across all areas and processes of an organization. Digital transformation aims to improve operational efficiency, agility, and overall productivity by modernizing outdated systems.

As the demand for modernization and digital transformation increases, the adoption of cloud infrastructure services is displacing classic data center hardware and software investments for a large number of firms worldwide, as highlighted by the statistic in Figure 1.3.

[1] 1 Zettabyte equals approximately 1 trillion Gigabytes (1 ZB $\approx 10^{12}$ GB).

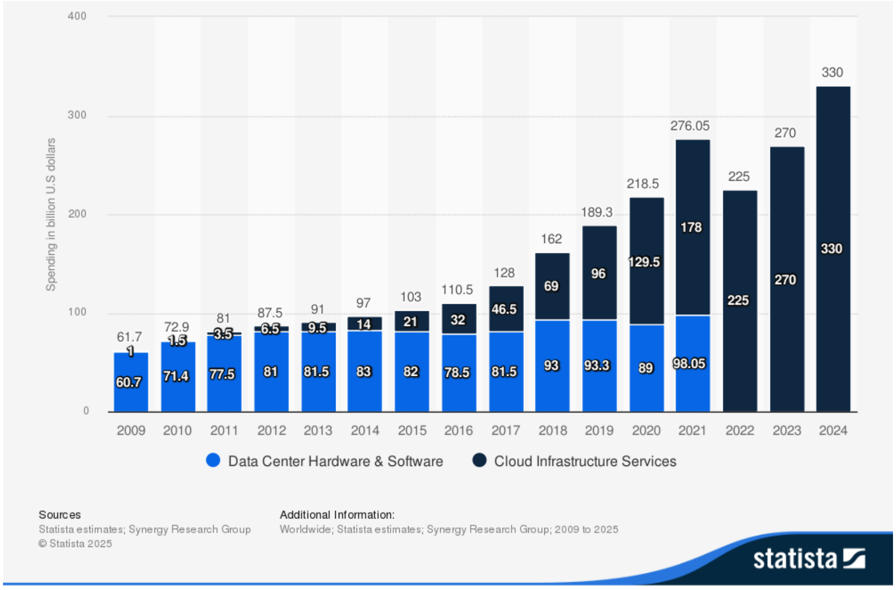


Figure 1.3: Enterprise spending on cloud and data centers from 2009 to 2014 [6].

1.1 Why Multi-Cloud?

Despite the benefits offered by cloud computing, relying on the infrastructure of a single cloud provider still presents significant limitations, particularly for enterprises offering a wide range of distributed services across multiple regions. For instance, an organization leveraging infrastructure services from a single cloud provider may need to adopt a solution from a different vendor to address specific workload requirements.

The dependency on a single cloud provider, commonly referred to as vendor lock-in, constitutes a critical barrier to the integration of services from multiple vendors, as they often rely on different technology stacks and Application Programming Interfaces (APIs) that can obstruct interoperability and portability across distinct environments. As a result, the competitive dynamics within the cloud computing market resemble an oligopoly model, with a limited number of competitors capable of autonomously defining governance rules for their digital platforms, leveraging vast computational resources, and exploiting significant benefits from economies of scale. From the consumer's perspective, this scenario implies a lower degree of control over their data and digital assets.

To foster a fairer competitive landscape and enable companies to retain greater control over their data and infrastructure, the European Commission has proposed several initiatives as part of the European Digital Strategy, aiming to favor the standardization process among industries and cloud service providers.

Within this process, the open source ecosystem plays a pivotal role, offering a variety of cloud-agnostic tools that enable interoperability and application portability across heterogeneous infrastructures. Among these, the integration of Kubernetes (K8s) [7] with Liqo [8], an open-source project developed at Politecnico di Torino, proposes a cutting-edge solution for organizations to leverage a pool of shared computing resources and manage the environment as a unified system, typically referred to as a cluster federation.

This dynamic and flexible approach enables the computing continuum, an innovative paradigm that provides seamless workload portability within cloud and edge environments, shaping the transition toward federated multi-provider ecosystems.

1.2 Thesis Structure and Research Methodology

The thesis was conducted in collaboration with ArubaKube, a spin-off founded through the collaboration between Politecnico di Torino and Aruba S.p.A. [9]. The company focuses on developing innovative cloud-native solutions based on K8s and other open-source technologies, including Liqo, to which it is a major contributor. This study complements the internship activity carried out with ArubaKube to support the Go-To-Market (GTM) strategy of an enterprise platform designed to facilitate operations and promote resource sharing within a multi-provider Kubernetes cluster federation enabled by Liqo. In particular, the work offers an analysis of the cloud and multi-cloud ecosystems dynamics, from both the customer and provider perspective.

The first part of the work focuses on understanding how migration toward cloud solutions can generate significant organizational benefits for a broad range of firms, regardless of their size and technological capabilities. However, to fully exploit the potential of cloud-based solutions, it is essential for organizations to align their Information Technology (IT) strategy with long-term business goals when migrating workloads to cloud environments. This section provides a comprehensive literature review on strategy, innovation, and digital transformation, and an analysis of the current market landscape, characterized by a limited number of cloud providers, to identify the sources of hyperscalers' competitive advantage. Additionally, it offers a holistic perspective on the cloud computing value system, including emerging paradigms such as cloud-native architectures, edge computing, and artificial intelligence, which are progressively reshaping centralized approaches.

The second part of the thesis introduces multi-cloud adoption and Kubernetes federation as key enablers of federated multi-provider ecosystems, which shift the governance perspective from cloud providers to organizations consuming their services. This transition is facilitated by key architectural principles such as infrastructural modularity and interoperability, which enable organizations to maintain greater control over ecosystem governance dynamics. This section presents the dynamics characterizing the emerging ecosystem, as well as highlighting the role of European initiatives fostering an agile data economy and a competitive multi-provider landscape to favor the transition. Additionally, a study was conducted based on a patent analysis and a Bass diffusion curve estimated from available statistical data to assess the diffusion dynamics of K8s-related technologies in the European landscape. Finally, the work proposes an adoption framework to illustrate the federated multi-provider ecosystem from the perspective of organizations adopting those services, and a complementary survey for IT practitioners addressing strategic drivers, operational challenges, and infrastructural implementations of multi-cloud adoption. Figure 1.4 provides a schematic representation of the described structure.

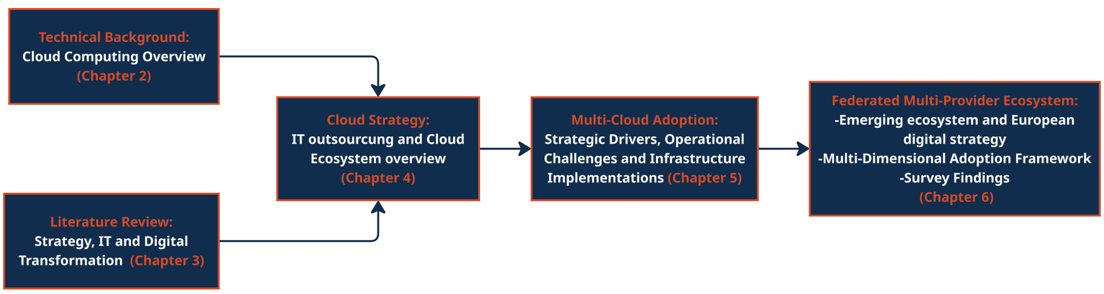


Figure 1.4: Structure of the thesis.

Chapter 2

Cloud Computing Overview

This chapter provides a detailed introduction to cloud computing as the technological foundation of this work. In particular, it explores the main categories of cloud computing services and deployment models, in accordance with the definitions provided by the National Institute of Standards and Technology (NIST). To illustrate the trade-offs associated with different cloud solutions, two practical use cases are presented at the end of the chapter.

2.1 Definition of Cloud Computing

NIST defines cloud computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources [...] that can be rapidly provisioned and released with minimal management effort or service provider interaction" [10].

Examples of computing resources include networks, servers, storage, applications, and services, organized through a layered stack. The sharing of these resources, owned by Cloud Service Providers (CSPs), through a pay-as-you-go model enables organizations and individuals to access advanced technological capabilities "as a service" without owning and maintaining the underlying hardware and infrastructure, which would imply higher capital expenditures. According to NIST [10], cloud computing services present five essential characteristics:

- **On-demand self-service:** users can automatically provision computing resources without needing to interact with the service provider. Services should always be available except for unexpected outages or scheduled maintenance.

- **Broad network access:** capabilities should be accessible over the network through standard mechanisms and from multiple devices (e.g., phones, laptops, or workstations).

- **Resource pooling:** computing resources are pooled to serve multiple consumers (multi-tenancy) and are dynamically assigned based on demand. This model enables economies of scale and cost efficiency for both users and providers. However, the consumer generally does not control the exact location of the provided resources.
- **Rapid elasticity:** resources can be elastically provisioned and released to scale according to demand. From the consumer's perspective, resources may appear virtually unlimited, as they are available in any quantity at any time.
- **Measured service:** cloud systems optimize and control resources through metering capabilities (typically on a pay-as-you-use or charge-as-you-use basis), providing transparency for both users and providers.

2.1.1 Deployment Models

Organizations adopting cloud infrastructure services can leverage cloud computing through different deployment models, based on their needs and internal expertise in managing IT systems and organizational requirements. In many cases, CSPs and consultancy companies offer managed services to support the process of cloud migration and implement the most suitable environment. The three main deployment models are public cloud, private cloud, and hybrid cloud.

Public Cloud

In the public cloud model, the CSP owns, manages, provisions, and maintains the entire infrastructure. The provider delivers computing resources to the consumer over the Internet, offering a wide range of cloud services on demand (e.g., servers, networking, or storage). The technology enabling this process is virtualization, where a hypervisor manages multiple operating systems on a single physical host. This process enables multi-tenancy: different customers can access multiple Virtual Machines (VMs), which are logically isolated instances of a physical server. CSPs usually operate under pay-as-you-go or subscription-based pricing models, bearing capital, operational, and maintenance expenses for the underlying infrastructure, implying a significantly reduced total cost of ownership for customers.
Figure 2.1 shows the high-level architecture for a public cloud deployment.

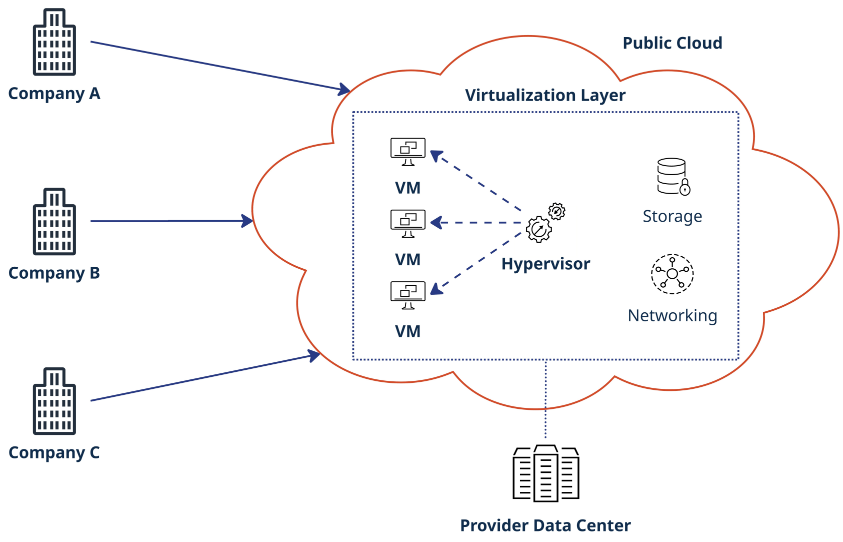


Figure 2.1: Public cloud architecture.

The on-demand provisioning of public cloud services enables scalability of the infrastructure, which can adapt to fluctuating demand through the dynamic allocation of computing resources. However, outsourcing infrastructure management reduces the organization's control over data and computing environments. Additionally, performance, reliability, and security entirely depend on CSP standards. As a result, many IT departments prefer to outsource only the management of less critical and standardized applications to public cloud providers.

Private Cloud

A private cloud refers to a cloud environment entirely dedicated to a single organization or business unit. It can be delivered both on-premises, when the hardware is installed directly in the organization's own data center and is managed and maintained by the organization itself or by an external provider, or off-premises, hosted by an external provider.

In the case of an off-premises configuration, the dedicated environment can be provisioned through bare-metal servers, which grant single-tenancy through dedicated hardware, or a Virtual Private Cloud (VPC). A VPC logically isolates a portion of the public cloud environment, enabling the customer to deploy VMs on top of the public provider's infrastructure within a virtual network with its own dedicated IP range.
Figure 2.2 shows the high-level architecture for a private cloud deployment.

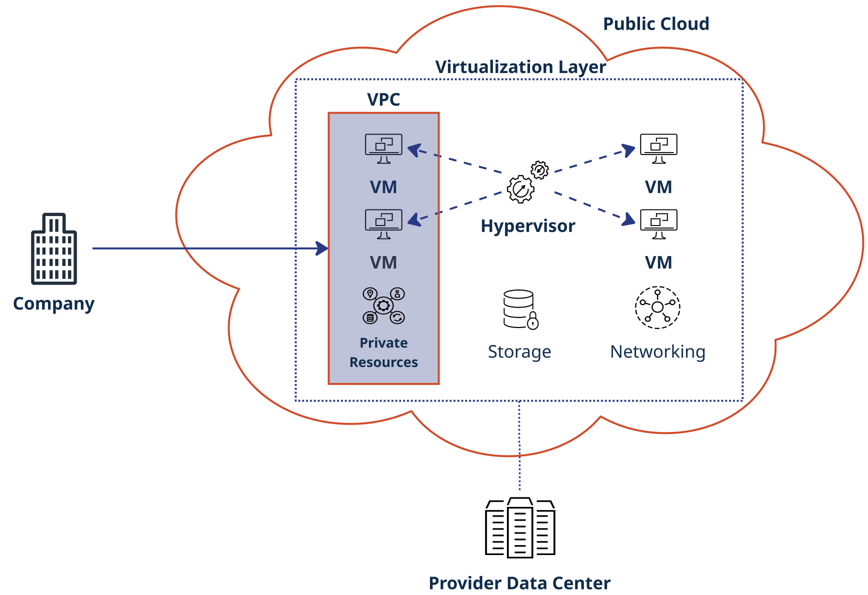


Figure 2.2: Private cloud architecture.

This solution provides a secure and isolated environment within the public cloud, allowing organizations to maintain greater control over the infrastructure while still leveraging the reduced operational complexity resulting from the outsourcing of physical hardware management.

Hybrid Cloud

A hybrid cloud model combines a private cloud and a public cloud deployment into a single flexible infrastructure, allowing interoperability and workload portability between the two environments [9]. This solution enables organizations to leverage the advantages of both deployment models, with the flexibility and scalability of the public cloud and the security and control of the private cloud for critical workloads. A typical hybrid cloud use case is the deployment of mission-critical applications or regulated workloads containing sensitive data that need to comply with specific regulations in the private cloud while deploying less sensitive and dynamic workloads in the public cloud environment (see 2.2.1).

An additional capability of the hybrid cloud is cloud bursting, which enables workloads running on the private cloud to leverage additional capacity from the public cloud resources in case of demand spikes.

Figure 2.3 presents a possible high-level architecture of a hybrid cloud environment.

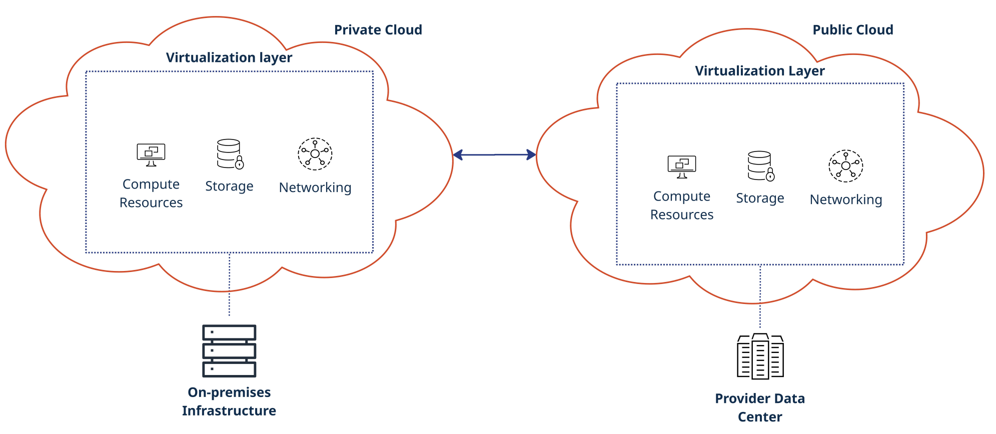


Figure 2.3: Hybrid cloud architecture.

A correctly configured hybrid cloud environment ensures performance, compliance, and flexibility for organizations. However, the model introduces operational complexity and requires significant expertise for management and configuration. For this reason, many organizations rely on Managed Service Providers (MSPs) for the configuration and maintenance of the organization's hybrid cloud deployment. As of 2024, the hybrid cloud has become the most widely adopted deployment strategy among enterprises, as highlighted by Figure 2.4.

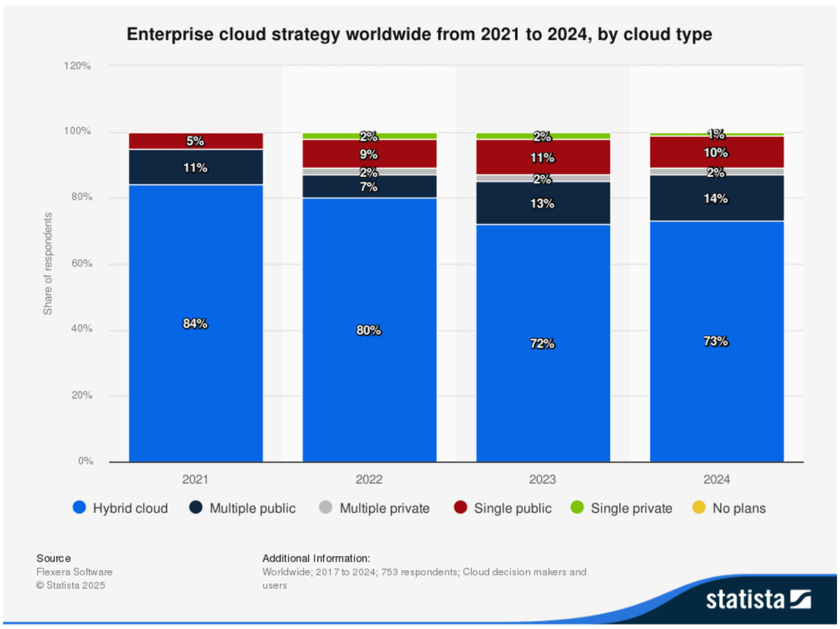


Figure 2.4: Enterprise cloud strategy worldwide from 2021 to 2024 [11].

2.1.2 Service Models

Cloud providers offer a wide range of services, which differ in terms of costs, responsibilities, and control the consumer exercises over the underlying infrastructure. In each service model, the provider is responsible for managing large-scale physical data centers, organized into Availability Zones (AZs). Each AZ provides redundancy and high availability for the provider's services within a broader geographical area, called a cloud region, which is distributed to reduce latency and meet data residency and compliance requirements.

According to NIST [10], the layered technological stack offered by cloud providers, which ranges from physical data centers and raw computing resources to complete applications, is organized into three main service models: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS). Figure 2.5 illustrates the hierarchical distribution of services and resources managed by the provider across each layer of the stack.

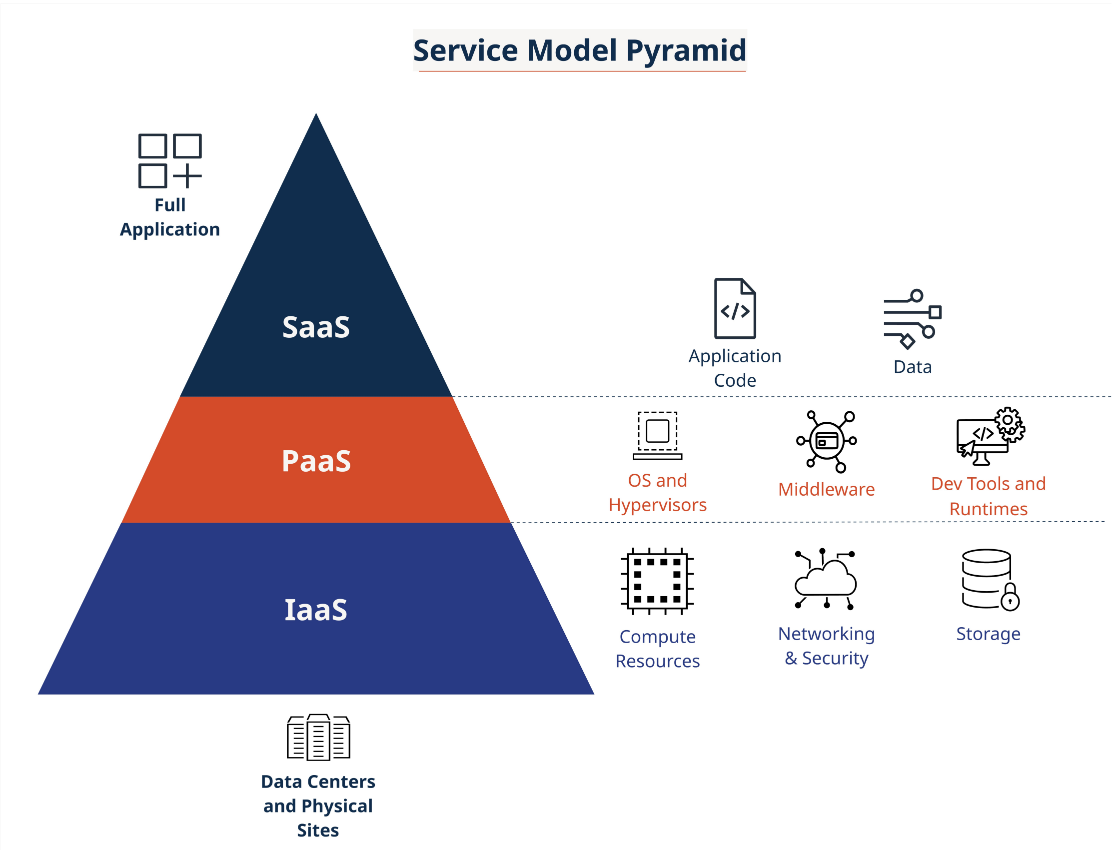


Figure 2.5: Service model pyramid.

Moving upward in the service model pyramid, the provider's responsibility increases. From the consumer's perspective, this is typically associated with lower costs and operational complexity, but also with less flexibility and control over the configuration. Therefore, organizations should carefully evaluate the trade-offs in terms of costs, flexibility, control, and expertise to choose the most suitable solution.
Figure 2.6 illustrates the distribution of responsibilities between the provider and the customer across the three service models.

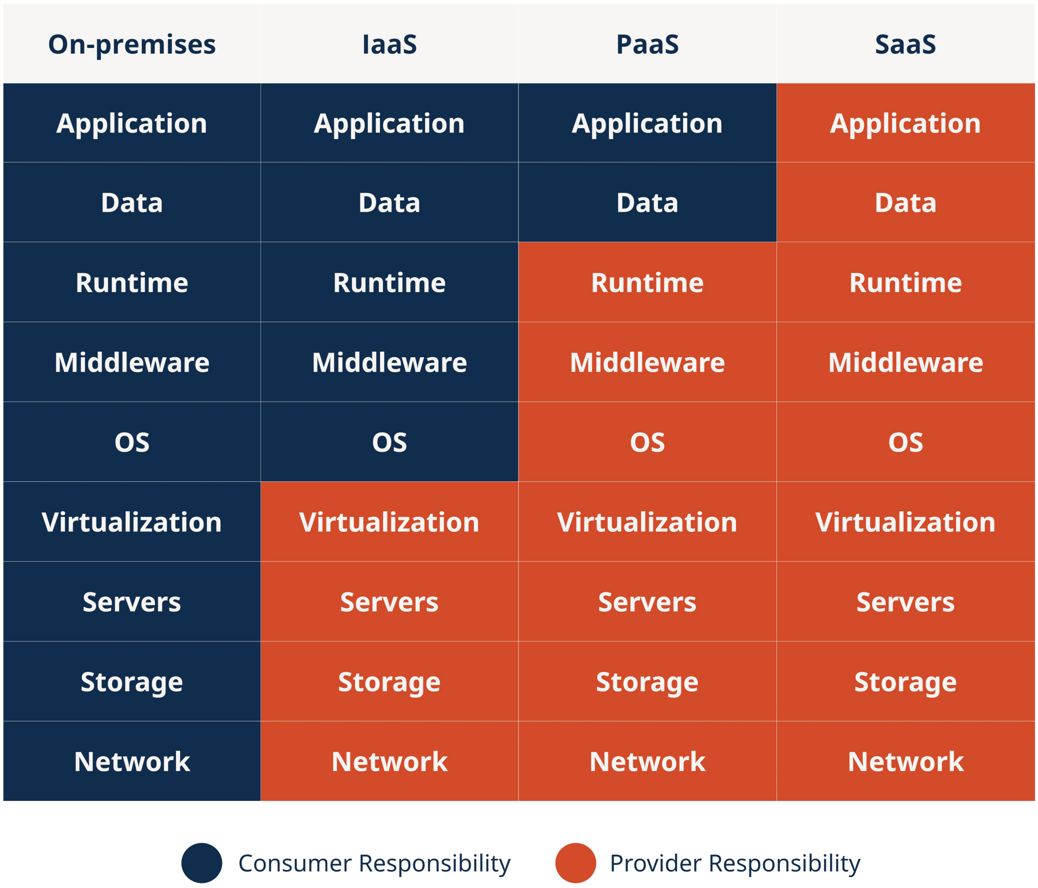


Figure 2.6: Distribution of responsibilities for cloud service models.

Infrastructure-as-a-Service (IaaS)

In the IaaS model, the provider is responsible for managing physical infrastructure and computing resources, as well as the virtualization layer. The consumer can provision VMs within the provider's AZs. Each VM runs on the provider's shared infrastructure and is managed by the vendor hypervisor. VMs, or Virtual Private Servers (VPSs) specifically, provide the customer with an isolated virtual environment, including its own Operating System (OS) and a share of the total physical server's resources, which can be increased or decreased according to the demand. For higher compliance or performance needs, organizations can opt for a "bare-metal" server, which is dedicated, single-tenant physical hardware. Typical examples of IaaS offerings include:

- **Compute:** VPS, bare-metal servers.
- **Storage:** File, Object, or Block Storage.
- **Networking:** VPC, Firewalls, and Load Balancers.

Platform-as-a-Service (PaaS)

In the PaaS model, the provider manages operating systems, middleware, and runtime environments in addition to the underlying infrastructure and computing resources. Platform services abstract lower-level system components from the users, resulting in faster time to market and offering a variety of capabilities, such as the ability to experiment with different operating systems, frameworks, and languages without the need to handle installations and configurations directly. PaaS environments are typically used by software engineering teams and Development and Operations (DevOps) professionals, who can leverage a ready-to-use platform to develop, test, and deploy application code with scalable resource provisioning.

This service model seamlessly integrates with DevOps workflows, providing built-in tools that support Continuous Integration and Continuous Delivery (CI/CD) pipelines. The consumer is still responsible for their own application and data, and each deployment requires an initial configuration of resources tailored to specific application requirements. Examples of PaaS solutions include Red Hat OpenShift and Google App Engine. Additionally, some platforms, like SAP BTP, extend existing Enterprise Resource Planning (ERP) functionalities with AI-based analytics or data-driven workflows.

Software-as-a-Service (SaaS)

SaaS service models represent the highest level of abstraction among cloud providers' offerings, providing ready-to-use applications accessible over web interfaces or APIs through a subscription-based or pay-as-you-use business model. The provider manages the full application stack, from logic and data to physical infrastructure.

The 2025 BetterCloud State of SaaS report indicates that U.S. companies use an average of 106 SaaS applications [12], reflecting the preponderance of this model across business functions. To overcome the complexity of managing a

growing number of SaaS services from multiple providers, referred to as "SaaS sprawl", organizations are increasingly adopting SaaS Management Platforms, which centralize governance and management of SaaS applications.
An emerging trend in the SaaS market is the integration of AI capabilities within enterprise applications. According to the same report, 95% of respondents have already invested in AI-related SaaS, with startups and mid-sized companies favored in the integration process due to their higher agility and reduced bureaucratic pressure. The AI SaaS market is projected to grow from a valuation of USD 71.54 million in 2023 to 775.44 million USD by 2031, with a Compound Annual Growth Rate (CAGR) of 38.28%, signaling strong momentum [13]. Typical SaaS offerings include general-purpose end-user services such as Google Workspace (e.g., Gmail, Google Docs, Google Drive) and enterprise-grade applications such as Salesforce, the leading Customer Relationship Management (CRM) system used for sales, marketing, and service automation.

Market Outlook

Overall, the global cloud computing market exhibits solid growth, with an estimated CAGR of 21% between 2018 and 2029 [14]. Despite SaaS representing the largest market size, it shows a slightly lower estimated CAGR from 2023 to 2029 compared to IaaS and PaaS, as highlighted in Figure 2.7.

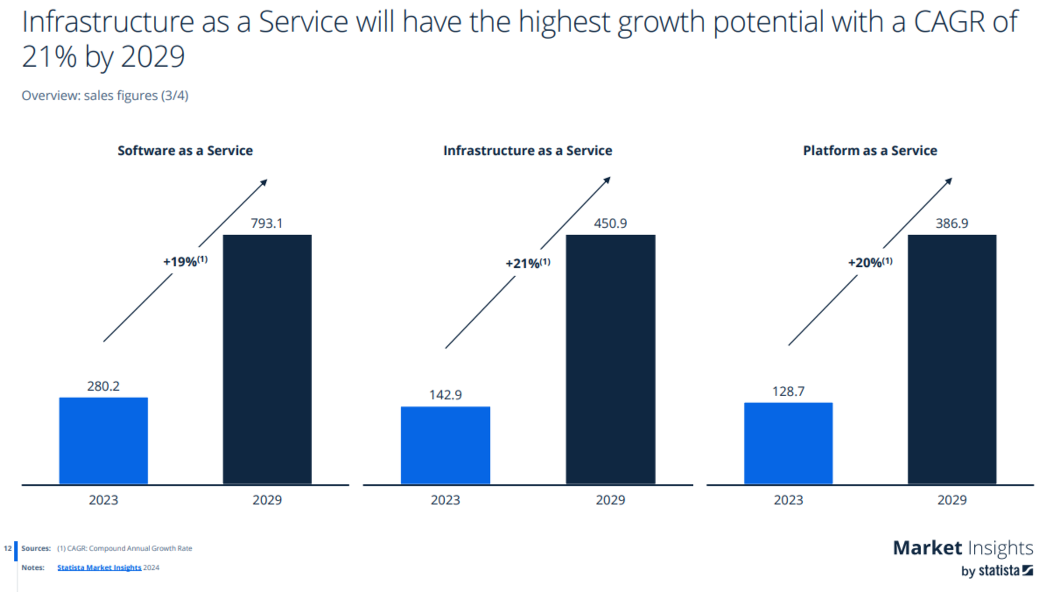


Figure 2.7: Comparison of cloud service models CAGR [14].

Emerging factors that will potentially further slow down the growth of SaaS services in the following years are the mentioned SaaS sprawl and the more recent developments concerning user programming [15]. AI-assisted coding (commonly referred to as “vibe coding”) and low-code/no-code development platforms, which enable end-users to build applications themselves using natural language, may shift part of the demand from SaaS solutions toward IaaS and PaaS environments.

2.2 Use Case Scenarios

To better understand the impact of different service and deployment models in terms of costs, flexibility, and control, this section provides two enterprise-grade use cases: the first one concerning the implementation of a data analytics platform in a hybrid cloud deployment, while the second one compares IaaS, PaaS, and SaaS solutions for cloud-based data storage.

2.2.1 Scenario 1: Hybrid Cloud

Consider a company that uses a CRM platform to manage terabytes of customer information on-premises. The dataset contains both Personally Identifiable Information (PII) and behavioral data, such as preferences, attitudes, activities, and feedback, which are particularly valuable for sales and marketing purposes. The organization wants to integrate big data analytics and machine learning models to classify customers, identify behavioral patterns, and build clusters that can support the GTM strategy of a new product. Processing this data requires significant computational resources that may not be available on the on-premises infrastructure. Instead of investing in expensive hardware, the company can leverage public cloud scalable resources. Since PII is subject to strict regulations, such as the General Data Protection Regulation (GDPR) in Europe, and because the CRM system is integrated with legacy on-premises workloads, the company must retain part of this data on the local infrastructure.

To successfully deploy the new analytics service, the organization can adopt a hybrid cloud strategy, processing anonymized customer data useful for the GTM analysis in the public cloud while retaining sensitive PII on the on-premises infrastructure. If the strategy is successful, the product launch will generate additional sales and new customer data. Thanks to the scalability of public cloud resources, the company is able to support this growth without incurring huge additional investments, as resources are provisioned on demand. Furthermore, if the on-premises infrastructure reaches its capacity limit, part of the workloads can be offloaded from the on-premises environment to the public cloud through the cloud bursting capability.

2.2.2 Scenario 2: IaaS vs PaaS vs SaaS

Table 2.1 shows the cost comparison between the three available solutions. For IaaS and PaaS models, the Azure Pricing Calculator [16] was used as a reference for cost estimation, based on the following assumptions and configuration:

- US Central was selected as a region because of price availability.
- Only fixed costs were considered; in the case of IaaS and PaaS services, additional charges for read and write or I/O operations should be considered for a more realistic estimate.
- To mirror OneDrive subscription terms, a one-year reservation discount was applied to IaaS and PaaS services.
- To avoid a single point of failure, the IaaS configuration relies on two VMs.
- Both IaaS and PaaS storage options were calculated using Locally Redundant Storage (LRS).

Table 2.1: Cost comparison among IaaS, PaaS, and SaaS solutions.

Option 1: IaaS	
Configuration	2 VMs
VM Price ($/unit-month)	217.08
Windows License ($/unit)	268.64
VM Cost ($/month)	971.44
Storage Capacity (TB)	200
Storage Cost ($/month)	3090
Total Monthly Cost ($)	**4061.44**
Option 2: PaaS	
Configuration	200 TB capacity
Hours per Month	730
Cost ($/TB-hour)	0.015
Total Monthly Cost ($)	**2190.00**
Option 3: SaaS	
Configuration	200 users
Subscription Fee ($/user)	5
Total Monthly Cost ($)	**1000.00**

As highlighted by the results, SaaS offers the most cost-effective solution for this scenario. Additionally, IaaS services introduce additional labor costs and operational complexity, as the organization must handle the setup, configuration, and maintenance of the underlying infrastructure. As the company operates lower in the stack, it must also provide the proper level of security and compliance.

Despite the higher costs and operational overhead, IaaS offers the greatest flexibility and enables the integration of the storage solution with existing workloads or on-premises environments. Furthermore, IaaS solutions enable the deployment and replication of resources across multiple cloud providers. To facilitate this process, the company should aim to develop a cloud-agnostic infrastructure, which minimizes dependency on a single vendor. This approach is favored by the adoption of open-source technologies and Infrastructure-as-Code (IaC) tools such as Terraform [17], which allow the definition of both cloud and on-premises resources in human-readable configuration files that can be versioned, reused, and shared.

While IaaS and PaaS configurations provide tools to manage policies, data residency, and run custom compliance controls, SaaS solutions typically offer less control over data and infrastructure, which makes the model less suitable for regulated or sensitive data such as PII, financial transactions, or government data.

To illustrate the trade-offs among the three service models, Table 2.2 provides a score ranging from 1 (low negative impact) to 5 (high negative impact) across multiple criteria such as cost, flexibility, complexity, control, and compliance, assigned to each solution, while the radar chart in Figure 2.8 provide a graphical representation of the comparison.

Table 2.2: Trade-off evaluation among IaaS, PaaS, and SaaS models.

Factor	**IaaS**	**PaaS**	**SaaS**
Cost	4	2	1
Lock-in	2	4	5
Complexity	4	2	1
Control	1	3	5
Responsibility	4	3	2

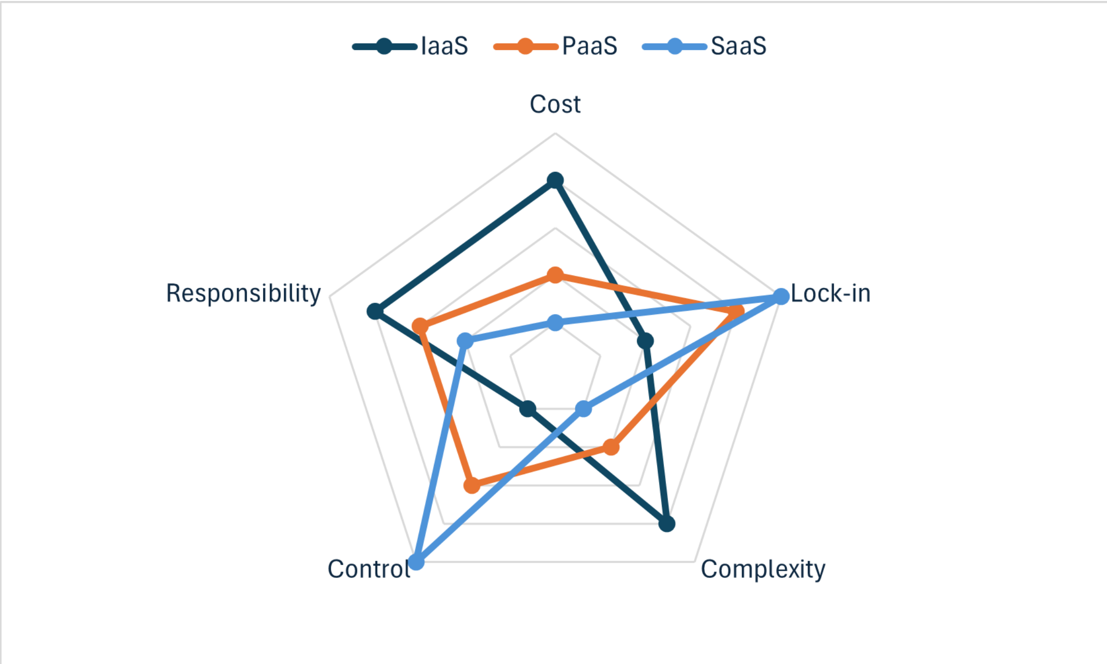


Figure 2.8: Multi-factor service model comparison.

Chapter 3

Strategy and Digital Transformation

This chapter presents a timeline of the most relevant literature and frameworks about strategy and innovation, and clarifies the role of IT in digital transformation. This theoretical perspective provides the basis for illustrating the dynamics of cloud ecosystems and understanding the sources of competitive advantage for firms operating in the industry.

Michael Porter defines strategy as "the creation of a unique and valuable position, involving a different set of activities" [18].

In the early stages of the proposed timeline, it is possible to identify two main perspectives with respect to the firm's internal resources and capabilities: the static view, which defines a firm as an organized association of complementary resources, and the dynamic view, which conceives the firm as a bundle of organizational routines and processes capable of adapting to environmental and technological changes. [19]

The timeline begins with the Competitive Forces approach, mainly derived from Porter's theories. According to this perspective, a competitive advantage is established by positioning products in a profitable industry, pursuing a generic strategy. The firm's internal organization is perceived as a set of activities that should be aligned with the predefined objective strategy to maximize value creation.

With the emergence of the Resource-Based View (RBV), the focus shifts from the industry structure and the product–market[1] fit to the internal resources and activities carried out by the firm, which are considered a potential source of sustainable competitive advantage. The dynamic view of the firm extends the RBV, considering the emergence of a competitive advantage as determined by the ability

[1]When a product successfully satisfies the needs of a specific market.

of the firm to adapt its capabilities to external changes, particularly in rapidly evolving environments.

More recently, the theories about platforms and ecosystems explored the dynamics of the modern IT industry to understand how a firm can maintain a relevant position within a turbulent and dynamic environment. The theory suggests that within those environments, a competitive advantage emerges through effective governance of the firm's ecosystem, which involves the orchestration of partners, users, and complementors[2]. In parallel, scholars investigated how the innovation process and innovation management impact the firm's competitive advantage and, more recently, how IT and digital transformation reshape the firm and ecosystem dynamics. The key milestones contributing to the theoretical background are highlighted in Figure 3.1 and reviewed in depth within this chapter.

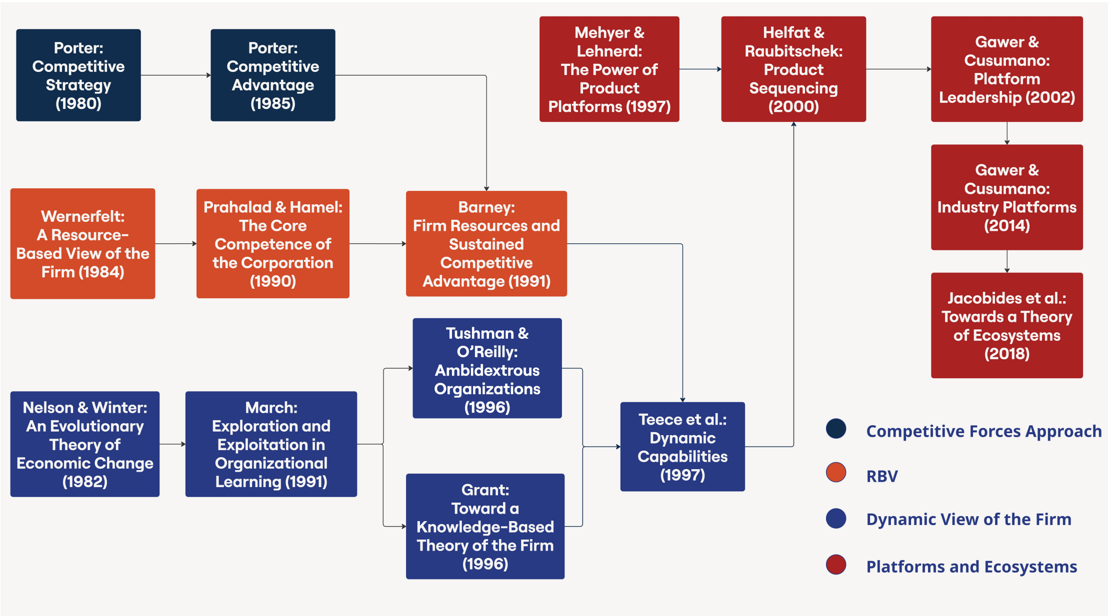


Figure 3.1: Timeline of strategic literature.

3.1 Competitve Forces Approach

In the book *Competitive Strategy* [20], Michael Porter developed the well-known Five Forces framework, analyzing industry structure and market dynamics to explore the nature of competition among incumbents.

[2]The developer of a complementary product: the value of two complementary products together is greater than the value of selling the individual products separately.

Porter proposed three actionable generic strategies for a firm to gain competitive advantage:

- **Cost leadership:** the firm competes on quantities and volume, exploiting lower unit costs. This strategy enables the firm to set lower prices with respect to competitors while retaining positive margins and acquiring a larger share of the market.
- **Differentiation:** based on offering a superior product with unique capabilities. By adopting this strategy, a firm is able to charge premium prices, as customers perceive higher value when buying the firm's product with respect to competitors.
- **Focus:** enables the firm to apply either cost leadership or differentiation within a limited market segment or niche.

Firms can be categorized according to their size and growth stage since their strategic priorities evolve over time. Generally, three main stages can be identified[21]:

- **Startups:** small companies in the early stage of development, usually focused on launching an innovative product aimed at finding the product–market fit.
- **Growth-stage firms:** organizations that have established product–market fit and seek to scale operations to acquire a larger market share.
- **Established firms:** mature organizations that have succeeded in scaling and need to preserve their position in the market.

Cost leadership is typically applicable to established firms, as competing on volumes allows them to reach the minimum efficient scale and fully exploit economies of scale, which, according to Porter, "always leads to a cost advantage for the large-scale firm over small-scale firms" [20]. In the Five Forces model, economies of scale are classified as one of the main factors favoring incumbents' positions against new entrants. Startups usually pursue focus or differentiation strategies, targeting market niches or proposing innovative and superior products. On the other hand, growth-stage companies often face organizational complexity and coordination costs, which "can lead to diseconomies of scale in a value activity as scale increases." For these reasons, scaling the business represents one of the most critical phases in a firm's lifecycle.

In *Competitive Advantage* [22], Porter shifts the focus to the activities performed by the firm, which determine the effective implementation of the objective strategy. The Value Chain Framework (see Figure 3.2) classifies the activities that generate value for the firm into primary and support activities.

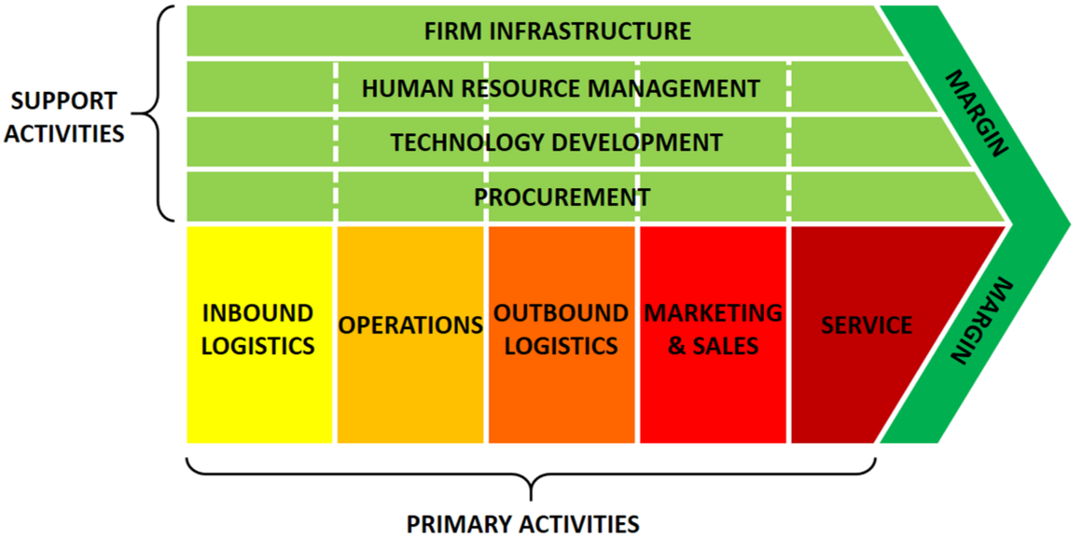


Figure 3.2: Porter Value Chain [23](adapted from [22]).

Primary activities are directly involved in product development and value creation for the firm. Support activities provide the necessary input to carry out the primary ones and can constitute themselves a source of competitive advantage.

Among support activities, technology development plays a critical role. Indeed, technological change can either create or destroy competitive advantage by reshaping learning curves and production processes. Consequently, each company should define a proper technological strategy aligned with the overall firm objective strategy. Additionally, Porter introduced the concept of timing and technological leadership: pioneering a new technology can lead to first-mover advantages, such as the ability to set standards and benefit from early profits, but it also involves a higher up-front investment (pioneering cost) and greater technological uncertainty.

3.2 The Innovation Process

The model proposed by Porter fits very well in describing competition dynamics within existing and relatively stable industries. However, it does not fully address situations in which technological change not only modifies the existing equilibrium among competitors but may also radically change the structure of industries and markets, as well as create new ones.

The concept of radical transformations of economic structures was first articulated by Schumpeter (1942), who described capitalism as a process of "creative destruction", an intrinsic characteristic of modern economics driven by what he categorized as revolutionary changes [24]. By contrast, evolutionary changes are small and incremental improvements that do not alter the fundamental structure of the economy. This distinction was later codified by scholars in the context of technology and product development, distinguishing between radical and incremental innovations, which refer to the nature of technological change without necessarily implying a shift in the competitive dynamics within an industry.

3.2.1 Innovation in Product Development

The concept of innovation is closely related to product development. The first relevant discussion on this topic was proposed by Abernathy and Utterback (1975), who investigated the relationship between innovation, competitive strategies, and production processes [25]. Their analysis revealed statistically relevant results showing a strong correlation between the innovation process and the firm's chosen competitive strategy. Figure 3.3 shows the pattern of relationships between the stage of product development and the rate of innovation.

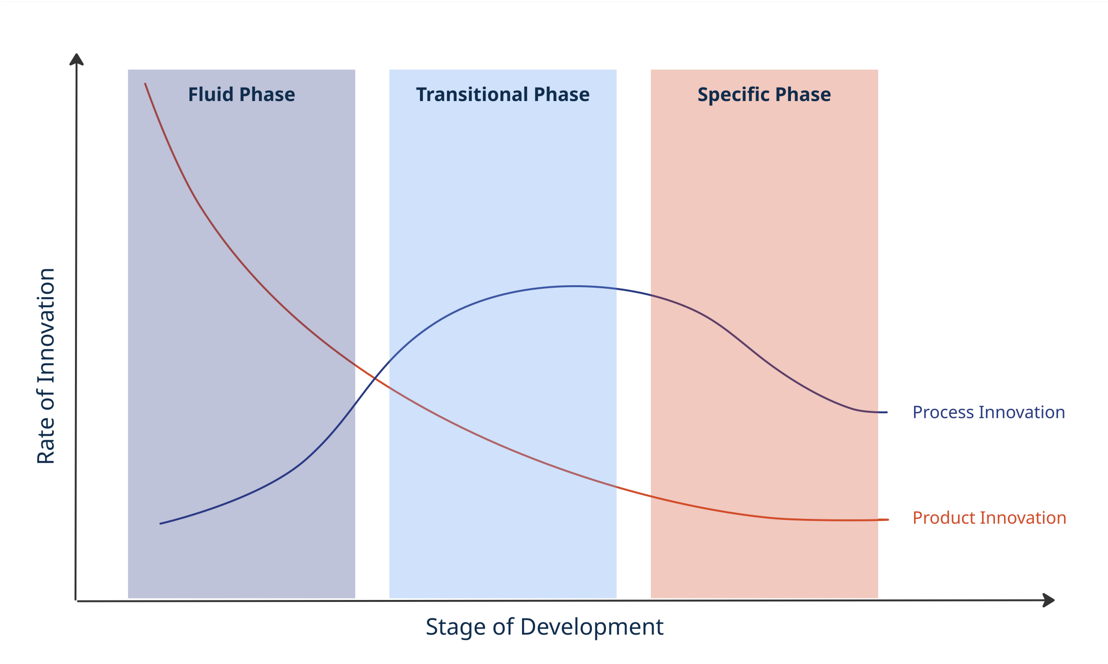


Figure 3.3: Process and Product innovation phases (adapted from [25]).

The correlations between the product lifecycle and the rate of innovation highlight three distinct phases [26]:

- **Fluid phase:** characterized by a high rate of product innovation, as multiple firms experiment and compete to find the dominant design. In this phase, firms are more likely to adopt differentiation strategies, maximizing perceived value and performance of the product.

- **Transitional phase:** the emergence of a dominant design slows down product innovation, moving the effort toward the optimization of production processes through the development of internal technical capabilities.
- **Specific phase:** in the last phase of the product lifecycle, the firm is able to define a systemic and standardized process that typically allows only small and incremental changes to the product. Standardization enables cost reduction and consequently economies of scale, leading firms to prefer cost leadership strategies with minimal effort on differentiation.

In *Reindustrialization and Technology*, the innovation process was described as a complex network of internal and external communication paths, linking the firm's capabilities with the market and technological community. Within the framework in Figure 3.4, the idea for a new product or technological development can originate from either a market need (market-pull) or a technological breakthrough (technology-push) [27].

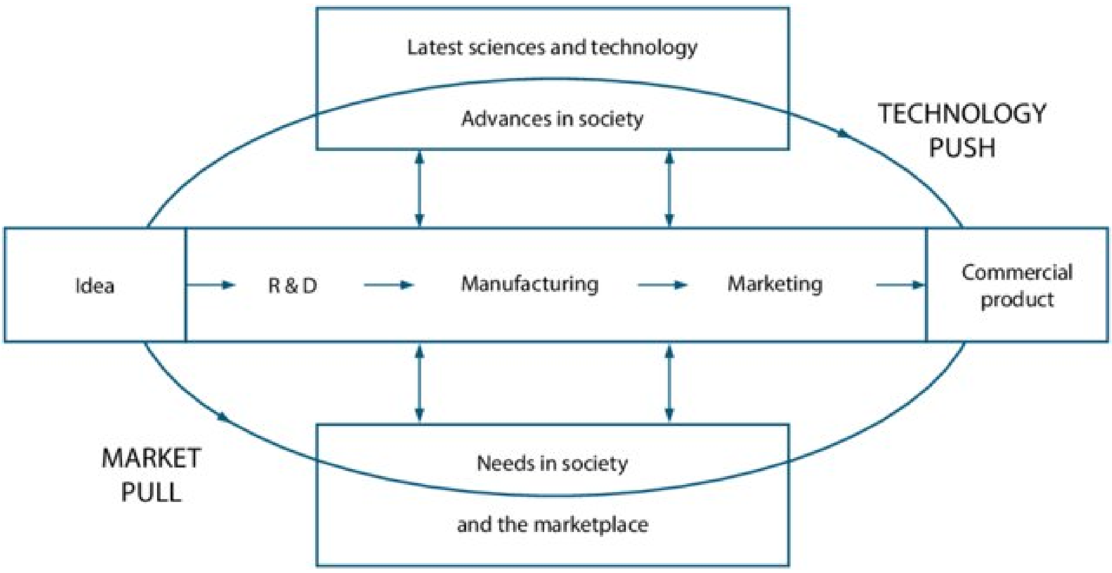


Figure 3.4: Technology Push vs Market Pull [28](adapted from [27]).

In industries characterized by a high rate of innovation, technology-push dynamics often anticipate market demand. When technological progress is driven by frequent and radical innovations, customers may not yet perceive the need for a new revolutionary product, instead suggesting incremental improvements to existing technologies. As Steve Jobs observed in an interview with *Business Week*: "A lot of times, people do not know what they want until you show it to them"[29]. Despite the absence of explicit market demand for the radical innovations introduced by the iPhone, such as the multi-touch interface and the App Store, the product turned out to be a transformative success for Apple, reaching approximately 15% of global shipment market shares by 2010, three years after its launch [30].

Managing Technology: A Strategic View (1989) emphasizes the importance of managing technology and innovation to establish a sustainable competitive advantage. The book provided actionable guidance for innovation management, highlighting that small and multifunctional teams can significantly accelerate innovation cycles and decrease time-to-market [31]. This approach is consistent with Conway's Law, which suggests that a product's design mirrors the communication structure of the organization that develops it [32]. Cross-functional communication and tighter feedback loops enhance agile innovation processes and promote modular architectures, which are particularly relevant in the context of software engineering and cloud-native development (see Section 4.4).

In *Architectural Innovation* (1990), Henderson and Clark proposed a framework (see Figure 3.5) for classifying different types of innovations, using the product as the primary unit of analysis [33].

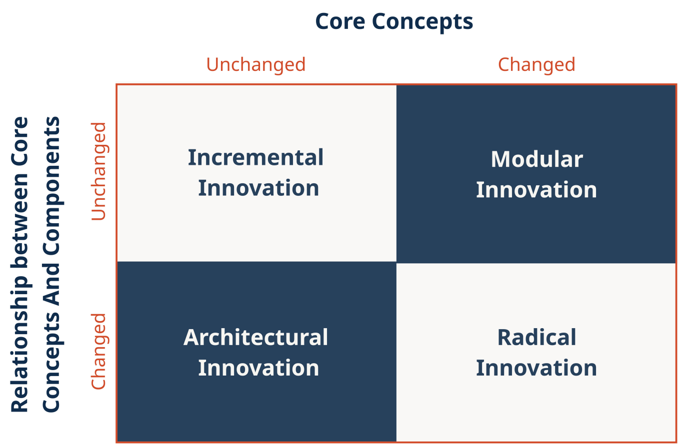


Figure 3.5: Classification of product innovations (adapted from [33]).

The distinction between radical and incremental innovation already existed, but the authors considered it incomplete. Therefore, they introduced two additional categories:

- **Modular innovation:** changes only the core design concepts of a technology.

- **Architectural innovation:** changes only the relationships between core concepts and components.

Architectural innovation, in particular, does not modify the core concepts of a product but can still create substantial challenges for incumbents, as it erodes the architectural knowledge embedded in their organizational routines.

3.2.2 Technology Diffusion and Innovations Dynamics

In *Diffusion of Innovations*, Everett Rogers provided a graphical and conceptual representation of innovation dynamics, framing the adoption of new technologies by individuals and corporations as a social process driven by communication over the network of adopters [34]. Figure 3.6 illustrates the number and distribution of adopters through a bell-shaped frequency curve and an S-shaped cumulative curve.

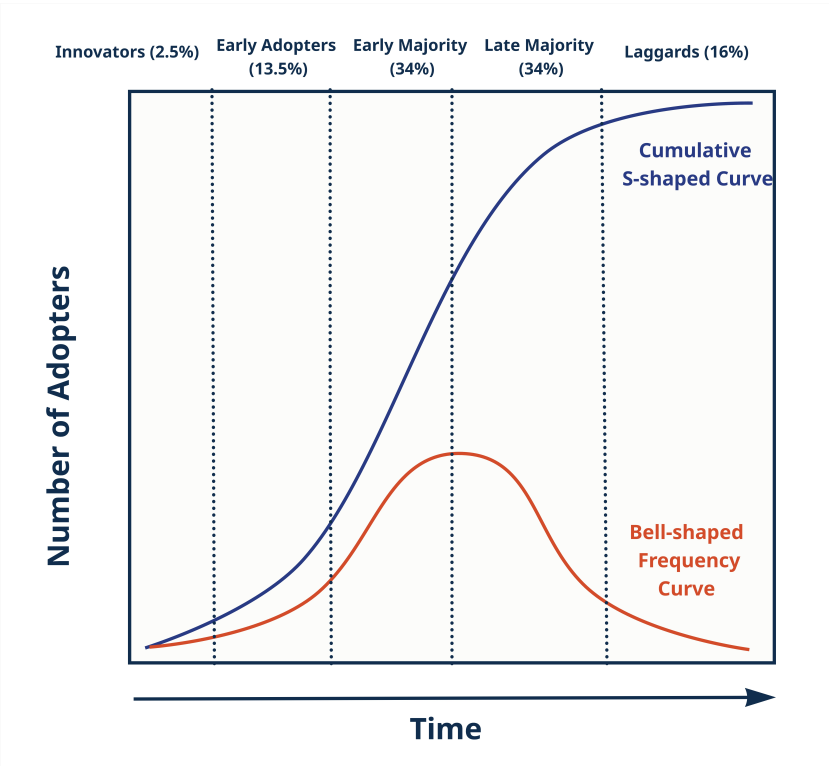


Figure 3.6: Innovation diffusion pattern (adapted from [34]).

New adopters can be classified into five categories:

- **Innovators:** a venturesome niche of customers with sufficient financial resources who act as gatekeepers, importing the innovation within the system boundaries.

- **Early Adopters:** a larger group in the society that provides the reference knowledge for future adopters.
- **Early Majority:** a larger group that prefers to adopt the technology when use cases and benefits are well defined.
- **Late Majority:** more skeptical users who adopt the new technology driven by network effects or economic needs.
- **Laggards:** the group of more conservative and traditional users who adopt the innovation only when it has become mainstream. Their adoption usually corresponds to the adoption of a new idea by innovators.

In *The Innovator's Dilemma* (1997), Christensen distinguished between sustaining and disruptive innovation, introducing a market and industry perspective on the impact of innovations. Sustaining innovations improve product performance without altering the competitive dynamics within the industry, while disruptive innovations initially underperform when evaluated using incumbents' criteria but introduce unique features that substantially enhance perceived customer value [35]. Indeed, according to Christensen, most disruptive innovations emerge in small or new markets with low profit margins. However, as customers recognize the additional value and demand grows, these innovations are able to displace established incumbents' revenues. For this reason, many firms suffered despite their expertise and qualified management. The author presented the case of Xerox, a market leader in photocopiers during the 1960s. The company was focused on the development of large and highly profitable photocopiers. When Canon introduced smaller personal copiers, they initially produced lower margins for the company; however, the demand for the new product was able to disrupt Xerox's profits over time.

In the field of innovation management, the strategic relevance of external knowledge emerged in the early 2000s. In *Open Innovation*, Chesbrough highlighted the need for firms to strategically manage external sources of knowledge, defining the new paradigm as "the use of purposive inflows and outflows of knowledge to accelerate internal innovation, and to expand the markets for external use of innovation, respectively" [36]. Open innovation enables firms capable of leveraging both external and internal knowledge as a way to mitigate the risk of disruption by radical innovations. However, to fully exploit the benefits of this approach, firms should be able to integrate both internal and external resources into coherent systems and architectures.

Chesbrough's paradigm shows strong correlations with later developments in platform and ecosystem theories (see 3.5), which stressed the importance of integrating external knowledge and resources within the firm's boundaries. In this context, a business model provides a framework to connect technical decisions with economic outcomes. According to the author, the primary functions of a business model include: defining the value proposition, identifying the market segment, describing the value chain structure, estimating margins and cost structure, defining the position within the value network and relationships with customers and suppliers, and finally, formulating a strategy to create and sustain a competitive advantage through effective innovation management [37].

3.3 Resource-Based View (RBV)

Studies about the innovation process have traditionally focused on product innovation and the evolution of technology. However, these perspectives alone cannot fully explain the sources of competitive advantage. In this direction, an alternative approach to Porter's theories was proposed by Wernerfelt (1984) through the RBV [38]. This model shifted the focus of strategic analysis from product-market positioning to the internal resources of the firm, aiming to explain how its competitive position changes over time. Following similar principles, Barney (1991) noted that the existing literature had focused on the analysis of the external environment, specifically markets and industries [39]. Figure 3.7 highlights the conceptual distinction between the resource-based model and antecedent environmental models of competitive advantage.

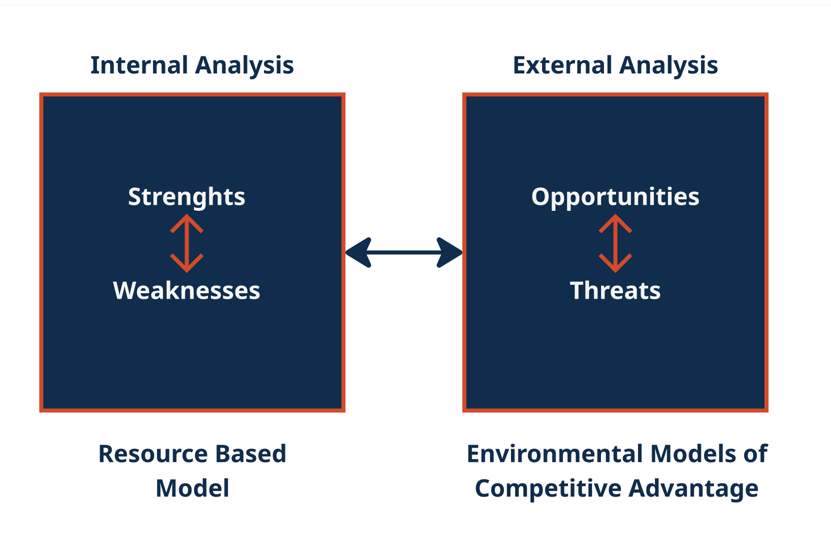


Figure 3.7: Comparison between RBV and competitive force appoach (adapted from [39]).

According to previous theories, competitive advantage derives from identifying a profitable position within an industry according to its threats and opportunities.

This approach implicitly assumes a similar set of strategically relevant resources for firms competing within the same industries. The author, instead, argued that to explain the competitive position of a firm, it is essential to analyze its idiosyncratic attributes.

Barney defined resources as "all assets, capabilities, organizational processes, firm attributes, information, and knowledge, etc., controlled by a firm that enable the firm to conceive of and implement strategies that improve its efficiency and effectiveness"[39]. To explain how some resources can provide a sustainable competitive advantage, he proposed the VRIN framework (later evolved into VRIO [40]: Valuable, Rare, Inimitable, and Organized; see Figure 3.8).

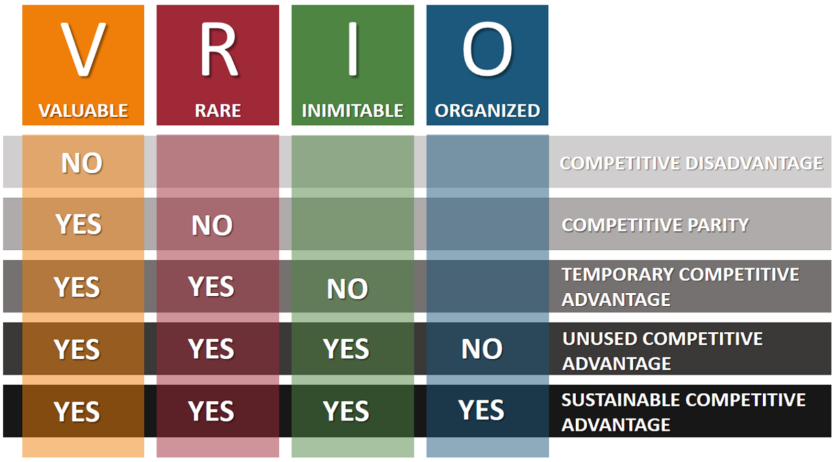


Figure 3.8: VRIO framework [41] (adapted from [40]).

The framework provides a valuable model to understand which resources should be prioritized by a company to sustain a competitive advantage. However, it remains limited in fast-moving environments with a high innovation rate, where it is challenging to define stable sources of competitive advantage.

Prahalad and Hamel (1990) made an important step in this direction, defining core competencies as the collective knowledge and integration capabilities that enable the firm to coordinate multiple production skills and technologies. The authors also introduced the concept of core products, which represent the physical embodiments of the firm's core competencies [42]. Core products should not be intended as end-customer goods but rather as components or subsystems that allow the firm to extract value from multiple end products across different markets, enabling economies of scale and scope.

Differently from Porter's view, where the strategic priority of the firm lies in finding a profitable market and then aligning the organization accordingly, the core competencies perspective focuses on building solid internal capabilities and identifying the markets in which they can be exploited to generate a sustained competitive advantage.

3.4 Dynamic View of the Firm

In its original formulation, the RBV assumes heterogeneous resources that are relatively stable over time. However, this assumption does not hold in turbulent and highly competitive environments, where valuable resources or competencies may be eroded by the emergence of a radical innovation. To address this limitation, scholars focused on the evolving dynamics of resources and capabilities within the firm.

The first roots of this dynamic perspective were proposed by Nelson and Winter (1982), who criticized the orthodox neoclassical economic view in which the firm is continuously adjusting price or quantity produced to reach equilibrium. Instead, the authors proposed a dynamic and "evolutionary" interpretation of the economic system, borrowing the term from biology and Darwinian selection. Following the analogy, organizational routines are identified as the firm's "genes." Therefore, the firm is considered path-dependent and capable of surviving market competition only if it can gradually adapt routines to technological and environmental changes [43]. Building on this perspective, March (1991) explored the trade-off between exploration and exploitation in organizational learning. Because the firm has limited resources, it must find an optimal allocation between the two categories of activities [44]:

- **Exploration:** involves experimenting and learning through innovative opportunities, which allows the firm to enter new markets. The risk associated with this set of activities lies in allocating resources to build competencies characterized by a high degree of uncertainty concerning long-term returns.

- **Exploitation:** in contrast, involves extracting value from activities and resources that already constitute a source of competitive advantage. However, this strategy may lead the firm into a competency trap, where it relies only on routines that were effective in the past without being able to adapt over time [45].

Figure 3.9 highlights the emergence of competencies with respect to the development of the firm's technological capabilities.

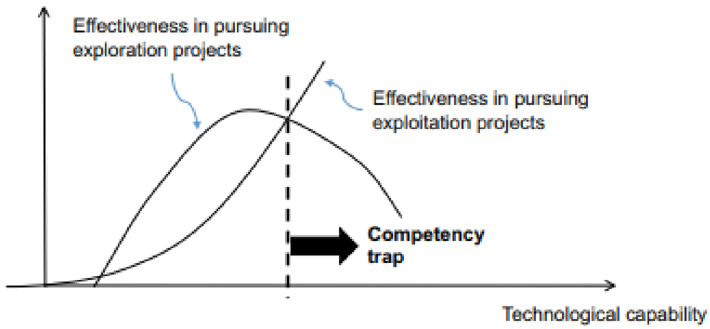


Figure 3.9: Competency trap [19].

When facing a competency trap, firms are more likely to be disrupted by radical innovations. The attribute of organizations capable of surviving both incremental and radical innovations is referred to as *ambidexterity*. Ambidextrous firms are characterized by distributed organizational structures with small and autonomous units and a mixture of different cultures, consistent with the optimal structure identified in [46] with respect to innovation management. At the same time, their overall size enables them to leverage economies of scale and scope.

In *Hypercompetition*, D'Aveni introduced the concept of *temporary advantage*, arguing that in fast-moving and hypercompetitive environments, it is not possible to aim for a sustainable competitive advantage. Instead, incumbents must continuously "leap" from one temporary advantage to another to avoid disruption [47].

In 1996, Robert Grant proposed the *knowledge-based view of the firm*, where knowledge represents "the most strategically important of a firm's resources". The firm is conceived as a "knowledge-integrating institution," coordinating and combining individual knowledge to create value [48]. The author distinguished between explicit knowledge, which can be codified and transferred among individuals, and tacit knowledge, which is embedded in organizational routines. In this context, information technology (IT) plays a critical role in aggregating explicit knowledge and facilitating its transferability. Competitive advantage arises when the firm is able to coordinate both explicit and tacit knowledge into inimitable organizational capabilities.

The importance of knowledge and organizational learning was further developed in the seminal paper on *Dynamic Capabilities* (1997). The authors defined dynamic capabilities as "the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments" [49]. This study represented a crucial step in the evolution of the RBV into a dynamic model, specifically addressing the challenge of understanding the sources of competitive advantage in high-technology industries such as semiconductors, information services, and software, where Time-to-Market (TTM) is critical and the competitive environment is difficult to predict due to constant technological changes. The authors illustrated their framework by comparing it with existing strategic perspectives, namely the Competitive Forces approach and the RBV [49].

The Competitive Forces approach is composed of three steps: finding an industry, choosing a strategy based on competitors, and acquiring the necessary assets. This perspective assumes a relatively straightforward acquisition of assets and does not consider the firm's unique internal resources and capabilities.

The RBV focuses on identifying unique resources, determining the market in which they can be effectively exploited, and extracting value either through direct participation, the sale of intermediary assets, or the transfer of resources to other firms. In this view, resources are sticky and heterogeneous, and organizational learning is not taken into account.

The dynamic capabilities framework, on the other hand, is based on three pillars:

- **Organizational processes:** involving coordination, learning, and reconfiguration of routines and resources.
- **Position:** built on the assets and the technological base of the firm.
- **Path-dependence:** as the firm's future trajectory is influenced by its past choices.

3.5 Platform Leadership and Ecosystems

Before introducing the theories, it is important to clarify the distinction between company or product platforms, which are the collection of assets a firm uses to develop a product, and industry platforms, which are products, services, or technologies that provide the foundation upon which an entire business ecosystem can develop complementary products, technologies, or services [50].

3.5.1 Product Platforms

The first codification of product platform methodologies was provided by Meyer and Lehnerd (1997). The authors emphasized that using a platform to build families of products that share commonalities such as architecture, processes, and components can lead to significant advantages in terms of value creation and cost efficiency [51].

One of the most relevant principles in the development of a product platform is *modularity*, which is particularly relevant for the software industry. This concept was later codified in *Design Rules: The Power of Modularity* [52], as a design principle that involves decomposing a complex system into smaller, functionally independent modules. Using this principle in product design enables both parallel innovation, allowing teams to develop and improve individual modules independently from the overall system, and interface standardization, which facilitates innovation on the core product while minimizing resources invested to manage the interconnections between modules.

The concept of modularity was later extended beyond product design to organizational structures. In the context of organizational modularity [53], modules are represented by the firm's business units. This decentralized configuration favors economies of scope and business unit autonomy, enabling them to experiment and innovate more rapidly while leveraging a pool of shared assets, technologies, and knowledge.

In 2000, Helfat and Raubitschek [54] proposed a dynamic framework to explain how a firm's knowledge base, capabilities, and products co-evolve over time. *Product sequencing* is the central mechanism of this process and describes how firms develop a path-dependent sequence of related products, whose shape is determined by the firm's existing knowledge and learning processes. Effective management of innovation across the sequence, as well as the strategic linkage among successive products, can constitute a source of competitive advantage over long time spans.

The framework in Figure 3.10 illustrates an interconnected cycle linking three core components: the system of knowledge, the system of learning, and product sequencing.

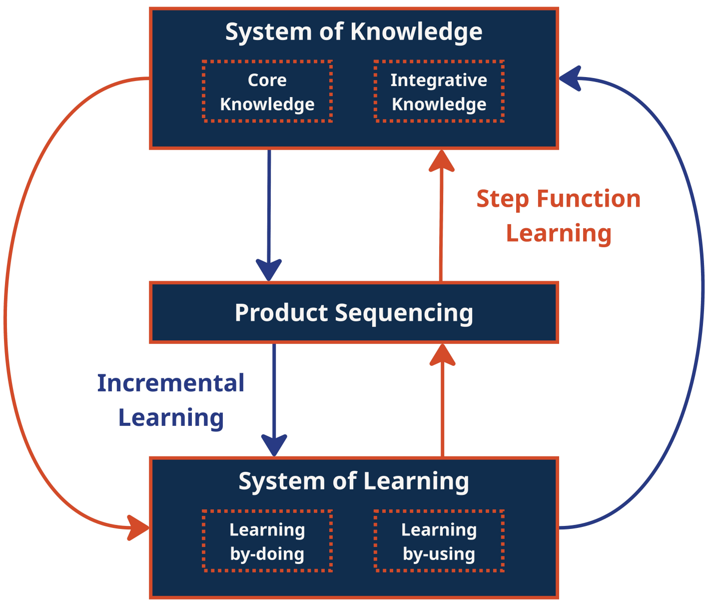


Figure 3.10: Product sequencing framework (adapted from [54]).

The System of Knowledge involves the firm's technological and organizational knowledge base, which includes both core and incremental knowledge.

- **Core knowledge** refers to the technical or scientific know-how that constitutes the foundation for a firm's product or service and is particularly critical in high-technology and similarly complex industries.

- **Integrative knowledge** reflects the firm's ability to coordinate and align different sets of activities (e.g., marketing, R&D, manufacturing) through a vertical chain, or by integrating multiple technologies in complex systems. In this context, IT is particularly effective in enabling integration and information flows across multiple stages of the vertical chain.

Product Sequencing describes the evolution of a firm's products over time. Each product in the sequence can concurrently rely on existing knowledge while creating opportunities for the learning process. The direction of product sequencing can either be *horizontal*, if knowledge and resources are leveraged across related markets or product lines, or *vertical*, when the firm's activities are linked within a vertical chain through multiple stages. *The System of Learning* enables two distinct cycles within the framework:

- **Incremental learning** occurs when the firm relies on existing knowledge to develop new products. This process generates feedback loops that can directly flow into the sequence to develop new products or refine the firm's existing knowledge base. During the product lifecycle, the firm can experience two distinct forms of incremental learning: learning "by doing" during the production process, or learning "by using" after the product is released and available in the market.

- **Step-function learning** arises when the firm must radically modify its set of core or integrative knowledge to develop new products. This learning process is usually enabled by environmental feedback, such as technological breakthroughs or benchmarking of competitors. Before being able to exploit the acquired knowledge base for product development, the firm needs to acquire such knowledge and assimilate new competencies, which is typically carried out through dedicated teams or business units.

The two processes are not mutually exclusive and usually occur in parallel. Effective management of product sequencing implies the constitution of two interrelated portfolios for products and knowledge, enabling the firm to pursue both exploration and exploitation strategies [44]. As firms develop successive products, the need to integrate knowledge across different stages of the value chain may lead to increased vertical coordination. Consequently, product sequencing is influenced by the firm's degree of vertical integration, which reflects the strategic decisions to "make" (produce a product or service internally) or "buy" (outsource a product or service to a supplier). These choices are explained by *Transaction Cost Economics (TCE)*, which focuses on minimizing transaction and coordination costs [55]. High levels of asset specificity and uncertainty imply higher transaction costs and greater risks of supplier dependency. Additionally, the firm's supply chain should be aligned with product characteristics to ensure effective governance and coordination across both internal and external activities involved in the development and distribution of the product [56].

3.5.2 Industry Platforms

Within industries that produce complex technological products composed of various interdependent components, platform products and modularity are essential for enhancing production process efficiency and enabling faster innovation. These principles enable parallel innovation, which may originate from the firm's internal business units but also from suppliers and external partners to whom specific modules are outsourced or purchased.

In 2002, Gawer and Cusumano introduced the concept of *platform leadership* to explain how firms such as Microsoft and Cisco preserved their leadership positions while managing an extensive network of suppliers and external interdependencies[3]Gawer and Cusumano [58]. Their strategy relied on establishing themselves as the central platforms for which external actors deliberately produce complementary products and services. Firms pursuing similar strategies must manage network externalities[4] in addition to their internal processes to extract the maximum value from their products.

From an economic perspective, Rochet and Tirole [59] proposed a model to describe platform competition, focusing on markets characterized by a strong influence of network externalities. In particular, *two-sided markets* emerge when an organization provides the central platform for two distinct user groups, each benefiting from the participation of the other. A typical example is game consoles, which serve as the common platform for both end customers and game developers. Differently from multi-product markets, for firms operating in two-sided markets the profit-maximizing price is not symmetric between the two user groups. Instead, platform leaders often subsidize one side to maximize joint participation and extract value from the other.

[3]A modern example of this strategy is Apple, which relies on 187 major suppliers for its products as of 2023 [57].

[4]A network externality occurs when the value or utility of a good increases in parallel with the number of users.

For instance, in the video game industry, consoles are typically sold at a relatively low price to expand the user base while generating revenues through royalties and licensing fees from game developers. A similar dynamic also applies to open source software, which will be further explored in Subsection 4.3.1.

Platform leaders benefit from significant entry barriers, including network externalities and customer switching costs. However, their advantage can be challenged through two strategic mechanisms:

1. A revolutionary technological change that introduces superior functionalities.

2. *Platform envelopment*, where a competing firm embeds the functionalities offered by an incumbent's platform into its own ecosystem [60].

A well-known example of platform envelopment occurred when Microsoft incorporated the streaming media services provided by RealNetworks into its operating system, launching Windows Media Player. RealNetworks was operating as a two-sided market, distributing the software free of charge to end users while monetizing it through audio and video providers.

The approach of absorbing the functionalities of other platforms is referred to as *bundling*. In this strategy, the attacking platform offers a bundle that incorporates the functionalities of a rival's platform through targeted pricing strategies, which depend on the degree of user overlap between the two platforms. When the platforms present a strong customer base overlap, the attacker can set a price close to the combined value of both platforms, as the user is more likely to prefer accessing all functionalities within a single, integrated platform. Figure 3.11 provides a graphical representation of the process.

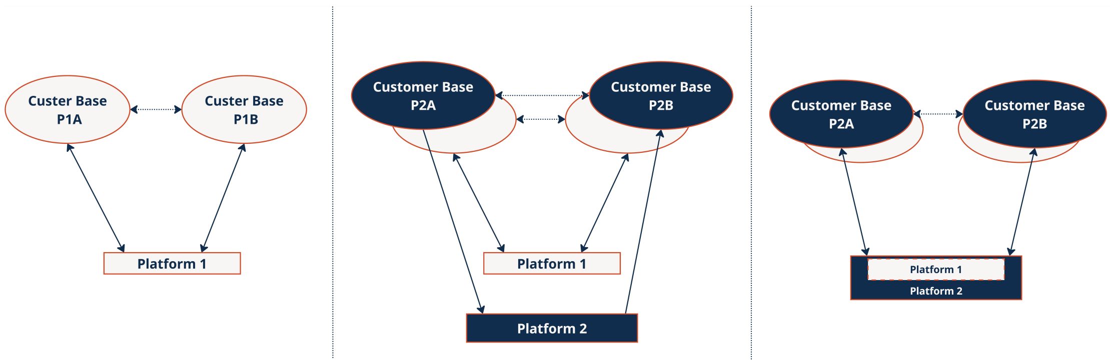


Figure 3.11: Platform envelopment process (adapted from [59]).

If successful, this strategy allows the attacking firm to benefit from both economies of scope, when the absorbed platform shares common components with the attacker's base platform, and economies of scale, resulting from increased network externalities as the new customer base is approximately the combined size of the two platforms.

The concept of two-sided platforms was subsequently expanded into *multi-sided platforms*, which enable direct interactions between two or more distinct groups affiliated with the platform [61].

In 2014, Gawer and Cusumano provided a connection between platform leadership strategies and innovation management. Their focus was dedicated to *industry platforms*, where a technology or product developed by a firm becomes the foundation for complementary innovations within a broader business ecosystem [50].

According to their distinction, industry platforms share similar characteristics with product platforms, such as the presence of common components and architectures, but differ in their degree of openness. The interfaces of an industry platform should be sufficiently open to enable complementors to *plug in* their products or components and foster *open innovation*, yet sufficiently closed to enable the focal firm to capture the most value and preserve its profits. These platform governance decisions are crucial for the firm to maintain a sustainable competitive advantage.

The authors proposed the example of Intel, which evolved from being a microprocessor producer into an industry platform leader for personal computers. The most important decisions enabling this transformation were: determining which complements to produce internally, defining the degree of modularity and openness of the platform, managing relationships with external complementors, and designing an effective organizational structure.

For instance, Intel established internal laboratories to strengthen its technological leadership within the industry, developing industry standards that enabled the company to improve technological capabilities while concurrently stimulating competition among complementors.

Platform leadership, despite representing an important source of competitive advantage, does not make the focal firm immune to the innovator's dilemma and disruptive technological changes [35].

3.5.3 Ecosystems

Industry platforms foster an ecosystem where innovation, knowledge, resources, and capabilities co-evolve among complementors. In 2018, Michael G. Jacobides and Gawer [62] provided a description of the attributes characterizing an ecosystem, the conditions in which it emerges, and how the focal firm creates and captures value. The authors defined ecosystems as "a group of interacting firms that depend on each other's activities" [62]. Their framework integrates three complementary perspectives:

- **The business ecosystem view**, which analyzes the ecosystem from the perspective of the focal firm (or hub) and the surrounding community of external entities.
- **The innovation ecosystem view**, which focuses on the system of innovations that enable customers to use end products, supported by a set of components and complements.
- **The platform ecosystem view**, which consists of "an array of peripheral firms connected to the central platform via shared or open-source technologies and/or technical standards (which, for IT-related platforms, can be programming interfaces or software development kits)."

The key characteristics enabling the formation of an ecosystem are modularity, which enables different producers to develop interdependent components, and the ecosystem structure, which is non-hierarchical and differs from market-based value systems or vertically integrated supply chains, as highlighted in Figure 3.12.

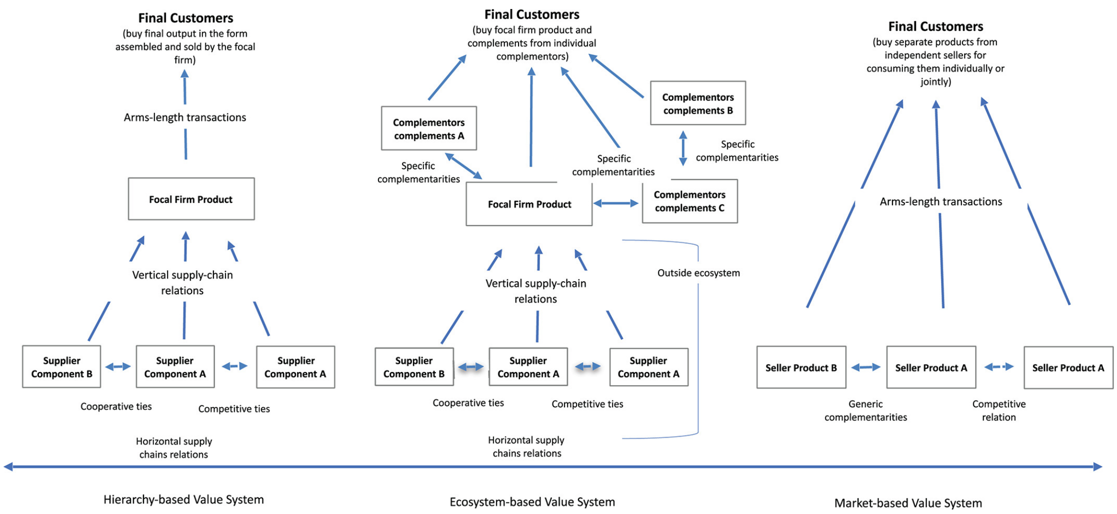


Figure 3.12: Ecosystem structure and value creation model [62].

Within an ecosystem, end customers can select among complementary components, such as a mobile phone user choosing which applications to install. Although ecosystems are characterized by non-hierarchical structures, they still require a degree of governance from the perspective of the focal firm to maintain alignment and coordination among participants.

An effective ecosystem strategy requires the focal firm to define rules for components and interfaces to coordinate complementor activities. For instance, in IT ecosystems, the definition of specific APIs imposes nonfungible investments on complementors, generating a platform lock-in. The strength of an ecosystem lies in its ability to achieve coordination between all participants and complementors without the need for vertical integration, since their performance depends on the success of the platform.

Notably, not every ecosystem emerges as the result of strategic decisions. Sometimes, a firm that introduces modular architectures can unintentionally create an ecosystem. In these cases, ecosystem evolution is influenced by the competitive dynamics and changes affecting competitors' ecosystems (such as iOS adapting to Android's growth), which remains an open field of research.

3.6 The Role of IT in Digital Transformation

In the proposed timeline, it becomes evident that technology plays a crucial role in explaining how firms develop strategies capable of adapting to rapid environmental changes. In particular, the role of IT gained significant strategic relevance from the early 2000s, as reflected in the formalization of platform and ecosystem theories.

With the advent and proliferation of internet connections, scholars increasingly recognized the need to explore the strategic and organizational implications of IT, whose functions can be described as follows [63]:

- **Automate:** substitutes human labour through automation.
- **Inform-up:** supports the management decision-making process.
- **Inform-down:** enables sharing of information and knowledge with employees.
- **Transform:** changes processes and business models of the firm.

Because of the impact IT has across all organizational functions (e.g., HR, marketing, sales), it is considered to have a cross-functional role, emphasizing the importance of developing digital business strategies [64]. Digital technologies enable business transformation across all industries, beyond the technology sector, leading to enhanced efficiency and competitive advantage for firms capable of becoming "digital masters" [65].

Digital transformation can be defined as "a process that aims to improve an entity by triggering significant changes to its properties through combinations of information, computing, communication, and connectivity technologies" [63]. Digital transformation is a process of continuous change and disruption, conceptually aligned with the theory of dynamic capabilities, as both emphasize the need for firms to adapt to new environments through the reconfiguration of routines and resources. Figure 3.13 presents a framework outlining the fundamental building blocks of digital transformation.

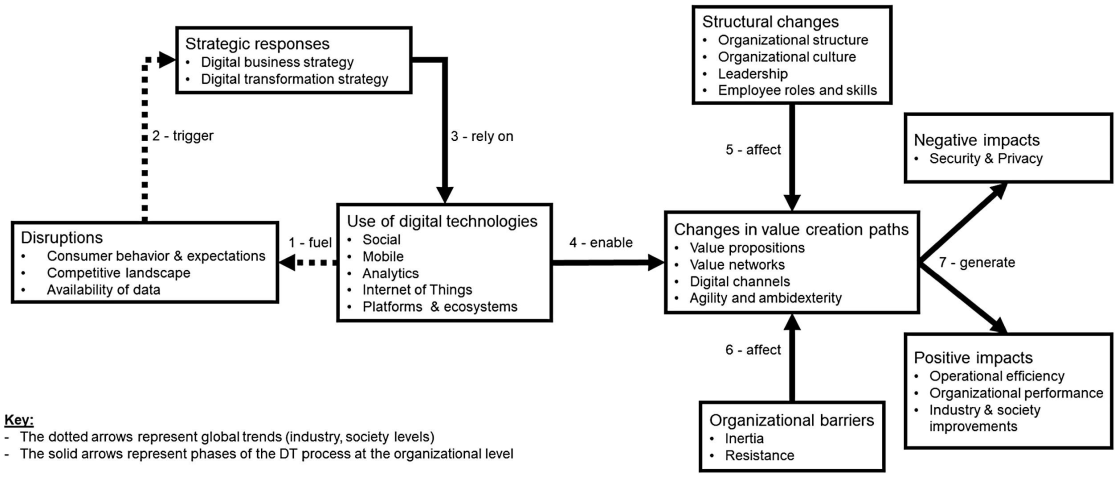


Figure 3.13: Inductive framework of digital transformation [63].

Although the literature remains fragmented regarding the co-evolution mechanisms and interactions between innovation management and digital transformation, a strong connection between the two fields is evident. The disruptive nature of digital transformation forces firms to accelerate the process and rapidly adapt to avoid being displaced by radical innovations in their industries [66].

This dynamic is observable with emerging technologies such as the Internet of Things (IoT), Artificial Intelligence (AI), and Big Data, which require firms across all industries to undertake the process of digital transformation. The paradigm that foresees the realization of digital transformation within an industry, enabling smarter and more efficient processes, is called *Industry 4.0* [67].

These technologies not only demand increased data storage and computational capacity, but they are fundamentally reshaping firms' business strategies by modifying how resources, routines, processes, and knowledge are managed. Furthermore, in the context of ecosystems and multi-sided platforms, digital transformation enables new firms to join digital platforms, accelerating the innovation process through collaboration with complementors, often encouraged and orchestrated by the focal firm.

Chapter 4

Cloud Strategy

Cloud computing can be considered a key enabler for the distribution of digital technologies across organizations [63]. A systematic literature review of peer-reviewed sources [68] published between 2014 and 2024 identified several factors illustrating the impact of cloud computing adoption on the digital transformation process, as shown in Figure 4.1.

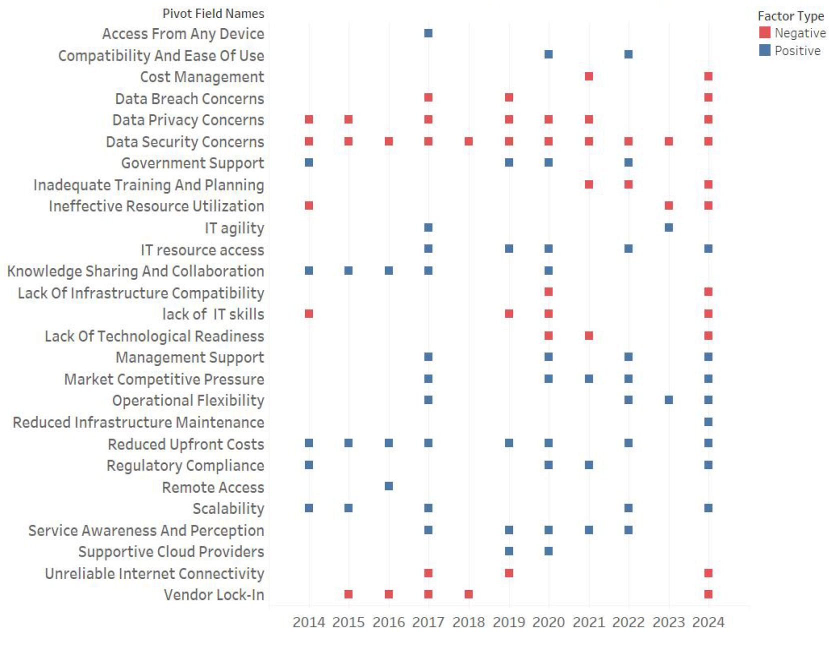


Figure 4.1: Positive (blue) and negative (red) factors influencing the adoption of cloud computing from 2014 to 2024 [68].

As highlighted by the study, cloud migration does not automatically generate tangible benefits. On the contrary, it can worsen costs, security, and performance if not supported by a robust and carefully designed strategy [69]. Because of its technological and strategic relevance, cloud computing can be interpreted as a coevolution of computing technology and business models to deliver digital resources [70]. This chapter explores the role of cloud computing at a macro level, focusing on how providers manage digital platforms and ecosystems to gain competitive advantage, and what managerial implications and opportunities arise for firms adopting cloud services. The analysis primarily focuses on IaaS and PaaS providers, as these models offer firms the highest degree of flexibility in terms of managerial and strategic choices.

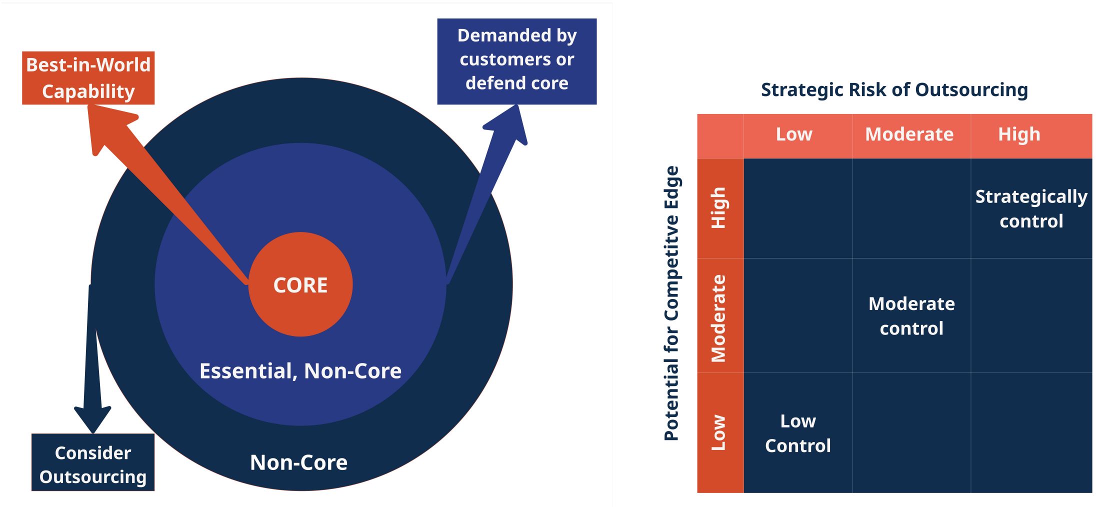
4.1 IT Outsourcing

One approach that firms can leverage to facilitate their digital transformation journey is IT outsourcing, defined as the organizational arrangement through which a firm manages IT services and resources provided by an external partner to support its business goals [71]. Outsourcing not only offers potential cost advantages but also represents a relevant strategic decision involving the firm's knowledge and capabilities.

This practice enables firms to leverage open innovation and access external knowledge and expertise, particularly within an ecosystem characterized by deep interdependencies among complementors and providers. Outsourcing can also stimulate internal innovation, allowing firms to focus resources on their core competencies. However, it also implies risks related to the lack of control over the outsourced activities and dependency on the provider, commonly referred to as *vendor lock-in*. For these reasons, companies must develop governance strategies to understand which activities and functions should be outsourced [72]. As illustrated in Figures 4.2a and 4.2b, companies should avoid outsourcing activities that provide a competitive edge, particularly those leveraging the firm's core competencies. The outsourcing process requires a structured decision-making and implementation process, guided by the following considerations:

- **Why:** Evaluation of the strategic motivations, advantages, and risks associated with outsourcing.
- **What:** Identification of processes and activities that should be outsourced.
- **Which:** Definition of procedures, guidelines, and metrics involved in the decision-making process.
- **How:** Structure of contracts and relationships with the provider.

- **Outcomes:** Evaluation and analysis of the results from the outsourcing decision.



(a) Representation of core competencies. **(b)** Outsourcing strategic guidelines.

Figure 4.2: Strategic Outsourcing (adapted from [72]).

4.1.1 Cloud Migration

Cloud computing emerged as the dominant model for IT outsourcing across different service layers (IaaS, PaaS, and SaaS), and as a fundamental enabler of digital transformation [73], serving as the technological backbone of various industries worldwide. From 2011, with the introduction of the *Federal Cloud Computing Strategy* in the United States [74], governments started to promote a cloud-first approach, prioritizing cloud solutions over traditional IT strategies. Europe adopted a similar position, identifying cloud as a key enabler for the *European Commission Digital Strategy* [75].

Through cloud migration, companies can seamlessly integrate emerging technologies on cloud environments while leveraging providers' expertise and computational resources. In the manufacturing sector, IoTs sensors collect data that are stored in the cloud, where AI-driven analytics services can use providers' resources to process them and enable predictive maintenance. In financial services, hybrid cloud architectures support the development of fraud detection systems, allowing the processing of vast volumes of transactional data on the provider's infrastructure, while maintaining sensitive customer data in the on-premises environment. In the automotive industry, data spaces facilitate the secure exchange and processing of sensor and telemetry data among manufacturers and suppliers, enhancing collaboration across the value chain. These examples highlight the critical role of well-designed cloud strategies in driving efficiency, innovation, and compliance across multiple industries and regions.

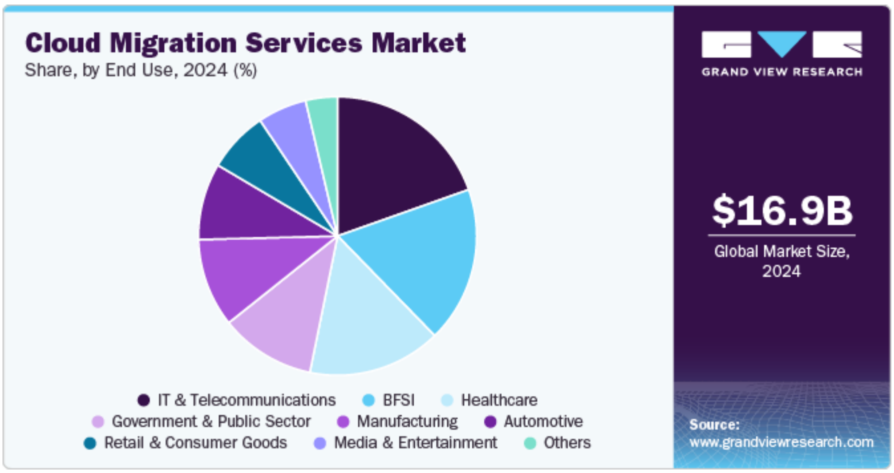


Figure 4.3: Distribution of cloud migration services market by sector [76].

The *Technology-Organization-Environment (TOE) Framework* provides a holistic view for analyzing the adoption of new technologies within organizations [77]. The framework, shown in Figure 4.4, identifies three critical dimensions influencing the adoption decision, namely: technological, organizational, and environmental contexts.

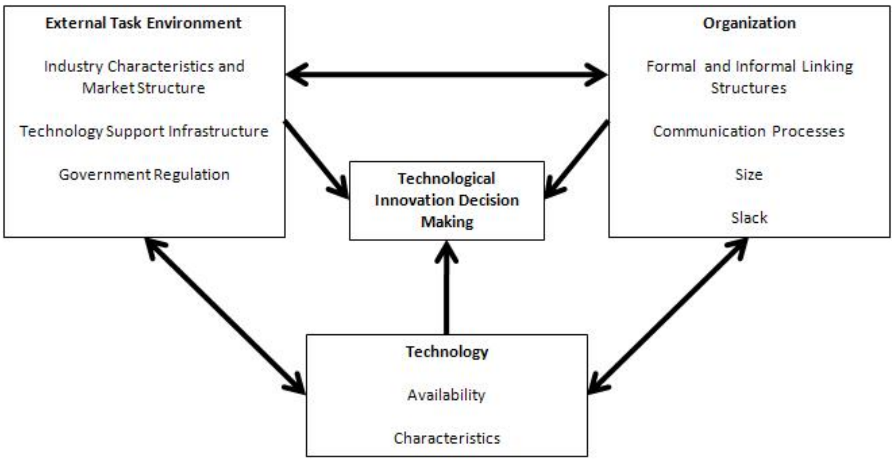


Figure 4.4: The TOE framework for technology adoption.

In the case of cloud computing, the technological context concerns the firm's infrastructure and its ability to integrate cloud services with existing legacy systems to determine the most suitable deployment or service model [69]. The organizational context includes the firm's internal activities, knowledge, and resources that are either involved in the process or outsourced when implementing cloud solutions. It also includes the organizational structure, which may involve provisioning dedicated cloud environments or teams to manage innovation more effectively. Finally, the environmental context relates to external factors such as the industry structure, market positioning, regulatory environment, and ecosystem dynamics. These factors should be carefully evaluated when planning the migration to cloud environments and will be examined in greater depth later in the course of the chapter.

4.1.2 Benefits of Cloud Computing

From an economic perspective, cloud computing enables a shift from capital expenditures (Capital Expenditures (CapEx)) to operational expenditures (Operational Expenditures (OpEx)). CapEx refers to long-term investments in physical or digital assets made by the firm to generate value in the future. These expenditures are listed on the balance sheet and are subject to amortization. On the other hand, OpEx covers day-to-day operations and is recorded on the income statement, allowing firms to reduce their net profit, thus generating a tax benefit.

This shift can be particularly advantageous for a wide range of organizations. For startups and Small and Medium-sized Enterprises (SMEs), it enables the setup of high-performance compute clusters through resource pooling, while for established companies, it allows faster learning and more effective experimentation, without committing additional resources to IT infrastructure. A 2015 case study [78] compared the investment required for an on-premises IT infrastructure with an equivalent Amazon Web Services (AWS) IaaS environment, demonstrating significant cost reduction and greater flexibility for SMEs. Specifically, the study reported an investment of approximately €90,000 for the on-premises setup, while an equivalent AWS deployment incurred a total cost of roughly €58,000 over a three-year period.

Moreover, cloud computing allows firms to scale digital services elastically in the case of fluctuating demand. This capability is particularly valuable for startups and growth-stage companies, as it supports economies of scale without requiring additional investments in physical infrastructure. Lower infrastructure costs also reduce barriers to entry and enable better resource allocation across the value chain, with more capital employed in primary activities.

Concerning organizational complexity, cloud computing simplifies infrastructure management by outsourcing the configuration, maintenance, and monitoring of IT infrastructure and services to providers. Given the rapid and standardized provisioning of cloud resources, DevOps teams can deploy virtual instances or testing environments almost instantaneously, choosing the preferred OS and configuration.

The standardization of infrastructure services transforms IT infrastructure into a commodity [79], shifting differentiation to higher levels of the technological stack, such as PaaS and SaaS, or to interoperability capabilities.

For DevOps teams, serverless computing offers a convenient approach that completely abstracts infrastructure management. While PaaS provides a platform that is always active and requires a manual configuration of resources, serverless architectures operate on an event-driven model in which resources are instantly and automatically scaled by the provider [80]. Figure 4.5 illustrates the comparison between these two models.

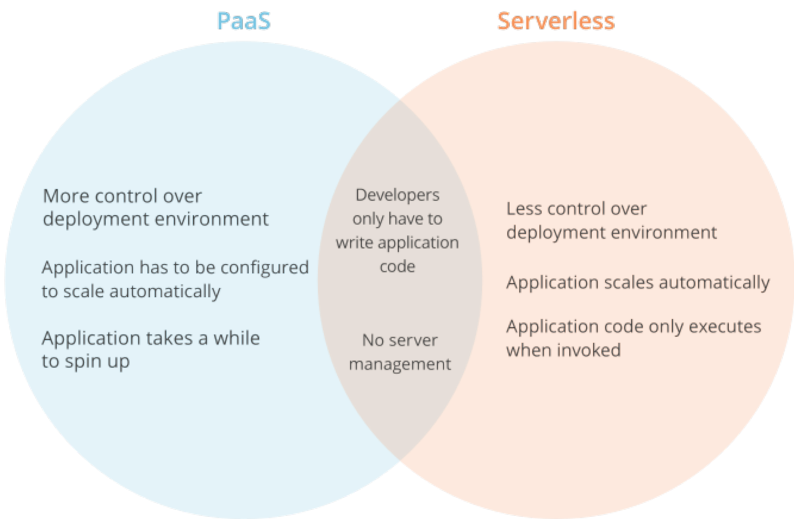


Figure 4.5: Comparison between PaaS and serverless computing models [80].

Serverless computing is typically implemented through Function-as-a-Service (FaaS), which enables the event-driven execution of application code. This model offers significant benefits in terms of cost efficiency, scalability, and flexibility compared to traditional platforms but implies less control and environment customization.

Beyond economic and organizational advantages, cloud adoption presents substantial managerial and strategic implications. As anticipated in Section 4.1, outsourcing IT infrastructure allows firms to focus on core competencies and innovation, minimizing time and resources involved in non value-generating activities. Organizations lacking the proper in-house expertise or computing capacity can delegate IT functions, such as database administration or network management, to providers. For organizations seeking both support and control, adopting managed IaaS services provides an effective balance, as it allows leveraging providers' expertise for infrastructure management and configuration, while retaining control over strategically relevant operations.

Additionally, cloud computing enhances organizational ambidexterity through faster testing environments and shorter innovation cycles. This allows firms to simultaneously pursue exploration strategies while exploiting their core competencies. Such flexibility is particularly valuable for firms operating in turbulent and dynamic environments, where adaptability and agility are crucial factors to establish a competitive advantage. The paradigm substantially reduces the time-to-market (TTM), particularly for software and digital solutions. For startups, this translates into a faster discovery of product-market fit, while for established firms, it fosters the development of dynamic capabilities that sustain innovation management and can prevent disruption as new technologies emerge.

For startups, innovation is a mission-critical factor and is often easier to achieve due to fewer bureaucratic constraints, while established enterprises frequently struggle to pursue effective innovation strategies. In the technology sector, many large enterprises "have already embarked on a slow death spiral" [21], with the exception of those capable of ensuring consistent product innovation. However, for all firms regardless of their size and growth stage, it is essential to align IT and cloud strategies with long-term business goals to fully exploit the benefits offered by cloud computing.

4.2 Hyperscalers and Competitive Advantage

The current cloud computing market is dominated by AWS, Microsoft Azure, and Google Cloud, which together account for approximately 63% of global market share for IaaS and PaaS services, as shown in Figure 4.6.

These data indicate that the cloud computing market structure for infrastructure and platform services exhibits oligopolistic characteristics [81], with a few dominant providers controlling most of the market share and computational resources, and with high barriers limiting participation for new entrants. In the context of IT infrastructure services, these firms are commonly referred to as *hyperscalers*, as their infrastructures are designed to process, store, and transport vast volumes of data efficiently, operating in more than a thousand globally distributed and large-scale data centers [82]. Their ability to scale capacity and resources enables them to meet the growing global demand and serve billions of users worldwide.

The high-tech industries in which these firms operate are characterized by hypercompetition and continuous radical innovation. The success of Amazon, Microsoft, and Google can be attributed to their ability to develop dynamic capabilities and implement platform leadership strategies. Despite their scale and complex organizational structures, these companies are able to continuously deliver cutting-edge solutions driving the diffusion of new technologies.

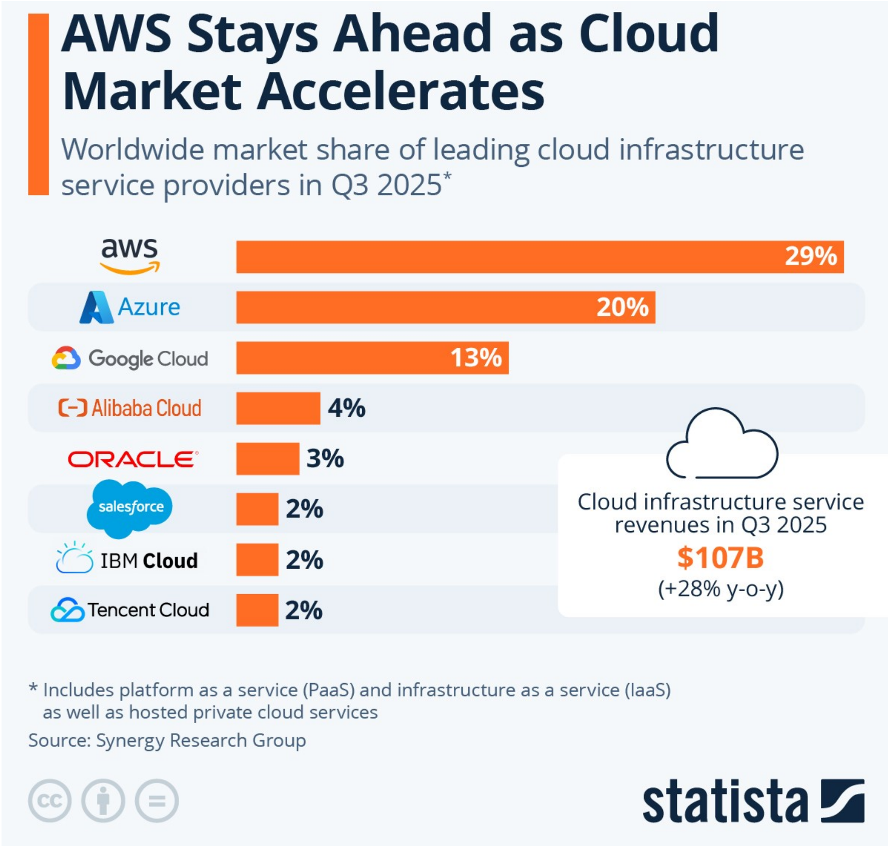


Figure 4.6: Cloud computing market shares [83].

Hyperscalers capture the largest share of value across the cloud services value chain. Through virtualization, multi-tenancy, and resource pooling, they achieve significant economies of scale with extremely low marginal costs. Their competitive advantage stems from a clear strategic vision and the effective governance of digital platforms and business ecosystems.

Despite the technologies enabling cloud computing have existed for several years, Amazon was the first to successfully commercialize infrastructure services over the internet, launching Amazon S3 (Simple Storage Service) and EC2 (Elastic Compute Cloud) in 2006. Before the public launch, the company faced an internal need to scale its infrastructure to support a rapidly growing volume of e-commerce operations. To address this challenge, Amazon developed internal tools and standardized APIs, providing reusable modules that improved the efficiency and scalability of its internal platform. The firm's management recognized that operating and scaling large-scale IT infrastructures was a common pain point across industries and that Amazon's internal solution could be transformed into a viable business model to deliver computing capabilities to other firms.

Currently, AWS accounts for approximately 15% of the company's total revenue and 50% of its operating income [1] [84], holding about 30% of the global cloud computing market share for IaaS and PaaS services [83]. The high profit margins generated by AWS support the cost leadership strategy in retail and e-commerce segments, which are able to maintain competitive prices and serve a larger portion of demand. This is evident in Figure 4.7, where the international operations representing about 23% of total revenues, contribute only 6% of operating income compared to the 50% generated by AWS.

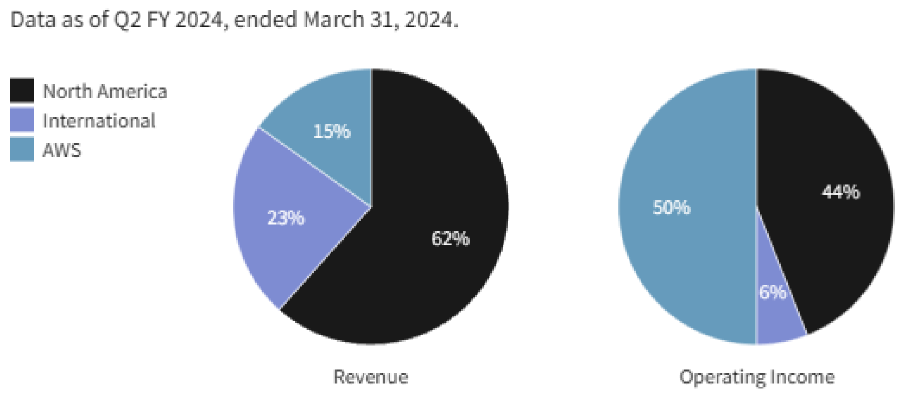


Figure 4.7: Amazon revenue and operating income by business segment [84].

Amazon's strategic vision evolved from being an e-commerce leader to becoming the orchestrator of a global digital platform and ecosystem, exemplifying how dynamic capabilities can be exploited to implement an effective platform leadership strategy. A similar dynamic can be observed in the shift undertaken by Amazon in the early 2000s, when the company transitioned from an online bookstore to an e-commerce platform. The company successfully identified its core competencies as the internal efficiency and the ability in scaling operations rather than retailing products online. In 2012, Amazon launched the AWS Marketplace, allowing external developers and software vendors to sell their applications and cloud solutions as complementary products for the AWS infrastructure, thereby reinforcing its ecosystem. The marketplace exhibits the dynamics of a two-sided platform, where customers are subsidized through free trials and integrated billing, while value is primarily captured from external software vendors.

Through AWS Marketplace, Amazon pursued a platform envelopment strategy, encouraging third-party software vendors to develop applications tailored to AWS infrastructure, which enabled the company to extend its scope from IaaS to the SaaS

[1]Operating income = Revenues – Cost of Goods Sold – Operating Expenses.

layer. To reinforce the ecosystem, Amazon applied strict governance mechanisms to make client investments in AWS non-fungible, such as proprietary APIs that make applications difficult to migrate to other platforms, integrated toolchains, specialized compliance frameworks, and dedicated certification paths. These strategies enabled Amazon to strengthen customer lock-in and consolidate its first-mover advantage.

Competitors such as Google and Microsoft integrated the cloud computing model into their digital platforms, leveraging existing computational resources and infrastructure originally developed to support their core services. When cloud computing first emerged, performance levels were not comparable to on-premises data centers provided by well-established IT firms such as Dell or HP. However, the real value proposition for customers lay in the ability to provision a server on-demand, allowing low-cost experimentation. Incumbents providing on-premises data centers tried to adapt to the emerging business model but ultimately discontinued their cloud activities. These evolutions exhibit the typical characteristics defining a disruptive innovation, as described by Christensen in [35]. The competitive advantage of current providers can be attributed to the structural characteristics of their business models, including [79]:

- Integration of PaaS and IaaS services, internally or via third-party vendors;
- Provision of managed IaaS solutions for clients seeking operational support;
- Hybrid and private cloud offerings addressing regulatory and performance requirements;
- Consulting activities to facilitate the transition from on-premises to cloud environments;
- High availability and broad geographical coverage;
- Presence of a large existing customer base;
- Effective ecosystem governance and partner relationship management;
- Economies of scale resulting from globally distributed operations.

4.3 Cloud Ecosystems

Cloud computing constitutes a complex, network-like ecosystem in which the provider acts as the focal firm. Within this environment, knowledge and innovations co-evolve, and every organization, either providing or adopting cloud services, unwittingly participates. Understanding the structure of the ecosystem and the firm's position within this environment presents significant managerial implications.

The Passau Cloud Computing Ecosystem Model (PaCE model) provides a framework to illustrate the dynamics of cloud ecosystems and, particularly, the process of value co-creation between participants [85]. The framework identifies 31 different roles, grouped into five categories:

- **Client:** person or organization that purchases or consumes cloud services.
- **Vendor:** providers of data, hardware, infrastructure, software, or connectivity.
- **Hybrid roles:** actors combining provisioning and orchestration (e.g., marketplaces, integrators, application providers).
- **Support:** entities enhancing trust or encouraging adoption (e.g., certification authorities, consultancy companies, identity providers).
- **Environment:** external institutions (e.g., universities, governments).

An empirical evaluation of 758 firms revealed that 403 of them simultaneously covered more than one of the 31 identified roles within the ecosystem. For instance, AWS, Microsoft, and Google are simultaneously infrastructure providers and marketplace operators. The 403 companies exhibiting overlapping roles were grouped into five clusters through a cluster analysis, resulting in 63 outliers and the distribution shown in Figure 4.8.

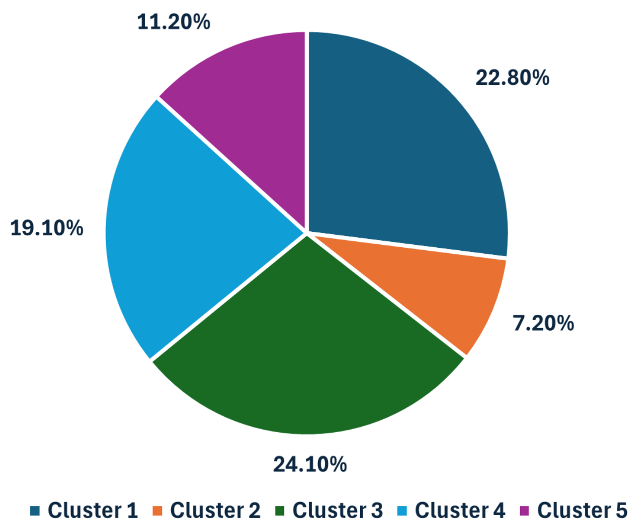


Figure 4.8: Distribution of firm clusters within the cloud computing ecosystem (adapted from [85]).

The clusters can be summarized as follows:

- **Cluster 1:** Service providers for new adopters (e.g., consultancy companies, managed service providers).

- **Cluster 2:** Major players and hyperscalers (e.g., Google, Amazon, Alibaba, Microsoft).

- **Cluster 3:** Telecommunications companies offering complementary cloud services (e.g., Vodafone, Verizon, Telecom).

- **Cluster 4:** SaaS-specialized developers and providers (e.g., Dropbox, Slack, Zoom).

- **Cluster 5:** Organizations providing training, certifications, or standardization (e.g., The Cloud Credential Council, Cloud Security Alliance).

For the purpose of this thesis and to improve illustrative clarity, the final PaCE model is adapted according to the structure proposed in [62] for ecosystem-based value systems. Figure 4.9 presents the resulting framework, which offers a holistic representation of the cloud computing value system. Double arrows indicate complementarity relationships among actors, while single arrows represent flows across the value chain.

Starting from the upstream of the value chain, physical infrastructure suppliers provide the necessary assets required to build and operate cloud providers' data centers. At the centre of the ecosystem, the cloud provider and its complementors interact through a typical *hub-and-spoke* structure. The effectiveness of this model depends on how effectively the provider orchestrates relationships with partners and complementors to enable value co-creation. Value co-creation can arise from the following mechanisms [86]:

- **Exchange:** the provider and its partners trade unique resources or knowledge. This mutual learning process becomes a source of competitive advantage and leads to a *combinatorial innovation*.

- **Addition:** a smaller partner adds a complementary service or module on top of the provider's offering. This mechanism can generate short-term benefits but lacks mutual learning effects.

- **Synergic Integration:** the provider shares resources and processes to co-create value with a partner. This combination establishes a strategic alliance relying on shared knowledge and capabilities.

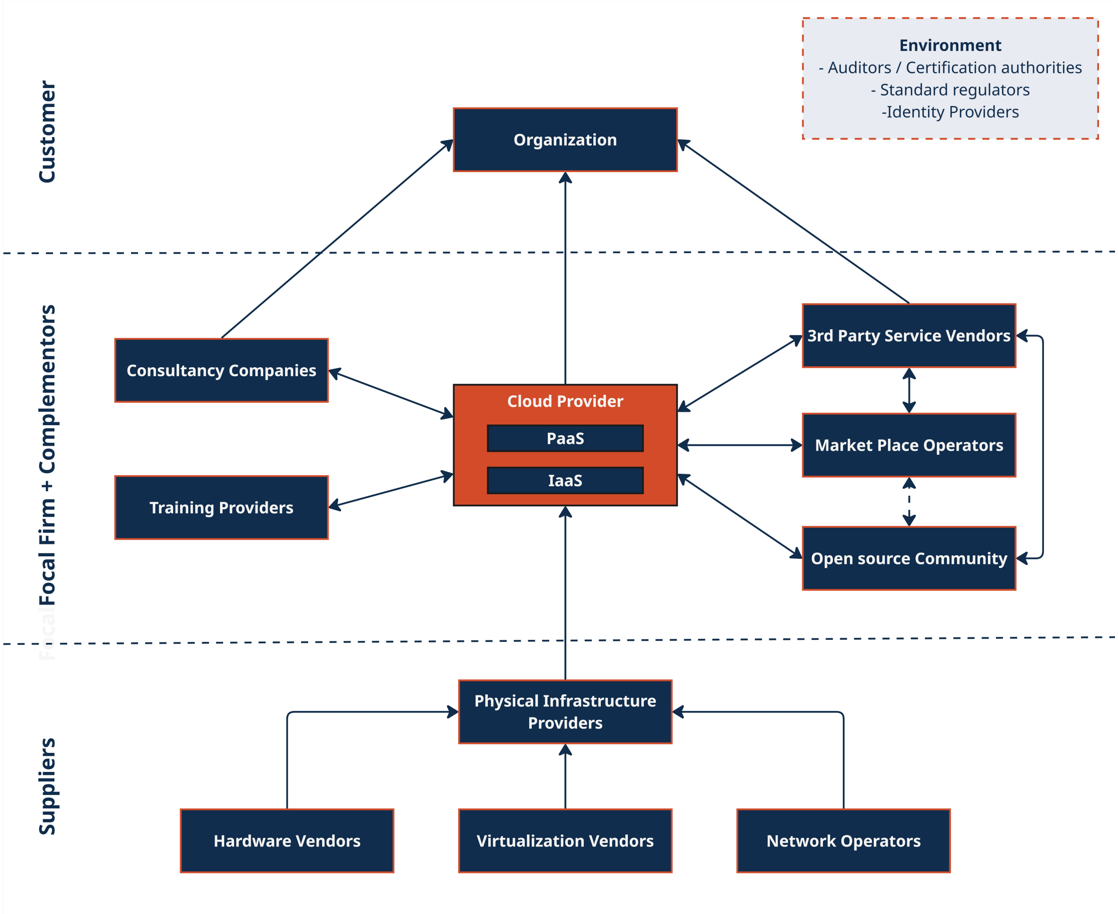


Figure 4.9: Cloud computing value system (adapted from [62] and [85]).

From a holistic ecosystem perspective, outcomes are determined by the provider's absorptive capacity, comprising routines and processes used to transform external knowledge into dynamic capabilities and alliance governance mechanisms. To mitigate opportunism within these relationships, the ecosystem relies on environmental actors such as regulators, auditors, and authorities, which ensure the definition of fair, transparent, and mutually beneficial contracts, as well as *de jure* standardization. The adoption of cloud services is further supported by training and certification providers, which help client organizations accelerate their learning curve and enhance operational efficiency. For the provider, these entities ensure the growth of the customer base and can reinforce vendor lock-in through certifications tied to the provider's specific environment. Within this framework, the adopting firm represents the downstream endpoint of the value chain. Chapter 6 will explore how this perspective shifts in multi-cloud ecosystems, where the user firm can mitigate vendor lock-in and assume a strategic role within the value network.

4.3.1 The Role of Open Source

Within this environment, the orchestration of external knowledge and resources is a determinant factor for effective innovation management, where open source software plays a pivotal role. The open source community develops software that can either be directly adopted free of charge by other users or commercialized by third-party vendors as managed services, which include configuration, maintenance, and technical support for enterprises and can also be distributed through the provider marketplace. Open source software follows the logic of a two-sided platform (see Subsection 3.5.2), as the software is distributed free of charge to users and developers who, in turn, may enhance its functionalities or contribute to its adoption and diffusion, while most of the value is extracted through managed hosting and paid enterprise support. This approach enable faster innovation cycles, as access to the source code allows developers to extend existing applications or reuse software components to accelerate the development of other projects.

Thanks to software's inherent modularity and the participation of external actors such as university researchers, customers, and developers, the open source ecosystem constitutes a fertile environment for open innovation (see Subsection 3.2.2) and enables the establishment of open standards favouring interoperability and reducing vendor dependency. Differently from traditional proprietary software, which requires low marginal costs to produce additional units but high initial investments, open source enables supporting organizations to leverage community-driven effort, allowing them to reach a minimum efficient scale more easily.

Within modern society, open source has contributed in advancing state of the art in various fields and has become a relevant part of the software industry. In this context, open source foundations emerged to provide effective governance models for individual developers, either by supporting communities and a portfolio of related projects, as as in the case of the Linux Foundation [87], or to regulate software licenses [88]. The use and redistribution of open source software are regulated through software licenses approved by the Open Source Initiative (OSI) [89]. Under permissive licenses, such as Apache, BSD, or MIT, source code can be freely reused in proprietary projects. Consequently, many open-source technologies, can either be adopted internally by cloud providers or offered as managed services.

Permissive licenses favour adoption and standardization, as in the case of Docker Engine and K8s (licensed under Apache 2.0), which have become the *de facto* standards for cloud-native development. However, these licenses can also cause friction between the organization maintaining the core software and major cloud providers offering managed solutions on top of it. Large providers can capture a disproportionate value share by leveraging their large-scale global infrastructure and computing capacity.

Figure 4.10 illustrates the user permissions associated with the main types of software licenses.

	Free software	Open-source software	Freeware	Public-domain software
Definition	"FREE" is a matter of liberty, not price	"OPEN" doesn't just mean access to the source code	"FREE" refers to price, while freedom of the use is restricted by creator	"PUBLIC DOMAIN" belongs to the public as a whole
Ground philosophy	Social movement	Development methodology	Marketing goals	Copyright disclamation
Ground rules	Four Freedoms https://www.gnu.org/philosophy/free-sw.html	Open Software initiative https://opensource.org/osd		Creative Common Organization https://creativecommons.org
Free of charge	Not necessary	Not necessary	✓YES	✓YES
Covered by copyright law	✓YES	✓YES	✓YES	✗ NO
Examples	Linux, ubuntu	MySQL, Apache	Skype, Adobe Acrobat	SQLite

Figure 4.10: User permissions under the main types of open-source licenses [90].

A notable example of this phenomenon is Redis [91], an open-source key-value store[2] widely used by technology companies such as Snapchat, Instagram, and GitHub, and offered by major cloud providers as managed services (e.g., AWS ElastiCache, Google Cloud Memorystore, and Azure Cache for Redis). In response, Redis Ltd. (formerly Redis Labs), which manages the Redis trademark, switched the software license from BSD to a tri-license model to incentivize disclosure of modifications and promote the purchase of commercial redistribution licenses.

4.4 Cloud Native

Cloud ecosystem is an evolving and dyamic evironment, characterized by multiple parallel innovation processes. Within this context, cloud native has emerged as the dominant design in software development, enabling a paradigm shift from traditional monolithic architectures to modular and service-oriented systems. The migration from legacy monolithic applications to microservices architectures represents a key driver of digital transformation [92].

[2] A type of database that stores data as a collection of unique keys and associated values.

A technological paradigm can be defined as an "outlook" on a system that involves a set of procedures, problems, and the knowledge required for their resolution [93]. For such a paradigm to emerge, it requires a set of theories, knowledge, tools, and methods that allow the transformation of a given technology into actual products and services [19]. The framework proposed in Figure 4.11 illustrates the cloud native paradigm from both the supply side (cloud providers and application developers) and demand side (firms adopting cloud-native services).

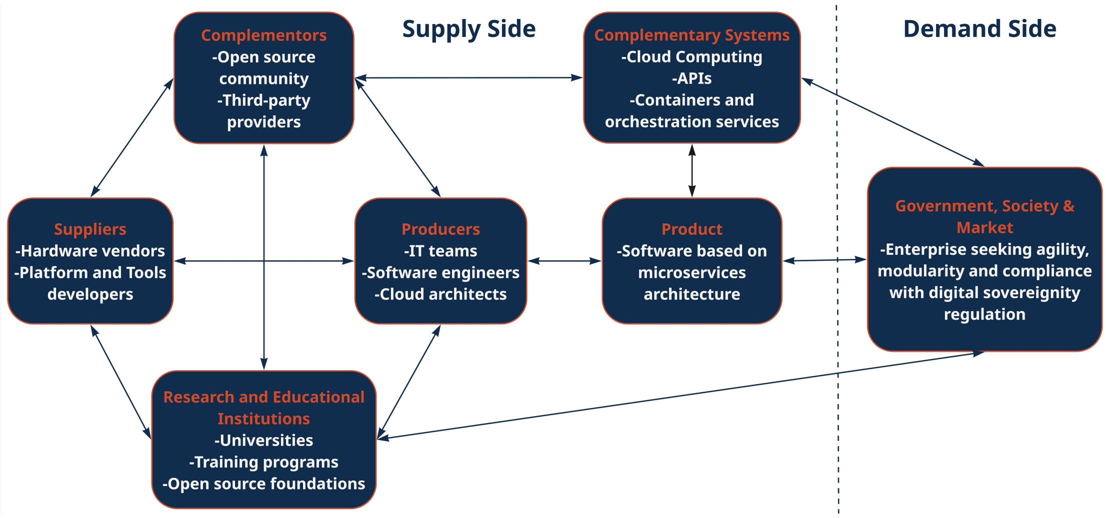


Figure 4.11: Cloud native paradigm (adapted from [19]).

This new approach fully exploits the inherent characteristics of cloud environments, such as scalability and flexibility, thus allowing a modular structure in which a service can be modified, replaced, or removed independently of the rest of the application. Modularity enables easier integration and adaptation of system modules, resulting in more efficient delivery with shorter iterations and reduced TTM [94]. Moreover, the high degree of modularity achieved by these systems allows organizations to adopt Agile frameworks and DevOps practices more effectively and favor the establishment of cloud-native architectures as the main model for cloud application development.

Agile practices aim to maximize customer value and product quality, prioritizing adaptability and short iteration cycles over long-term planning. This method typically relies on small, cross-functional teams, where customer feedback is continuously incorporated into iterative improvement cycles [95]. Through Agile practices, organizations are able to foster learning "by doing" and "by using" almost concurrently. DevOps, in turn, extends Agile principles to the delivery phase, involving all software stakeholders through a collaborative approach where development and operations teams work together to maximize efficiency and productivity. DevOps workflows are based on CI/CD pipelines, which automate and accelerate the integration, testing, and deployment of application code.

4.4.1 Containers

The key technology enabling this paradigm shift is containerization. In microservices architectures, containers represent the fundamental modular unit, providing flexibility and scalability across distributed systems. In traditional monolithic architectures, deploying an application on a server or device typically required hosting each application instance within a dedicated VM, each running a full guest OS coordinated by a hypervisor on the host OS. By contrast, containers provide lightweight and isolated instances of an application, packaging application code together with libraries and dependencies needed to run properly [96]. The lifecycle of containers is managed by a container runtime, such as Docker Engine, which has become the *de facto* industry standard.

In terms of portability, containerization shifts dependencies from the OS to the application layer. For instance, moving an application from a test to a production environment using VMs can lead to incompatibilities if the guest OS version or distribution differs. Containers, instead, ensure consistent performance as long as the container runtime is compatible with the host OS. Figure 4.12 illustrates the differences between VMs and container-based environments.

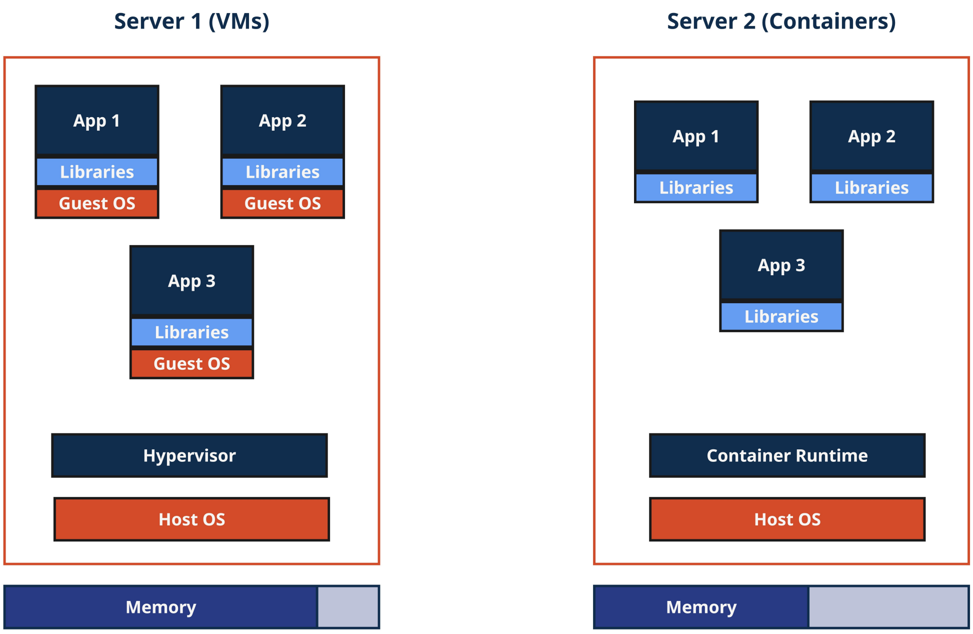


Figure 4.12: Comparison between VMs and container-based environments.

In modern applications based on microservices architectures and CI/CD pipelines, containers are essential for achieving faster service startup, efficient resource utilization, and high portability [97]. They enable applications to scale far more efficiently than VMs, with deployment and stop times faster by orders of magnitude, as shown in Table 4.1.

Table 4.1: Docker vs Virtual Machine Performance Comparison [97]

Type	**Time to Start**	**Time to Stop**
Docker	44 ms	28 ms
VM	59 s	33 s

The isolation between containers is achieved through namespaces, which create virtual environments within the host OS. While containers are lighter and faster than VMs, they do not provide the same level of privacy and security offered by a dedicated OS. Therefore, virtual machines are still necessary in scenarios requiring OS-level isolation, such as executing user-generated or third-party code within cloud provider infrastructure.

4.4.2 Kubernetes

Containers represent a powerful technology but managing their lifecycle, connectivity, isolation, and communication can become increasingly complex, particularly in large or distributed systems involving hundreds or thousands of instances. This complexity has driven the need for an orchestration tool capable of deploying, modifying, or scaling applications without manual intervention on container instances.

In 2015, two years after the release of Docker, Google introduced Kubernetes (K8s), an orchestration tool derived from its proprietary internal system, Borg. K8s is a system for automating the deployment, scaling, and lifecycle management of containerized applications. Its key innovation lies in the declarative model: users define the desired state of the system, and K8s automatically distributes container instances across the underlying infrastructure to achieve it.

As an analogy, K8s can be described as a post office for containers: the user specifies the destination (desired state), while Kubernetes autonomously determines the optimal way to deliver and manage containers to reach that outcome.

Shortly after its release, Google donated K8s to the Cloud Native Computing Foundation (CNCF), part of the Linux Foundation, which currently oversees its maintenance [98]. This strategic move promoted open standards while positioning Google as a technological leader in the emerging cloud native ecosystem, leveraging internal expertise to gain first-mover advantages. By releasing the project as open source, Google was able to draw upon external knowledge and encourage widespread adoption.

Kubernetes has since evolved into a flourishing ecosystem built on a highly modular architecture, continuously expanded by the open source community. Contributors enhance the system by adding functionality directly to the source code or by developing external tools such as Istio and Cilium, which facilitate service communication. These components are often essential to fully exploit K8s capabilities in complex and distributed environments.

Today, Kubernetes is among the most active and influential open-source projects, with over 100,000 individual contributors and more than 20,000 participating organizations on GitHub [99]. According to the CNCF 2024 survey [100], it is used in production by approximately 80% of companies worldwide, as shown in Figure 4.13.

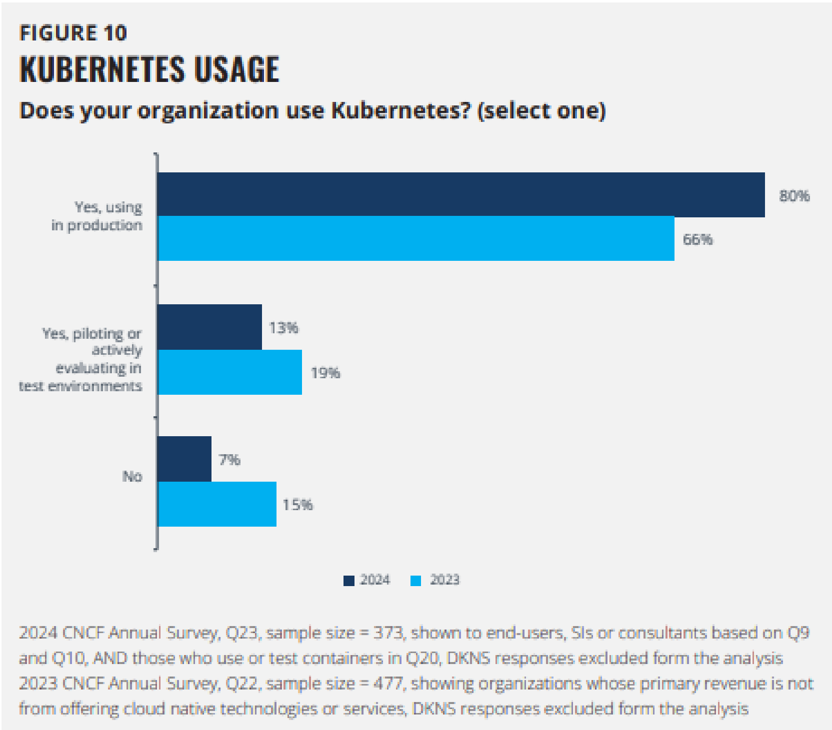


Figure 4.13: Global adoption of K8s [100].

Kubernetes Architecture

K8s enables the orchestration of containers across large-scale environments and clusters of machines. Each cluster functions as an independent unit grouping all K8s resources and modules. A K8s cluster is managed by a control plane, which includes the following core components, as illustrated in Figure 4.14:

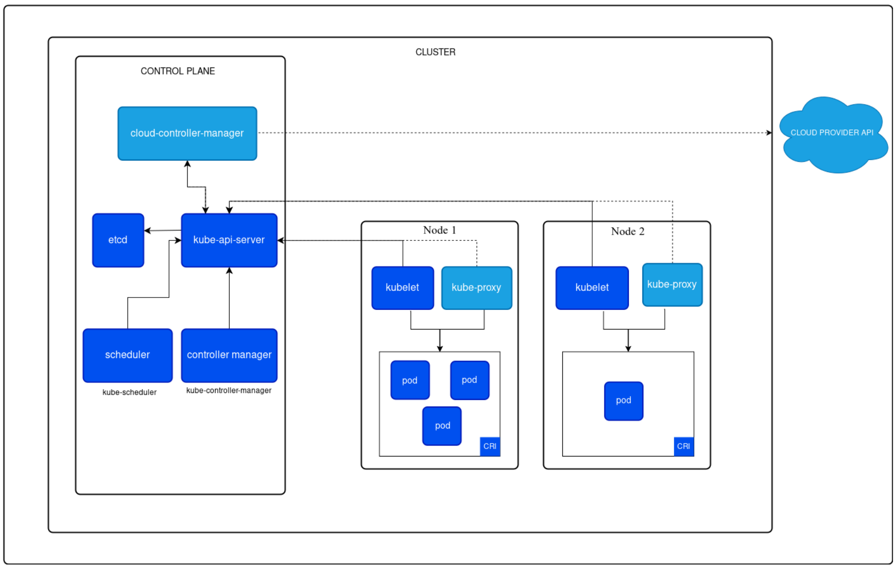


Figure 4.14: Kubernetes cluster architecture [7].

- **API Server**: Exposes the cluster APIs, acting as the central hub for communication between the control plane and worker nodes.
- **etcd**: A distributed key-value store maintaining the system's configuration data and the declared state of the cluster.
- **Controller Manager**: Runs control loops to continuously monitor the system and ensure it matches the desired state.
- **Scheduler**: Determines on which node deployments should occur based on predefined metrics such as resource capacity and availability.

A node can be either a physical or virtual server that hosts one or more pods, which are the smallest deployable units in K8s. A pod typically represents a single instance of an application or service and is managed internally by the *kubelet*, which communicates with the *kube-api-server* to ensure the node matches the declared cluster state. Because pods are ephemeral, and they can be destroyed, replaced, replicated, or recreated dynamically to reach the desired system state, Kubernetes uses *services*, which are logical abstractions providing stable endpoints for groups of pods to manage communication. Kubernetes supports three main service types:

- **ClusterIP**: Used for internal communication among pods within the cluster.
- **NodePort**: Provides an endpoint to expose the application externally, opening a common port shared across all node IPs.

- **LoadBalancer**: Distributes external traffic across multiple nodes and assigns a public IP address to the service. Load balancing is typically managed by cloud providers, as it is not natively supported in vanilla K8s clusters.

When a K8s cluster is hosted in the cloud, providers automatically supply and manage the underlying resources. In these environments, the *Cloud Controller Manager* enables K8s to interact with the provider's infrastructure through provider-specific APIs. According to a Dynatrace report, the number of K8s clusters hosted in the cloud increased from 45% in 2022 to 64% in 2024.

Despite its advantages, K8s is well-known for its operational complexity. Effective governance of K8s clusters requires careful configuration of security and resource management policies [101], such as Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) for user permissions; Resource Quotas to manage CPU, memory, and storage utilization; abd Network Policies to regulate communication between pods.

To address K8s configuration and operational challenges, all major cloud providers now offer managed services, including Google Kubernetes Engine (GKE), Amazon Elastic Kubernetes Service (EKS), Azure Kubernetes Service (AKS), and Aruba Kubernetes-as-a-Service (KaaS), within the European market. As reported by Flexera [4], European companies were initially less likely to adopt managed service provider solutions, but adoption rates have recently converged with global averages. The utilization of managed services correlates strongly with the scale and complexity of cloud infrastructure, with larger enterprises showing higher adoption rates, as illustrated in Figure 4.15.

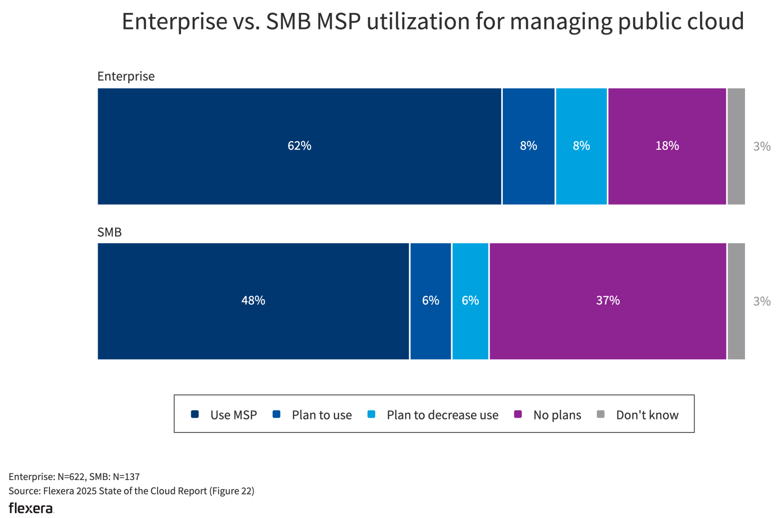


Figure 4.15: Adoption rate of managed K8s services by company size [4].

4.4.3 Business Impact of Cloud Native

Cloud-native technologies and infrastructure tools, such as Docker and K8s, have already had a significant impact on the modernization of different industries. One of the main benefits of cloud-native approaches lies in the shortening of innovation cycles, which requires organizations to rethink tailored business models to effectively support emerging technologies and cloud-native capabilities. Through a flexible and modular structure, cloud native can provide dynamic capabilities and enable organizations to adapt to technological innovations in the context of digital services. For instance, a cloud-native system can serve as the backbone for multiple emerging technologies, including AI/ML, IoT, and edge computing, which can be combined as modules within digital product platforms and service-oriented business models.

Across the technology stack (see Figure 4.16) cloud-native principles promote modularity and innovation at the application code layer, driving the commoditization of lower-level infrastructure. In this context, PaaS and FaaS solutions simplify the adoption of cloud native and DevOps practices by outsourcing to cloud providers the management of the underlying stack (from application runtime to infrastructure). This enables developers to focus on application logic and functionality, leveraging a hosted product platform to accelerate delivery and experimentation.

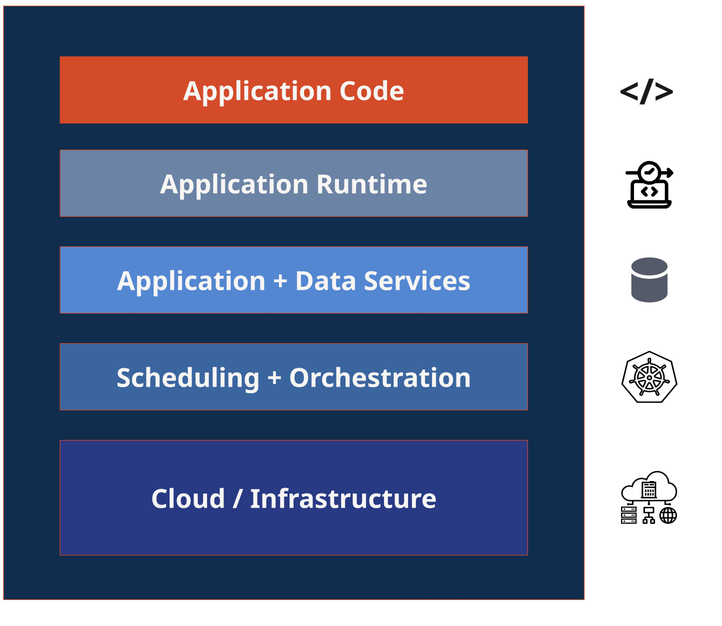


Figure 4.16: Cloud appication stack.

In the innovation process enabled by cloud-native technologies, startups tend to be favoured because of their organizational agility. Additionally, newly formed entities are able to implement the technological paradigm from day zero. As shown by CNCF survey results in Figure 4.17, cloud native adoption is usually more widespread in small-sized companies. Conversely, established enterprises often struggle to migrate toward cloud-native models because of legacy systems based on monolithic applications. As a result, cloud native can favour low-end disruption and lower barriers to entry in digital industries such as e-commerce, online streaming, and mobile applications, where flexibility and scalability are critical success factors.

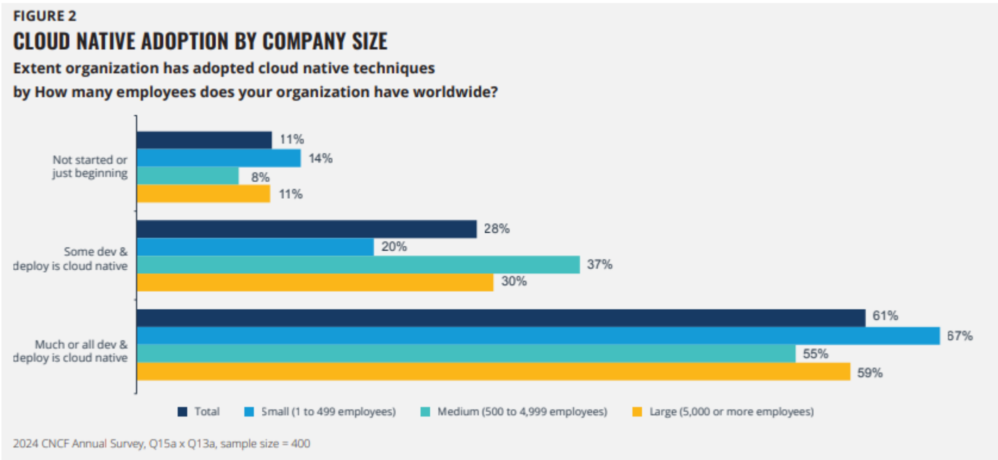


Figure 4.17: Cloud native adoption by company size [100].

Cloud native adoption enables an open and collaborative ecosystem, characterized by interactions between startups, incumbents, providers, and open source communities through reusable microservices, shared infrastructure, and distributed expertise. To accelerate effective cloud adoption, many large organizations have established a Cloud Center of Excellence (CCoE) (74% according to [4]). A CCoE is a specialized, multidisciplinary team responsible for centralizing governance and defining best practices to implement cloud-native solutions. In parallel, organizations are introducing Internal Developer Platforms (IDPs), which provide developers with a tool to unify the management of infrastructure, configurations, and deployments across multiple cloud environments, preventing the fragmentation of configurations and resources. These solutions foster organizational modularity, a capability particularly relevant for the governance of multi-cloud ecosystems which will be examined in detail in Chapters 5 and 6 .

4.5 Beyond the Cloud: Innovation in Cloud Ecosystems

As previously discussed, cloud-native approaches enable organizations to implement emerging technologies within their digital platforms without radically altering the system's structure. As illustrated in Figure 4.18, the technologies expected to have the greatest business impact over the next three years are AI and Machine Learning (ML) and edge computing.

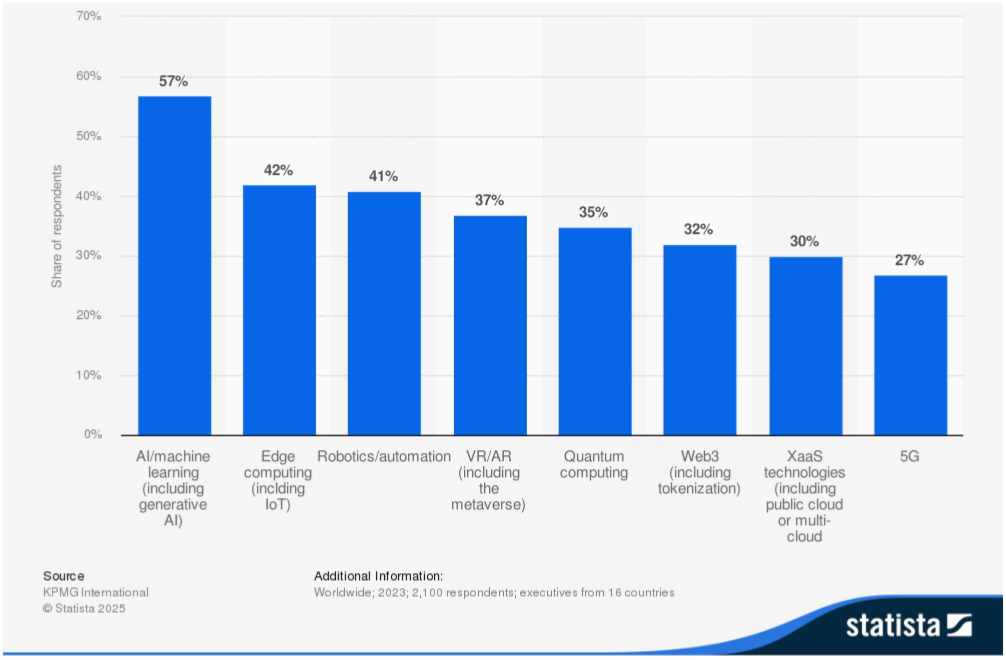


Figure 4.18: Emerging technologies expected to have the greatest business impact [102].

4.5.1 Edge Computing

Edge computing represents a shift from the traditional centralized cloud computing model, in which data are processed and stored in data centers operated by providers and distributed across fixed geographical regions. This centralized approach is not optimal for applications requiring ultra-low latency, such as real-time IoT systems, where data must be processed closer to its source (e.g., IoT devices and sensors).

To address this limitation, edge computing extends centralized cloud infrastructure to distributed edge nodes, allowing applications to perform data processing, storage, and caching closer to the source. Major cloud providers first implemented cloud Content Delivery Networks to reduce latency and server load, enabling caching and content delivery closer to the user. Over time, these nodes evolved into edge nodes, which include computational capabilities for localized data processing. The integration of edge nodes within computing ecosystems to create a distributed cloud architecture is commonly referred to as Edge-as-a-service. This model has been adopted by major cloud providers through offerings such as AWS Lambda@Edge, Azure Edge Zones, and Google Distributed Cloud Edge.

Similarly to how hardware virtualization enabled the emergence of cloud computing, edge computing is powered by Network Functions Virtualization (NFV) and Software Defined Networking (SDN). These technologies allow the virtualization of networking hardware (e.g., routers, switches, firewalls) and resources within software environments such as VMs or containers. For cloud and network providers, this enables the delivery of virtualized networking services at the edge through Virtual Network Functions (VNF), without relying on proprietary hardware.

The architectural model that extends cloud computing capabilities from provider data centers to edge nodes and IoT devices is known as *fog computing*, a concept first introduced by Cisco in 2012. Fog computing provides an end-to-end horizontal architecture that distributes computing, storage, control, and networking functions closer to users.

The benefits of fog computing can be summarized by the acronym **SCALE**:

- **Security**: Improves privacy by processing data closer to the user, reducing the risk of interception.
- **Cognition**: Enables context-aware provisioning of resources.
- **Agility**: Supports faster innovation at the edge without being tied to centralized infrastructure.
- **Latency**: Reduces delays by processing data near the source.
- **Efficiency**: Optimizes the utilization of distributed computing resources.

The seamless integration of functions and resources across distributed infrastructures is referred to as the *cloud-to-thing continuum* or *cloud-edge continuum*. In this model, edge computing and cloud converge, combining high-performance computing enabled by edge nodes with the capacity and scalability of cloud resources. This approach enables resource optimization and supports advanced capabilities such as AI services, while improving compliance and aligning with the principles of the European Digital Strategy (see Section 6.2).

4.5.2 AI: Sustaining or Disruptive Innovation?

First, it is important to clarify that the term Artificial Intelligence AI does not refer to a single technology or paradigm, but rather to an umbrella term which refers to systems capable of performing tasks with human-like capabilities. ML represents the most promising subset of AI, relying on algorithms and statistical models that enable systems to learn from patterns of data, adjust internal parameters, and apply their acquired "knowledge" on new inputs [103].

Within ML, deep learning techniques received substantial attention in recent years, particularly those based on Deep Neural Networks (DNNs). A major breakthrough came from Geoffrey Hinton, who in 1986 introduced the backpropagation algorithm, allowing multi-layered (or deep) neural networks to model complex pattern representations automatically. After a long period of skepticism, the deep learning approach emerged as the dominant design for AI systems following the publication of *Attention Is All You Need* [104] which introduced the building blocks for transformer-based architectures that serve as the foundation of modern LLMs. Since the release of ChatGPT-4 by OpenAI in March 2023, the adoption of LLMs has experienced one of the fastest technological diffusions in history. ChatGPT models are currently used by approximately 10% of the world's adult population, as reported by the company [105]. Considering similar services such as Claude by Anthropic, Gemini by Google, and Copilot by Microsoft, the overall adoption rate is likely higher.

While LLMs accelerated the diffusion of AI-based technologies to the early majority, AI and ML were already used by governments and enterprises (the early adopters) to optimize processes and operations in domain-specific use cases. Examples include fraud detection in finance, predictive maintenance in manufacturing, and clustering or principal component analysis in marketing.

Most emerging technologies follow a non-linear trajectory with respect to users' expectations. The *Gartner Hype Cycle* provides a visual framework to represent the maturity and expectations surrounding new technologies. Before reaching the plateau of productivity and being able to solve real business problems, a technology typically experiences a *peak of inflated expectations*, where hype and experimentation exceed actual technological capabilities. The 2025 Gartner Hype Cycle for AI-related technologies places Generative AI (LLMs), cloud AI services, and edge AI among those expected to reach the plateau of productivity within five years, as illustrated in Figure 4.19.

Artificial Intelligence exhibits the characteristics of a radical innovation. However, as discussed in Subsection 3.2.2, the disruptive potential of a technology depends on market dynamics and its impact on incumbents' positions. Since the release of ChatGPT-4 and the subsequent mass adoption of generative AI, the IT industry has undergone major structural changes, such as the rise of AI-based SaaS startups developing proprietary models or relying on existing LLMs to deliver their services. However, in the context of cloud infrastructure, competitive dynamics among major cloud providers have remained relatively stable.

Hyperscalers such as Google, Microsoft, and Amazon not only provide state-of-the-art AI solutions to enterprises, but also supply the underlying infrastructure used by other AI companies for model training and inference. For instance, OpenAI relied exclusively on Microsoft Azure for its computing capacity until early 2025. The company later adopted a multi-cloud strategy to reduce dependence on a single provider. First, it announced the *Stargate Project*, developed in collaboration with Oracle and other partners [106], and subsequently signed an agreement with Google Cloud Platform (GCP) for additional computational capacity [107].

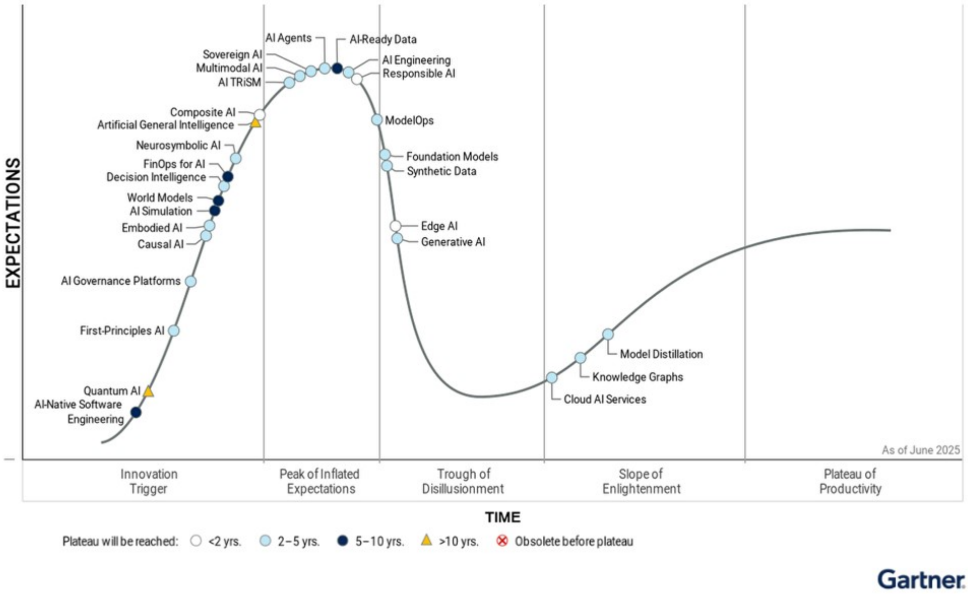


Figure 4.19: Gartner Hype Cycle for AI-related technologies [108].

4.5.3 AI-Cloud-Edge

In a recent presentation at Y Combinator, Andrej Karpathy, former Director of AI at Tesla, proposed an analogy between LLMs and operating systems [109]. He noted that LLM inference typically occurs in cloud environments, similar to how computation in the 1950s–1970s was centralized on mainframes before the era of personal computers. Following this analogy, the inference environment—whether in the cloud or on a local device—acts like an operating system kernel, serving as a bridge between the user interface (prompt) and the underlying hardware, where model parameters and computational resources reside.

In this architecture, computation is handled within a centralized backbone, while the LLM orchestrates external components such as APIs, Python code, and auxiliary models, similarly to how operating systems manage hardware and software subsystems. The interoperability between LLMs and external tools has been significantly enhanced through the development of the Model Context Protocol (MCP), an open standard that enables the connection between AI applications and external systems [110].

These interactions may represent a paradigm shift in software development, where the human-AI collaboration loop becomes a key principle in designing and developing new applications. Combining human contextual understanding with AI's analytical and generative capabilities will be crucial to achieving higher efficiency and solving real business challenges. Furthermore, the rise of AI-assisted coding enhances end-user programming through natural language interfaces, enabling organizations to build software internally using tailored platforms optimized for the human-AI collaboration loop.

Within this emerging paradigm, it is essential to create an AI-ready digital infrastructure supporting these capabilities, offering scalable computational resources and data management systems while ensuring interoperability and sustainability through efficient resource utilization.

In this context, a mutually beneficial relationship exists between cloud computing and AI: cloud and edge computing paradigms provide the computational backbone for AI-based services, while AI enhances the performance and efficiency of cloud environments. AI workloads, especially during training, demand massive computational capacity that is often infeasible on-premises. To meet this demand, major cloud providers offer tailored AI services that provide on-demand access to Graphics Processing Units (GPUs) and scalable compute resources [111].

Moreover, edge and fog computing enable distributed computation near data sources, such as IoT devices, allowing real-time inference and more efficient model training. Centralized cloud infrastructure, in turn, stores and processes the data collected by IoT devices, providing the input needed to train or refine AI models. On the other hand, the application of ML and AI within cloud environments yields several operational advantages [112]:

- **Dynamic resource scheduling:** optimizes allocation based on workload demand;
- **Network and compute optimization:** improves load balancing and reduces energy consumption;
- **Predictive monitoring:** enables early detection of anomalies or security threats.

Chapter 5

Multi-Cloud Adoption

Vendor lock-in represents a major concern for many enterprises, and the development of distributed multi-cloud and edge deployment models has become crucial to mitigate dependence on hyperscalers. As cloud ecosystems expand and organizations seek sustainable ways to experiment and innovate with emerging technologies, the demand for flexibility in cloud environments grows. While the cloud native paradigm provides a platform to build solutions optimized for cloud computing environments, multi-cloud enables organizations to establish a strategically relevant position within the broader cloud ecosystem. Through this approach, organizations can combine services from different providers according to their operational requirements and customers' needs, pursuing strategic and technological differentiation. This chapter explores the strategic drivers and operational challenges associated with multi-cloud adoption, as well as the technologies and architectures enabling seamless interoperability within multi-cloud environments. Additionally, it also presents the results of a study targeting IT professionals to investigate organizations' cloud strategies and adoption patterns. For the purpose of this thesis, multi-cloud refers to the utilization of infrastructure or platform services from two or more public cloud environments, regardless of the approach used to manage private infrastructure. According to recent studies, approximately 30% of organizations were already operating in hybrid multi-cloud or public multi-cloud environments. Based on the results, the percentage of organizations adopting multi-cloud deployments is expected to double over the next three years, reaching approximately 60%, as shown in Figure 5.1.

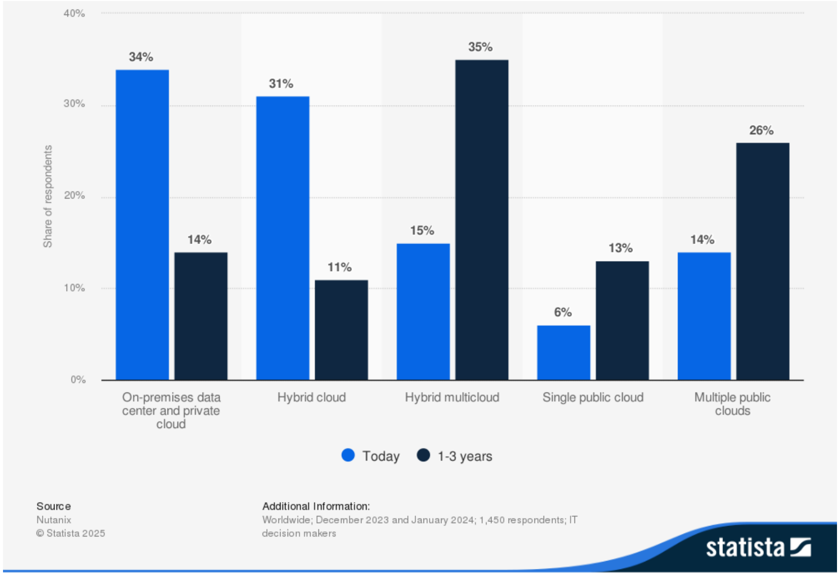


Figure 5.1: Usage of IT models worldwide in 2024, by type [113].

5.1 Strategic Drivers

A study conducted by Google identified the key drivers for the adoption of multiple public cloud provider services (see Figure 5.2).

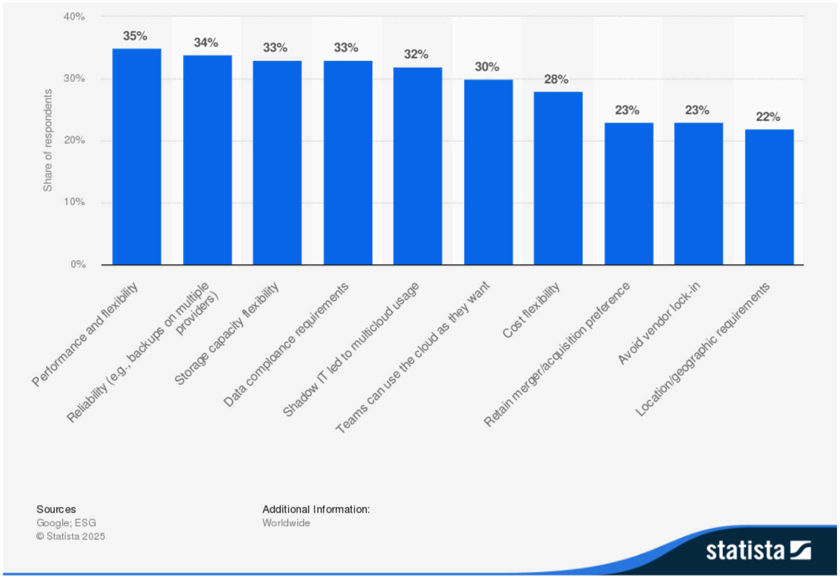


Figure 5.2: Reasons for using a multi-cloud infrastructure in 2023 [114].

From a strategic perspective, the most important drivers can be grouped into four main categories: flexibility, reliability, performance, and compliance.

5.1.1 Flexibility

The primary advantage of multi-cloud adoption lies in mitigating the risk of vendor lock-in, a condition in which the customer cannot easily move or integrate services of a different vendor without substantial costs, legal constraints, or technical incompatibilities [115]. This condition arises from the lack of standardization within different cloud ecosystems, where each major cloud provider deliberately restricts interoperability and workload portability as part of their ecosystem governance strategies.

However, organizations can overcome these challenges through the effective implementation of multi-cloud architectures. From a strategic perspective, avoiding vendor lock-in improves firms' bargaining power with suppliers, which is one of the key factors shaping the competitive landscape within industries, as highlighted in [20]. Moreover, multi-cloud flexibility allows organizations to combine best-of-breed services from different providers to develop tailored solutions aligning with their specific environments and business objectives. This approach fosters more effective exploration strategies and accelerates innovation in both products and processes. Although major hyperscalers offer comparable categories of core services, such as compute, storage, and networking, the maturity of their capabilities differs significantly across providers. For instance, AWS is widely recognized for the robustness of its S3 storage services, GCP differentiates itself through advanced AI/ML offerings, while Microsoft Azure stands out for its integration with the enterprise and productivity ecosystem.

For the customer firm, a seamless multi-cloud environment enables the organization to leverage an optimal combination of services, fostering dynamic capabilities across the enterprise. In such a configuration, the firm can operate across multiple provider environments as interoperable infrastructure modules while exploiting provider-specific technological specializations.

Finally, multi-cloud adoption can stimulate a fairer competitive landscape, increasing customer bargaining power and reducing hyperscalers' price rigidity. As technological innovation continues to accelerate interoperability and standardization, it is expected to drive the evolution toward an open and federated cloud ecosystem, which will be the focus of Chapter 6.

5.1.2 Reliability

One key driver for the transition toward multi-cloud environments is the need to ensure reliability and reduce the risks associated with dependence on a single cloud provider. Organizations relying on a single provider cannot guarantee business continuity in case of service outages or large-scale infrastructure failures, which may lead to significant financial and operational losses. A reliable multi-cloud architecture must be built around two fundamental principles: High Availability (HA) and Disaster Recovery (DR) [116].

HA refers to the system's capability to continuously operate without interruption, switching workloads to another infrastructure module when a failure is detected. In multi-cloud architectures, HA is achieved through:

- **Redundancy**, by retaining multiple instances of critical assets across different providers;
- **Geographical distribution**, which reduces the risk of localized infrastructure failures;
- **Load balancing**, which dynamically manages and distributes traffic to prevent overload on a single instance.

DR, on the other hand, refers to the strategies adopted to recover system functionality and data integrity after a disruptive event, such as a cyberattack, natural disaster, or security breach. DR is achieved through backups and data replication distributed across multiple providers' infrastructures and synchronized to ensure up-to-date information. Additionally, the system must be constantly monitored to support rapid fault detection and minimize recovery time. Other technologies that can be implemented to minimize Recovery Time Objectives (RTO)[1] include IaC, which reduces the time needed to reconfigure infrastructure services, and AI/ML predictive analytics, which anticipate potential failovers.

Implementing cloud-native principles during system design enables organizations to meet reliability requirements. In particular, K8s facilitates workload lifecycle management through declarative deployments and rolling updates. Rolling updates enable zero-downtime deployments in K8s, ensuring that the declared number of active pods is maintained through automatic updates. The following example illustrates a YAML deployment file used to deploy ten pod replicas with a rolling update strategy. The system guarantees that at least five pods (50%) remain active during the update, eventually creating additional replicas to accelerate the replacement.

[1]The necessary amount of time needed to recover the normal operations after a service interruption.

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: nginx-test
spec:
  replicas: 10
  selector:
    matchLabels:
      service: http-server
  minReadySeconds: 5
  progressDeadlineSeconds: 600
  strategy:
    type: RollingUpdate
    rollingUpdate:
      maxUnavailable: 50%
      maxSurge: 2
```

5.1.3 Performance

The performance driver concerns key network metrics such as the Round Trip Time (RTT), which measures the time required for a data packet to travel from the customer endpoint to the provider infrastructure and back. RTT provides an estimate of network latency, a crucial parameter for latency-sensitive applications such as real-time analytics, cloud gaming, and telecommunications services.

Although major cloud providers operate globally through multiple regions, performance can vary significantly across them. Consequently, no single provider consistently guarantees superior performance in all locations. For mission-critical applications requiring ultra-low latency, organizations often adopt a multi-cloud strategy to select the provider offering the best performance in each region.

Additionally, inter-cloud communications (data exchange between different providers) can sometimes achieve better latency results compared to intra-provider latency, as highlighted by Figure 5.3, which presents latency data collected from the Kentik Cloud Latency Map [117].

Another advantage related to performance arises from the integration of edge computing with multi-cloud environments, referred to as the *multi-provider cloud-edge continuum* [118]. This paradigm combines the benefits of both approaches, enabling real-time data processing across a distributed set of resources delivered by multiple CSPs which will be discussed in Subection 6.2.3.

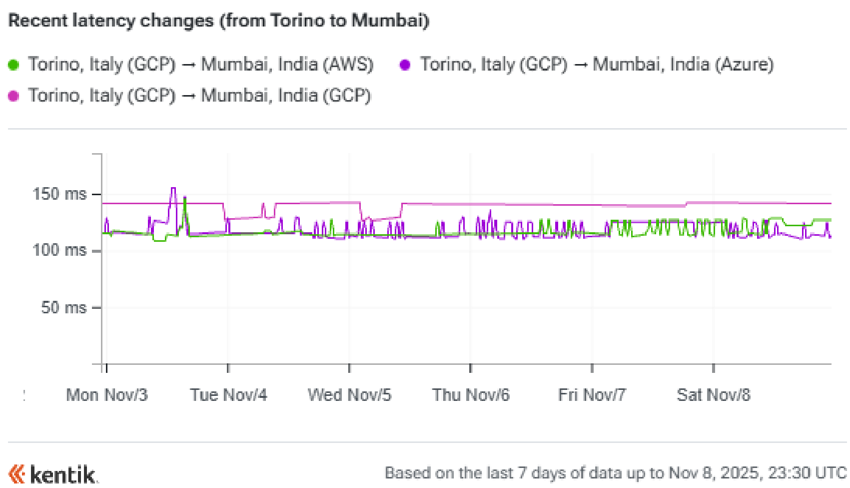


Figure 5.3: Observed inter-cloud and intra-cloud latencies between Turin and Mumbai [117]

5.1.4 Compliance

Another key driver behind multi-cloud adoption is compliance, which in this context refers to ensuring that an organization's IT infrastructure adheres to regulations and data residency requirements. Data residency refers to a business or government decision about the geographical location where data must be stored and processed [119]. For multinational enterprises operating across multiple jurisdictions, ensuring compliance involves managing sensitive data, such as customer Personally Identifiable Information (PII), in accordance with both local and international regulations. In Europe, for instance, the GDPR establishes a set of principles for the processing of personal data, including lawfulness, fairness and transparency, purpose limitation, data minimization, accuracy, storage limitation, and integrity and confidentiality [120].

The regulatory context has led organizations to adopt the principle of *compliance by design*, a preventive approach that integrates compliance considerations from the earliest stages of systems development. For cloud computing architectures, design decisions should ensure transparency on where data are accessed, stored, and processed, while meeting applicable jurisdictional and security requirements.

When operating in multiple countries, a critical issue arises concerning jurisdictional conflicts. For instance, Article 45 of the GDPR states that the transfer of personal data to a third country or an international organization may take place only when the proper level of protection is ensured [120]. Conversely, the U.S. *CLOUD Act* imposes obligations on U.S. tech-companies to provide data to U.S.

authorities in case of a legal infringement, even when such data are stored abroad [121]. This may lead to potential legal conflicts between European data protection and U.S. law obligations.

To mitigate these risks, major cloud providers established EU boundaries for European customers' data. In 2023, Microsoft introduced the *EU Data Boundary*, designed to minimize data outflows from EU data centers [122], followed by similar initiatives from GCP and AWS. Amazon, for instance, launched the *AWS European Sovereign Cloud*, a dedicated European cloud infrastructure logically and physically separated from global AWS regions. Alternatively, joint ventures between hyperscalers and European providers, such as *Bleu* between Microsoft, Orange, and Capgemini [123] or strategic partnership such as *Azure Local* between Microsoft and Aruba [124], seek to reduce the influence of external jurisdictions while leveraging hyperscaler capacity. From the perspective of organizations adopting cloud services, compliance is closely tied to the principle of data sovereignty, which refers to the control an organization maintains over its data across jurisdictions and industry-specific regulations [125]. Within this context, multi-cloud architectures enable a strategic response to meet regulatory requirements across different jurisdictions such as storing personal data with a local CSP while deploying less sensitive workloads on global hyperscalers. In the European landscape, this dynamic has contributed to the emergence of *sovereign clouds*, which are cloud environments specifically designed to comply with national or regional data sovereignty laws, ensuring that data are stored and governed under local legislation. Sovereign clouds are often integrated within multi-cloud strategies to ensure compliance, flexibility, and scalability of computing resources [126].

5.2 Operational Challenges

The adoption of a multi-cloud strategy is not straight-forward as implementing services from a single cloud provider and requires a specialized set of skills and expertise to address the key operational challenges, which include: lack of interoperability between provider infrastructures, fragmented security measures, and cost management across different cloud environments. Overcoming these challenges is essential for organizations to fully exploit the benefits of a unified multi-cloud environment. However, as reported by Flexera, in 2025 managing multi-cloud environments remains one of the major challenges for organizations. As illustrated in Figure 5.4, 68% of large enterprises identified multi-cloud management as a major concern, compared to 59% of SME. This result is explained by the higher diffusion of multi-cloud among large-scale organizations, facing specific drivers such as geographical distribution, integration of cloud environments following an acquisition, or access to exclusive best-of-breed services.

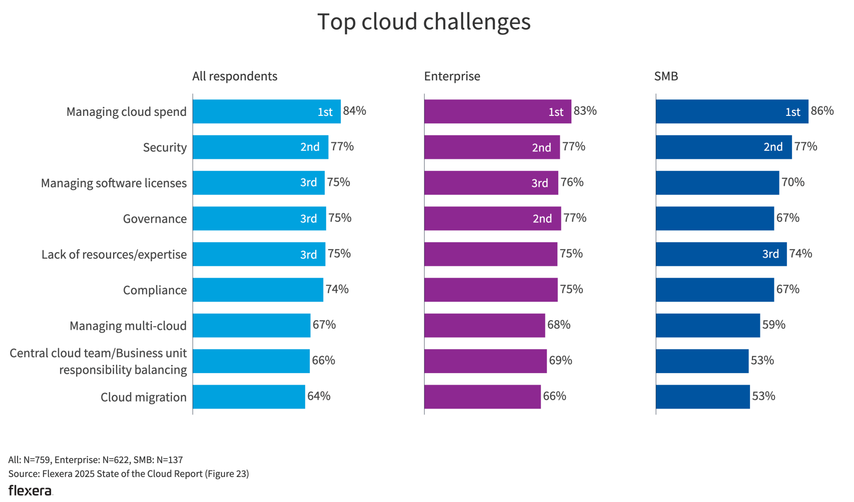


Figure 5.4: Top cloud challenges in 2025, by company size [4].

At the same time, larger enterprises face additional obstacles such as reduced agility, complex governance structures, and integration with legacy systems and applications. Despite the additional drivers, bureaucratic constraints and technical debt slows the transition toward multi-cloud-ready environments for large enterprises compared to digitally native firms. Moreover, as reported in Figure 5.5, current multi-cloud adoption often remains limited to applications siloed (57%) or replicated (50%) across different clouds rather than a true seamless integration, which includes advanced features like cloud bursting and intelligent workload placement.

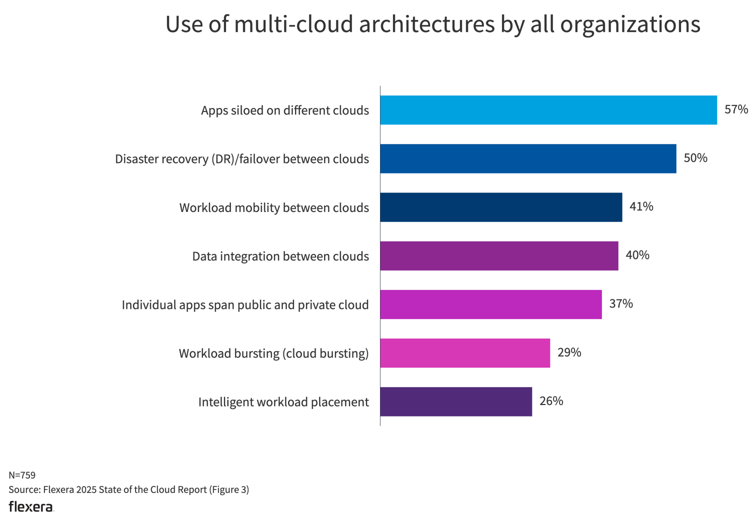


Figure 5.5: Multi-cloud adoption patterns among organizations in 2025 [4].

5.2.1 Interoperability

Interoperability between providers' infrastructures and APIs represents a primary challenge when implementing multi-cloud architectures. As mentioned in Subsection 4.2, cloud providers typically rely on proprietary APIs as part of their ecosystem governance strategy, which limits compatibility with external systems. This practice, combined with volume discounts and migration costs, reinforces vendor lock-in and discourages users from integrating multiple providers into their architecture. The lack of standardized interfaces forces organizations adopting multi-cloud architectures to manage different technology stacks, raising the need for highly skilled personnel within this environment.

In this context, open standards play a critical role in facilitating interoperability and driving innovation across the cloud industry [127]. Through open standards, it is possible to develop open APIs, which are standardized and cloud-agnostic interfaces allowing seamless communication across diverse cloud environments. These APIs are typically open source and defined through shared specifications and schemas. Examples include: *OpenAPI Specification (OAS)*, managed by the Linux Foundation, which provides APIs specifications encoded in a JSON or YAML document [128]; *GraphQL*, originally developed by Meta, which is a query language that enables the retrieval of data from multiple endpoints with a single API call [129]; *gRPC*, developed by Google and currently maintained under the CNCF, which is a high-performance framework designed to facilitate communication within distributed systems [130].

A particularly relevant European project is the *Sovereign European Cloud API (SECA)*, which seeks to establish open standards for APIs specification across major European providers such as Aruba, IONOS, and Dynamo [131]. SECA aims to enhance multi-cloud and IaaS interoperability while fostering a sovereign European cloud that mitigates dependency on hyperscalers' capacity.

5.2.2 Security

From a security perspective, multi-cloud environments expand the attack surface and increase the risk of policy misconfiguration, which can introduce vulnerabilities and expose organizations to cyber threats. Common security risks in multi-cloud environments include [132]:

- **Man-in-the-middle attacks**, in which the attacker intercepts or alters communication between cloud environments;

- **Side-channel attacks**, where vulnerabilities in a cloud environment are exploited to gain access to enterprise systems and exfiltrate data;

- **Ransomware**, where attackers encrypt data stored in cloud infrastructure and demand a ransom for its decryption.

A key challenge in multi-cloud architecture is maintaining consistent security measures across CSPs. Each provider has unique policies, encryption standards, Identity and Access Management (IAM) protocols, and compliance certifications. Within an IT system, IAM governs the provisioning of users' digital identities and determines which resources they can access. In multi-cloud architectures, IAM becomes particularly complex, as it requires developing a unified framework to integrate multiple identity systems. Organizations often adopt centralized Identity Service Provider (ISP) to manage user authentication and authorization across distributed systems.

Traditional access control methods, such as role-based access control (RBAC) or attribute-based access control (ABAC), do not provide the necessary flexibility and scalability in large distributed systems. For this reason, *Zero Trust Architectures* have emerged as a cybersecurity paradigm for cloud-native environments [133]. Zero trust provides continuous authentication and dynamic authorization for users but can be resource-intensive, as it involves real-time monitoring across the entire system. To ensure that security measures are applied throughout the software development lifecycle, DevOps and security operations must be coordinated across all stages of the process. This emerging approach is referred to as *DevSecOps* [134].

Furthermore, emerging trends such as post-quantum cryptography or AI-based anomaly detection offer promising solutions to current cybersecurity challenges and may enhance the overall security of multi-cloud environments [132].

5.2.3 Cost Management

In multi-cloud environments, each cloud provider applies its own billing system and usage metrics, which can result in complex and fragmented cost management for organizations. Moreover, the presence of several infrastructures can lead to reduced cost visibility and unused or underutilized resources, potentially generating shadow IT costs. To address these challenges, *Financial Operations (FinOps)* emerged as "an operational framework and cultural practice which maximizes the business value of cloud and technology, enables timely data-driven decision making, and creates financial accountability through collaboration between engineering, finance, and business teams" [135]. As suggested by the term, FinOps implies close collaboration between financial and DevOps teams.

These practices are guided by six key principles, which are:

1. Collaboration among cross-functional teams;

2. Alignment of business value with technology decisions;

3. Ownership for technology usage by every team;

4. Accessibility and accuracy of FinOps data;

5. Centralized FinOps reports;

6. Exploitation of the cloud variable cost model.

According to the 2025 *State of FinOps Report* [136], scaling of FinOps governance and policy emerged as the top priority for more than 30% of respondents, while multi-cloud cost reporting was cited as a priority for around 20% of organizations.

The decentralized ownership model, part of FinOps principles, helps organizations improve observability and accountability across multiple cloud environments. To support this goal, the FinOps Foundation is developing *FOCUS*, an open standard for cost and usage data across cloud vendors, which aims to reduce the complexity of operating with multiple providers.

Another financial determinant limiting multi-cloud adoption is the presence of switching costs, and particularly data egress fees. These fees are charged whenever data transit from a provider of data processing services to another and can result in a significant financial impact within the provider bill, sometimes representing more than 50% of total cloud costs for data-intensive workloads such as streaming or media services [137]. Egress fees limit flexibility and portability within multi-cloud environments and strengthen hyperscalers' bargaining power, reinforcing vendor lock-in. To promote fairer competition and enhance cloud portability, the European Commission introduced new conditions through the *EU Data Act*, which became applicable in September 2025. In particular, Article 29 mandates the elimination of switching costs for customers transitioning between cloud providers [138]. This regulation is part of the broader *European Digital Strategy* (see Section 6.2), which fosters a competitive multi-provider landscape.

5.3 Multi-Cluster Topologies

To mitigate adoption challenges and fully exploit the benefits of a unified multi-cloud architecture, organizations must develop systems capable of spanning across different providers' infrastructure. In this context, the cloud-native paradigm provides tools to package applications in a standardized and modular way, ensuring the application remains independent of the underlying infrastructure.

While cloud-native technologies enable seamless communication between microservices at the application level, achieving a truly unified multi-cloud environment also requires applying standardization layers at the level of the underlying infrastructure.

Within a single infrastructure, this challenge can be successfully addressed through K8s. However, as of 2025, there is no native K8s mechanism that allows transparent communication and interoperability between multiple clusters, which limits the development of applications distributed across multiple infrastructures.

The need for unified multi-cluster environments has therefore emerged to orchestrate and unify resources, policies, and workloads across multiple cloud providers while also avoiding a single point of failure. As reported in Figure 5.6, the use of multiple CSPs is increasing among cloud-native adopters [100], as well as the usage of multiple clusters, with approximately one third of organizations managing more than 50 clusters [139], potentially leading to cluster sprawl[2] and its associated management complexity.

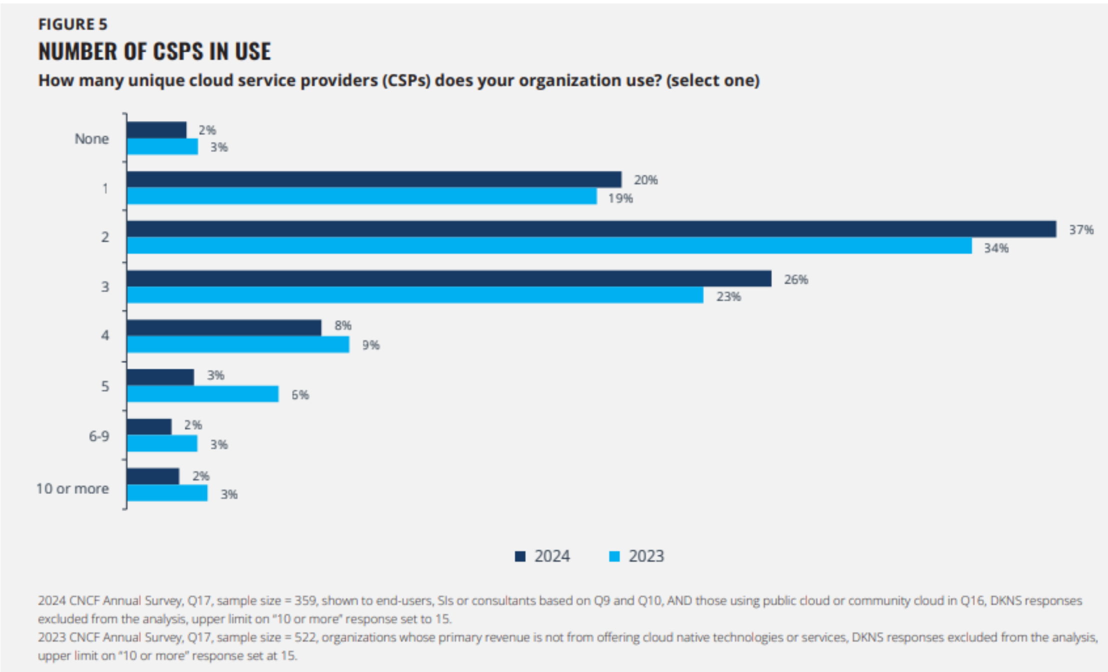


Figure 5.6: Number of CSPs in use by organizations. [100].

Additionally, 42% of organizations deploying workloads on K8s identified the management across multiple cloud environments as one of the primary operational challenges [113]. Creating a seamless multi-cloud and multi-cluster environment requires addressing multiple technical dimensions such as:

- **Networking**, to enable microservices communication across clusters;

[2]The proliferation of multiple kubernetes clusters in different environments.

- **Policy and Identity and Access Management (IAM)**, to unify authentication and authorization;
- **Orchestration and Scheduling**, to support workload offloading and replication;
- **Monitoring and Observability**, to ensure unified visibility and consistent metrics across the fleet[3].

Significant progress in this area has been driven by the Kubernetes Special Interest Group (SIG) Multi-cluster and the open source community, which are actively developing cloud-agnostic and multi-cluster solutions. SIG Multi-cluster is currently working toward a native K8s multi-cluster API, designed to abstract infrastructure dependencies and establish a standard control plane [140].

The open source ecosystem already offers several cloud-agnostic tools to address specific multi-cluster challenges, such as:

- **Submariner** and **Istio** for inter-cluster networking and service communication;
- **Open Cluster Management (OCM)** for governance and policy distribution;
- **Liqo** and **Karmada** for workload orchestration and scheduling across clusters;
- **Prometheus** and **Grafana** for monitoring and observability.

By adopting or combining these tools, organizations can create efficient multi-cluster environments. However, doing so requires high skills and deep knowledge, as each tool provides a specific set of rules, commands, and configurations.

In parallel, major hyperscalers introduced their own multi-cluster management solutions, such as Google Anthos, Microsoft Azure Arc, and EKS Anywhere. While Anthos and Arc can also manage external clusters provisioned by third-party providers, full automation and interoperability features are achieved only within the provider's own ecosystem.

5.3.1 Kubernetes Cluster Federation

Within the open source ecosystem, several projects have been developed to enable seamless multi-cluster environments, where a shared pool of cluster resources can be managed as a single system. This approach is typically referred to as a cluster federation. The term "federation" was originally introduced to describe a collection

[3]A logical group of multiple Kubernetes clusters.

of database systems unified into a loosely coupled federation to exchange and share information across distributed nodes [141].

In the context of K8s, this concept evolved into *Kubernetes Cluster Federation (KubeFed)*, an experimental project by Google and the SIG Multi-cluster community. KubeFed aimed to coordinate the configuration of multiple K8s clusters through a single API layer within a single control cluster. Despite its innovative approach, KubeFed has seen limited adoption because of its configuration, management complexity, and limited scalability, and currently, the project is not under active development [142]. In particular, KubeFed required the user to manually select the target clusters for each workload and define the associated policies individually. However, the project introduced foundational design principles, such as the use of K8s native APIs and centralized policy management, which were inherited by its successor *Karmada (Kubernetes Armada)* [143]. Figure 5.7 illustrates Karmada's architecture and main components, including the control plane, typically deployed in a dedicated cluster, and the agent running on the member clusters.

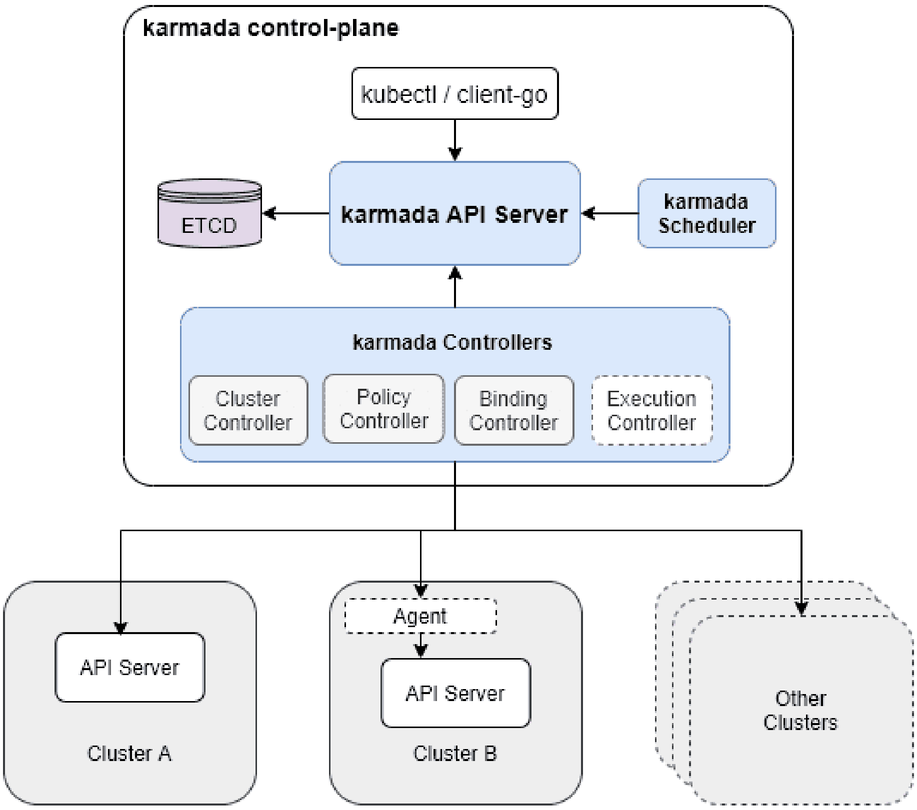


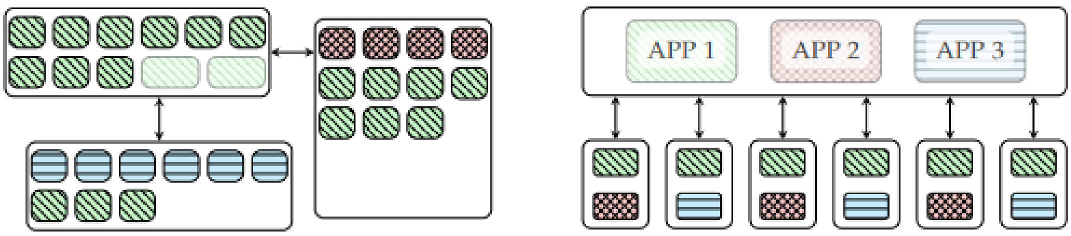
Figure 5.7: Karmada architecture and main components [144].

Karmada was released in 2021 and is currently an incubation project under the CNCF. The system extends KubeFed's original capabilities by introducing a global scheduler that dynamically assigns resources across clusters based on propagation policies defined by the user [144].

5.3.2 Liquidness Is All You Need

An alternative approach to centralized and hierarchical federation architectures was introduced with the open source project *Liqo*, developed at Politecnico di Torino and currently maintained within ArubaKube. The new approach builds upon the paradigm of liquid computing, which fosters a continuum of resources and services achieved through a peer-to-peer and decentralized architecture [145].

Liquid computing enables a fluid topology and provides flexibility in terms of deployment scenarios. Possible configurations include an "elastic cluster" (Figure 5.8a) in which a K8s cluster can dynamically offload workloads to external clusters, leveraging additional resources, and "super cluster" (Figure 5.8b), composed of multiple interconnected clusters orchestrated by a unified control plane as a single large distributed system, in which every cluster retains its autonomy.



(a) Elastic cluster configuration. **(b)** Super cluster configuration.

Figure 5.8: Two possible Liqo deployment scenarios [145].

Liqo represents the instantiation of the liquid computing approach in the context of cloud-native environments, enabling dynamic, decentralized, and seamless multi-cluster topologies. Compared to traditional federation frameworks, Liqo represents a cutting-edge solution integrating networking and portability of workloads at runtime, through dynamic offloading and cluster discovery [146]. Liqo follows a modular architecture which includes four main functional modules [8]:

- **Peering**: establishes peer-to-peer connections between clusters using VPN tunnels, without the need to manage complex manual configurations.
- **Offloading**: enables seamless workload offloading through the virtual node abstraction, which exposes remote clusters as local nodes within the home

cluster. This design allows scheduling workloads on remote clusters using native K8s tools.

- **Network Fabric**: facilitates pod-to-pod and service-to-pod communication across clusters, making distributed workloads appear as if they were running locally.
- **Storage Fabric**: provides a native multi-cluster storage layer to manage stateful workloads. In K8s, because of the ephemeral nature of pods, stateful workloads'[4]information are saved into persistent volumes which are typically tied to a single cluster. Liqo extends this capability allowing the requested volume to be created and provisioned on either the home or remote cluster, depending on where the stateful workload is scheduled.

Through these mechanisms, Liqo enables resource sharing and seamless workload portability across clusters based on the available capacity, while maintaining compatibility with native K8s approaches. The core enabler of this functionality is the virtual node abstraction, implemented through the *Virtual Kubelet* [147], which enables the control plane to see the remote cluster as a local node, as illustrated in Figure 5.9.

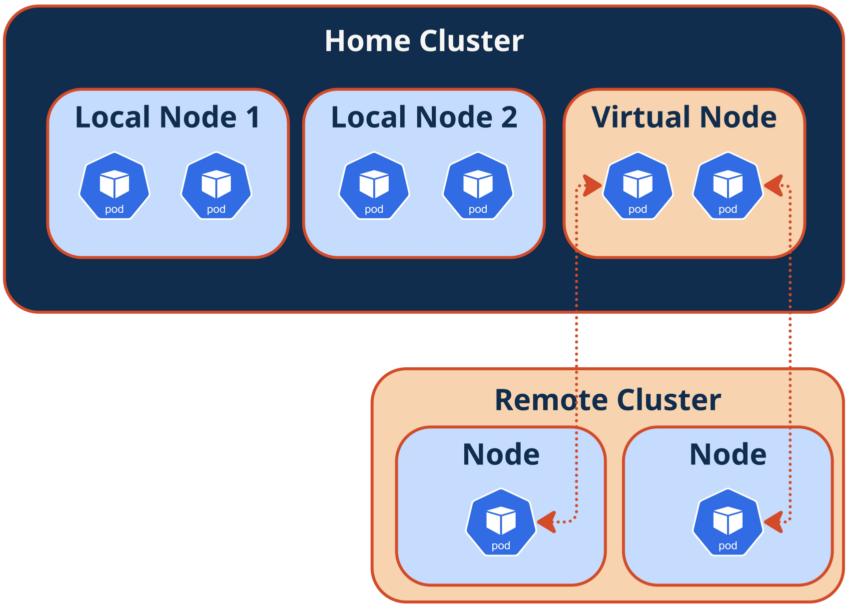


Figure 5.9: Liqo Virtual cluster architecture (adapted from [146]).

[4]Workloads requiring persistent data storage across user sessions (e.g. database instances)

Liqo's distributed and peer-to-peer architecture, combined with its dynamic offloading and discovery capabilities, makes it a natural fit for edge workloads, which are increasingly being deployed as K8s clusters. This approach enables the cloud continuum paradigm, which can be defined as "an extension of the traditional Cloud toward multiple entities (e.g., Edge, Fog, IoT) that provide analysis, processing, storage, and data generation capabilities" [148].

5.3.3 Diffusion of Kubernetes Federated Environments in Europe

By cross-referencing available market studies and industry statistics, it is possible to derive a plausible forecast for the diffusion of K8s federated environments in Europe. The main potential adopters of this technology are organizations that operate multi-cluster solutions spanning across heterogeneous infrastructure, either edge, on-premises or public cloud environments. Table 5.1 presents the forecasted Total Addressable Market (TAM) for Europe from 2024 to 2032, expressed in terms of the estimated number of potential adopters.

Table 5.1: Forecasted Total Addressable Market (TAM) for K8s Federated Environments in Europe.

Parameter	**2024**	**2032**
Market size ($B)	2.11	11.78
K8s adopters (worldwide)	61,397	263,995
European share	20%	20%
Potential federation adopters (% of K8s orgs)	45%	60%
TAM (K8s $\times$ EU share $\times$ federation adopters %)	5,525.73	31,679

The market size and the resulting CAGR (24%) were derived from the SkyQuest Global Kubernetes Market Report [149], along with the European market share. According to 6sense [150], the number of K8s customer firms is estimated to be 61397, while the number of adopters projected for 2032 was calculated adjusting the previously derived CAGR (based on market revenues), obtaining a 20% customer growth. The CNCF 2024 Annual Survey [100] reported that 80% of organizations run K8s in production, and 78% use more than one CSP. The intersection of these groups suggests a reasonable upper bound of approximately 60% for the potential adopters of K8s federation technologies; the remaining 40% is expected to run all K8s production workloads on a single cluster or within a single provider's managed service.

As of 2024, based on the average observed distribution for organizations operating in multiple environments [4], a lower bound of approximately 45% can be assumed for the potential adopters among K8s users. While overall K8s adoption will increase the absolute number of potential customers, the relative percentage growth is primarily driven by the rising importance of compliance and data sovereignty within the EU, as well as the role of multi-cloud, edge, and AI K8s deployments [139]. To illustrate the potential diffusion of K8s cluster federation in Europe, the Bass diffusion model [151] was applied to simulate yearly adoption patterns. The model produces an S-shaped cumulative curve (orange line) and a corresponding bell-shaped frequency curve (blue line), as illustrated in Figure 5.10.

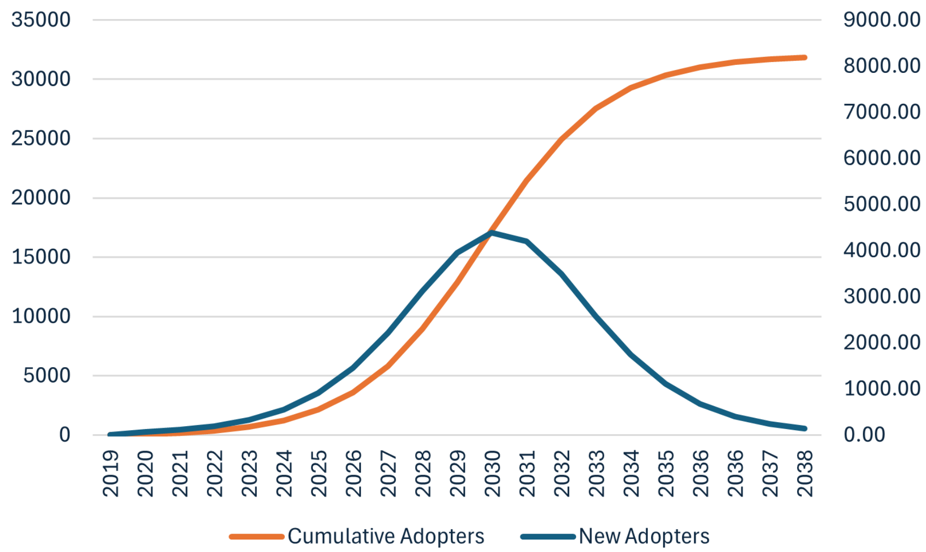


Figure 5.10: Bass diffusion model for the potential adoption of K8s federation.

The model is defined by the following formula:

$$N(t) = M \cdot \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}} \tag{5.1}$$

where:

- $N(t)$ represents the cumulative adopters at year t;
- M is the potential customer base at market maturity;
- p is the innovation coefficient;
- q is a parameter for the imitative diffusion.

The results from Table 5.1 were used to estimate M (approximately 32,000 potential adopters). Parameters p (0.0015) and q (0.55) were derived by calibrating the model so that the number of cumulative adopters in 2024 and 2032 matched the potential federation adopters, adjusted by a conversion rate between 20% and 25%. The considered time horizon is 20 years, which aligns with the maturity observed for similar B2B technologies, such as K8s itself, which is still transitioning toward late majority adoption in Europe nearly a decade after its release. The starting year ($t_0 = 2019$) corresponds to the first release of KubeFed. The resulting curve from the Bass diffusion model allows to chronologically categorize adopters (Table 5.2), following Rogers' adopter categories: innovators, early adopters, early majority, late majority, and laggards [34].

Table 5.2: Categorization of adopters based on the Bass diffusion model.

Category	Cum.%	Cum. adopters	Time interval	Adopters profile
Innovators	2.50%	800	2018–2023	K8s community; cloud-native startups
Early adopters	16.00%	5 120	2024–2027	Platform Engineers in compliance/latency-sensitive industries
Early majority	50.00%	16 000	2028–2030	Mid-to-large enterprises with mixed environments
Late majority	84.00%	26 880	2031–2033	Traditional enterprises and government IT
Laggards	100.00%	32 000	2034–2038	Legacy-bound organizations and SMEs

5.3.4 Innovation Phase Analysis

To complement the diffusion model, an analysis of innovation dynamics was conducted to assess the technological maturity of K8s federation and multi-cluster solutions. The study focuses on patent activity as an indicator of the innovation rate intensity, reflecting primarily the efforts of firms' R&D departments and research institutions toward the development of commercial technologies, rather than the innovation emerging from the open source community.

The search was conducted based on the priority date, using Google Patents for the period 2019–2025 (up to November 1, 2025) and the following domains and key-words:

- **Kubernetes:** Title + Abstract + Claims (TAC) = "Kubernetes";
- **Kubernetes multi-cluster:** TAC = "Kubernetes" + "multi-cluster" in full document;
- **Kubernetes federation:** TAC = "Kubernetes" + "federation" in full text.

Table 5.3: Patent publications by keyword, 2019–2025 (source: Google Patents, November 2025).

Key-word	2019	2020	2021	2022	2023	2024	2025*
Kubernetes	229	1136	1379	1876	1772	1322	470
Kubernetes multi-cluster	198	405	486	624	682	567	328
Kubernetes federation	44	116	146	133	149	136	154

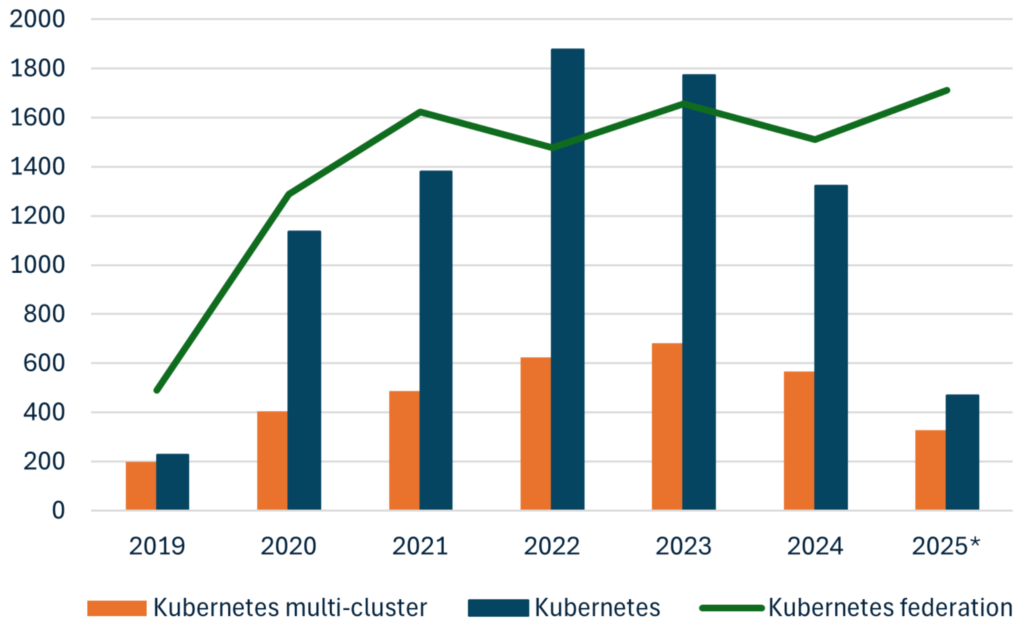


Figure 5.11: Patent trends for Kubernetes, Kubernetes multi-cluster, and Kubernetes federation (2019–2025).

The results in Table 5.3 and Figure 5.11 suggest distinct innovation phases for the three technological domains:

Kubernetes: The steady decrease in the number of patents from 2023 to 2025 (about 74%) suggests the stabilization of a dominant design. The technology is currently in the specific phase, with innovation efforts focusing primarily on the integration into enterprise workflows through the commercialization of mature solutions, such as managed K8s services.

Kubernetes multi-cluster: Patent trends show a similar pattern, with peak activity around 2022–2023. Despite a 48% decrease from 2023 to 2025, the relative share of multi-cluster patents increased from 38% to 70% among all K8s-related patents. This suggests a transitional phase for this domain, with vendors such as SUSE, RedHat, and VMware proposing similar solutions based on cluster fleet management platforms as the commercial dominant design.

Kubernetes federation: Within this domain, patent data exhibit an inconsistent pattern between 2021 and 2024, and it represents the only category showing a growth between 2024 and 2025. These results suggest that the technology is still in the fluid phase, with the major contributions and experimentation originating from open-source ecosystem niches and universities, while enterprise adoption is still at an early stage. A dominant design, as well as interoperability standards and commercial products, have not yet emerged in this domain.

Chapter 6

Federated Multi-Provider Ecosystems

As defined by NIST, "In its most general sense, federation could support the sharing of arbitrary resources, from arbitrary application domains with arbitrary consumer groups across multiple administrative domains." [10] The distinction between K8s federation and cloud federation is conceptually analogous to the distinction between a product platform and an industry platform (see Section 3.5). Multi-cluster K8s federation enables a shared infrastructure platform, where clusters represent modular resources used to deploy workloads across different providers. Cloud federation, similar to an industry platform, extends this concept to the governance of relationships and complementarities between cloud and edge providers, establishing common governance frameworks.

The governance of federations is a key factor in determining how they will exist within a broader ecosystem. A federation should operate within an administrative domain composed of an Identity Provider, a CSP, and a Cloud Consumer or user. The defining characteristics of a cloud federation include [10]:

- Distributed governance without a formal "owner";
- Membership and identity credentials managed by an Identity Service Provider (ISP);
- Participants discoverable and accessible to other members through the sharing of resources or metadata;
- The agreement on governance and common goals based on defined roles, attributes, and policies.

Figure 6.1 illustrates the architecture and value system of a Federated multi-provider environment. From the customer firm's perspective, providers no longer represent "the hub" but rather interchangeable modules providing compute, storage, and platform resources. The relationship between providers may be either cooperative (e.g., a Telco provides edge nodes for delivering cloud provider services) or competitive (e.g., two cloud providers offering similar services).

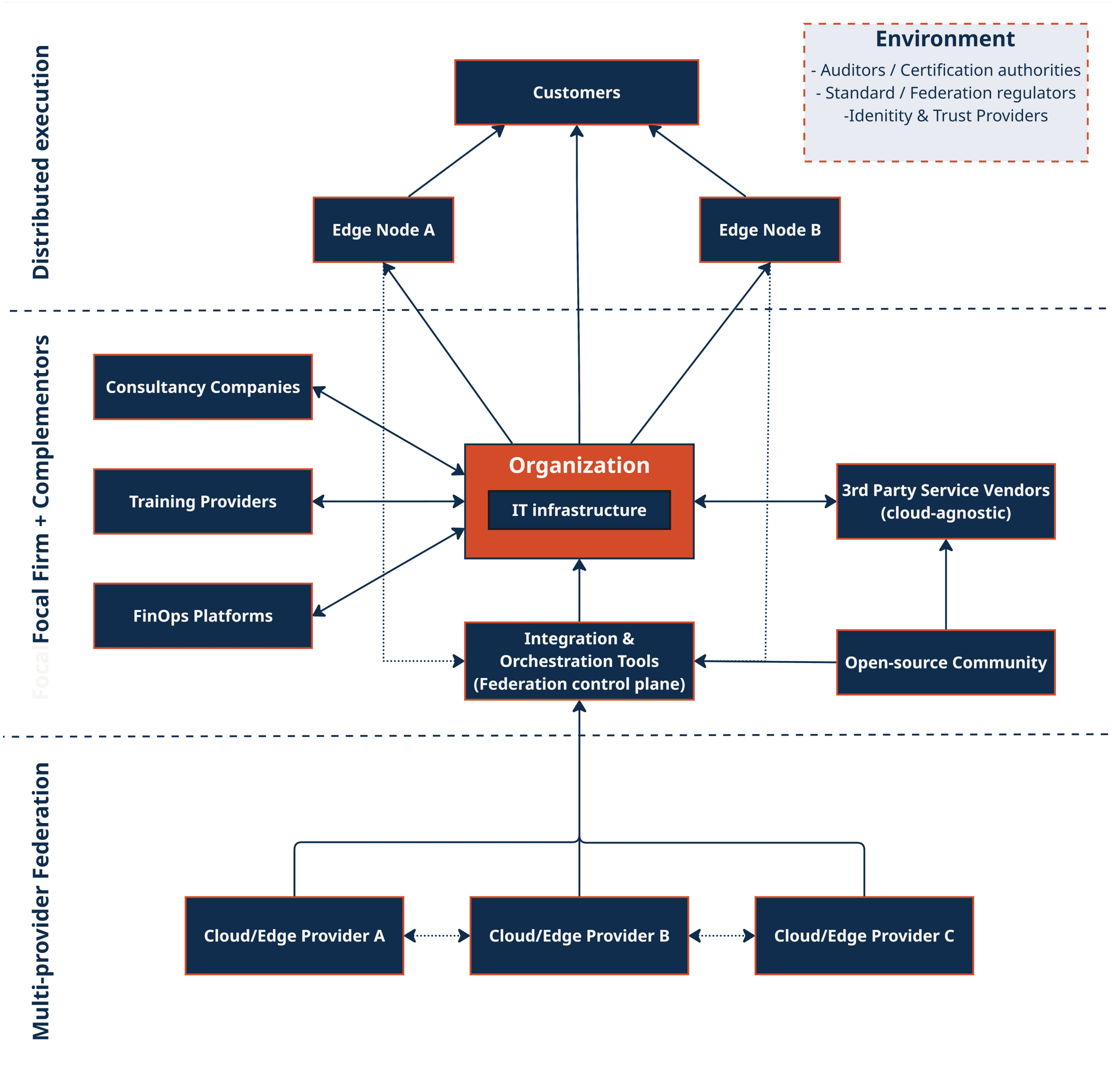


Figure 6.1: Federated multi-cloud architecture and value system.

This configuration enables a firm-centric ecosystem, where organizations adopting cloud services orchestrate resources and workloads across multiple clouds. Rather than being passive endpoints, they actively choose, mix, and arbitrate among cloud or edge providers via the federation control plane. Therefore, firms can strategically manage complementarities to build their own IT ecosystems, mitigating the dependency on single cloud providers. By leveraging edge nodes as distributed execution layers, the firm can deliver digital services and applications closer to end users. Edge nodes may also be managed as K8s clusters from the same federation control plane. Through the seamless integration between edge and cloud computing, the overall ecosystem follows the principle of the cloud continuum described in Subsection 6.2.3, enabling fluid interoperability and governance across infrastructures.

6.1 Five Forces and Competitive Dynamics

As anticipated in Chapter 3, Porter's Five Forces [20] provides a useful framework for analyzing the competition within a given industry. From the perspective of cloud providers, the emerging ecosystem model has the potential to reshape the existing competitive dynamics, as illustrated in Figure 5.4.

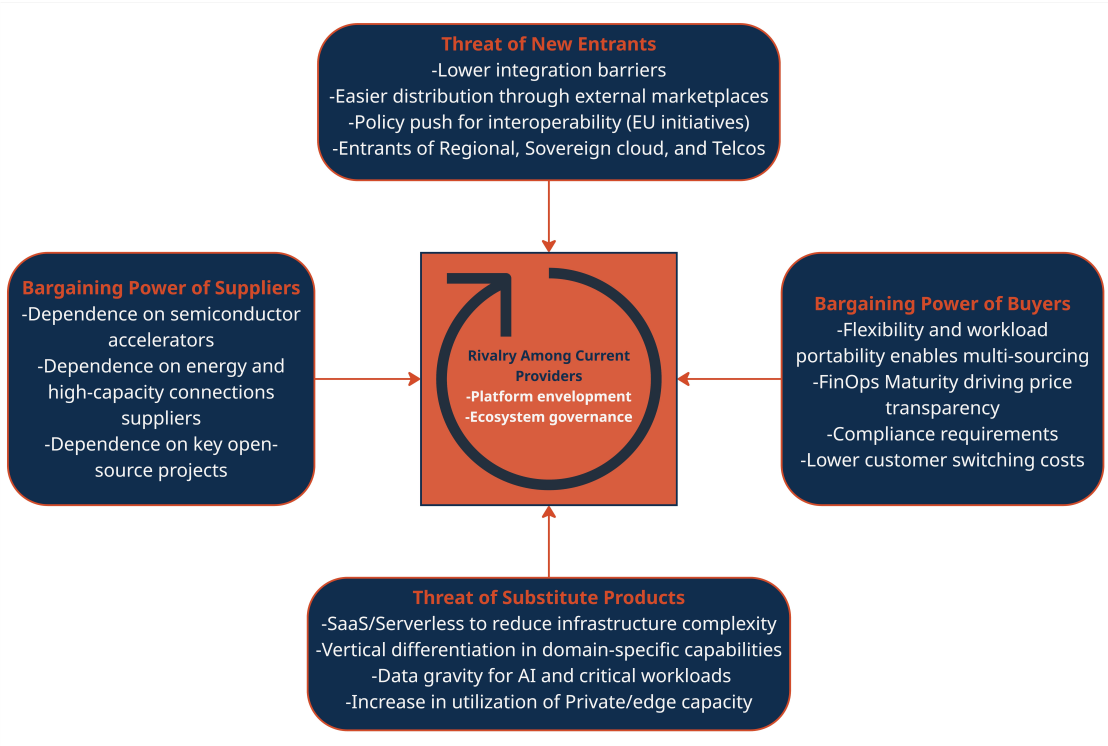


Figure 6.2: Competitive dynamics in cloud ecosystem, adapted from [20].

The current cloud market is dominated by a few hyperscalers (see Figure 4.6), yet interoperability between multiple infrastructures, enabled by federated environments and regulatory push, can lead to potential threats of new entrants. Regional, sovereign cloud, or Telco providers offering edge infrastructures can leverage standardized APIs, interconnected infrastructures, and federated cloud marketplaces to gain a relevant position within the industry. Concurrently, the growing demand for semiconductor accelerators (e.g., GPUs and TPUs) required by AI workloads, along with the increasing consumption of energy and networking resources within providers' data centers, can weaken hyperscalers' position, increasing the supplier bargaining power. Additionally, a great portion of their service offering and infrastructure depends on open source projects, which can lead to potential governance and licensing risks.

Concerning the threat of substitute products, the shortage of skilled personnel in organizations and the complexity of managing cloud infrastructure can favor the adoption of SaaS or serverless solutions. Moreover, specialized providers offering cutting-edge AI capabilities or adhering to domain-specific compliance requirements may offer a higher degree of vertical differentiation. Within a federated environment, their services can be deployed either on third-party, private, or edge infrastructures, which significantly reduces barriers to entry. Data gravity refers to the tendency of data to attract services and applications due to technical or economic constraints [152] and plays a crucial role in the competitive dynamics within cloud providers. While this principle can also reinforce hyperscalers' vendor lock-in, in dynamic and federated computing environments such as those enabled by Liqo, it is possible to offload computing resources and consume services leveraging the infrastructure of the data owner, as qill be discussed in Subsection 6.3.2 [153].

From the demand side, buyer bargaining power also increases as they are able to orchestrate multiple providers via the federation control plane. The adoption of open APIs enables flexibility, interoperability, and workload portability within multiple infrastructures. Additionally, the increasing maturity of FinOps practices drives higher price transparency and cost visibility, addressing one of the main pain points faced by organizations when adopting hyperscalers' cloud services 5.4.

To preserve their competitive position, existing providers can adopt two complementary strategies: The first is platform envelopment (see Chapter 3.5.2), achieved through the bundling of cutting-edge AI services with infrastructure plans, integrating multi-cluster orchestration in their marketplaces, and pursuing strategic partnerships and initiatives with sovereign clouds or regional providers. The second involves ecosystem governance through openness and modularity, balancing the need to attract complementors with preserving proprietary capabilities to ensure value capture.

Within the European landscape, policy initiatives play a crucial role in shaping competition and value distribution between providers. In particular, through multiple regulatory frameworks, the European Digital Strategy aims to reduce customer switching costs and favor an agile and dynamic competitive environment.

6.2 European Digital Strategy

The concept of data sovereignty refers to the principle that data collected, stored, and processed by an organization are subject to the law, regulations, and best practices of the jurisdiction in which these data are physically located. Beyond the regulatory dimension, data sovereignty concerns the protection of sensitive or private data, ensuring they remain under the control of their owner within the specified country. Data sovereignty is included in the broader concept of digital sovereignty, which extends the same principle to the entire digital ecosystem. Digital sovereignty can be defined as "the possibility of independent self-determination by the state and by organizations with regard to the use and structuring of digital systems themselves, the data produced and stored in them, and the processes depicted as a result" [154].

For organizations outsourcing their IT infrastructure through cloud computing services, digital sovereignty represents a key dimension shaping digital governance strategies. Furthermore, it also constitutes a major concern for policymakers and regulators. Since February 2020, the European Commission has promoted a comprehensive digital strategy spanning across multiple technological sectors, such as cloud computing, AI, cybersecurity, and quantum technologies. Within the European strategy, "digital sovereignty" refers to Europe's ability to act independently in the digital world, fostering innovation while reducing dependencies on non-EU hyperscalers to promote a fairer and more competitive cloud computing market.

The European Digital Strategy encompasses a set of regulations:

- The Digital Services Act [155] aims to regulate digital platforms and intermediaries such as online marketplaces, app stores, and social networks, focusing on transparency regarding algorithms and content generation for users.

- The Digital Markets Act [156] addresses the competitive environment within digital ecosystems dominated by large gatekeepers, promoting interoperability and fairness to prevent vendor lock-in. Gatekeepers exercise control over the ecosystem, acting as private rule-makers and controlling access to users and markets for other businesses.

- The Data Act [138] aims to stimulate data-driven innovation by improving data availability, enabling both organizations and individuals to exercise greater control over their data. Additionally, it defines the interoperability requirement for common European data spaces.

- The Data Governance Act [157], which complements the Data Act, establishes a common governance framework to ensure trusted data sharing and pooling across different industries, which are crucial determinants for the development of European data spaces.

6.2.1 Data Spaces

The concept of data spaces originated in Germany as part of Industry 4.0 initiatives, which led to the establishment of the International Data Spaces Association (IDSA). The term data space refers to “a type of data relationship between trusted partners who adhere to the same high-level standards and guidelines in relation to data storage and sharing within one or many vertical ecosystems” [158] In this context, the IDSA is a non-profit organization that aims to promote open standards, existing technologies, and governance frameworks to facilitate a trusted business ecosystem that favors observability and interoperability between actors. The reference standard for data space architectures is the IDSRAM (Reference Architecture Model) which is structured structured into five layers:

- **Business layer:** defines the roles, activities, and interactions within the actors involved in the data space.
- **Functional layer:** specifies the functional requirements of the International Data Space (such as trust, data sovereignty, interoperability, and value-adding applications) and the corresponding features to be implemented (e.g., identity management, authentication and authorization, data exchange, application provisioning, and implementation).
- **Information layer:** provides a conceptual model and common vocabulary to share common knowledge and favor interoperability among participants.
- **Process layer:** describes processes and interactions taking place among the components of the data space.
- **System layer:** decomposes the data space into logical software components, favoring integration, configuration, deployment, and extensibility (see Figure 6.3).

The overall data space is composed of a network of IDS connector, which act as the data endpoints for an organization without relying on central data storage. Data providers transfer owners’ data into the data space according to usage policies that ensure the owner’s control and sovereignty over data. The connector can be implemented as hardware or as a virtual machine hosting containerized applications. Applications can be either built by the organization or downloaded from a secure platform, the app store. Other main components include the clearing house, an IDS connector that manages financial transactions, billing, and usage the metadata broker, which provides information about an IDS connector, its owner, and metadata, ensuring availability for other participants; the Identity Provider, which validates the identity and manages the access of participants based on IAM.

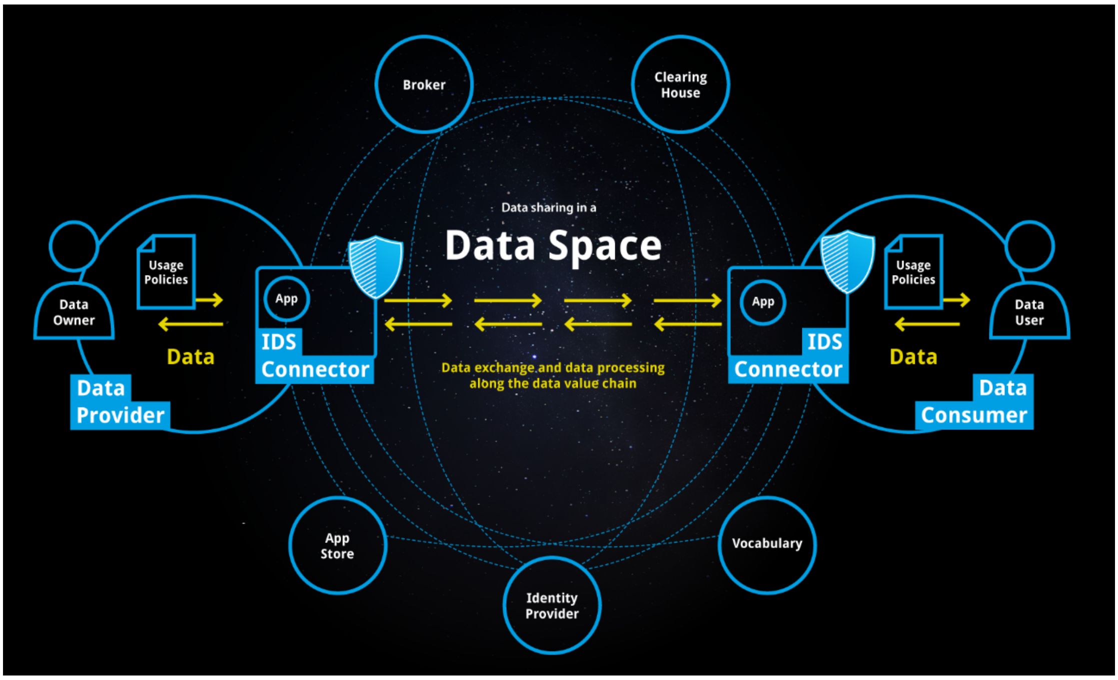


Figure 6.3: Reference architecture model of data spaces [158].

Data spaces are increasingly being deployed across various sectors as B2B platforms. Each data space can differ in size and purpose and requires a shared governance model between participants according to the specific use case. Governance models can be classified as centralized, federated, or decentralized (see Figure 6.4).

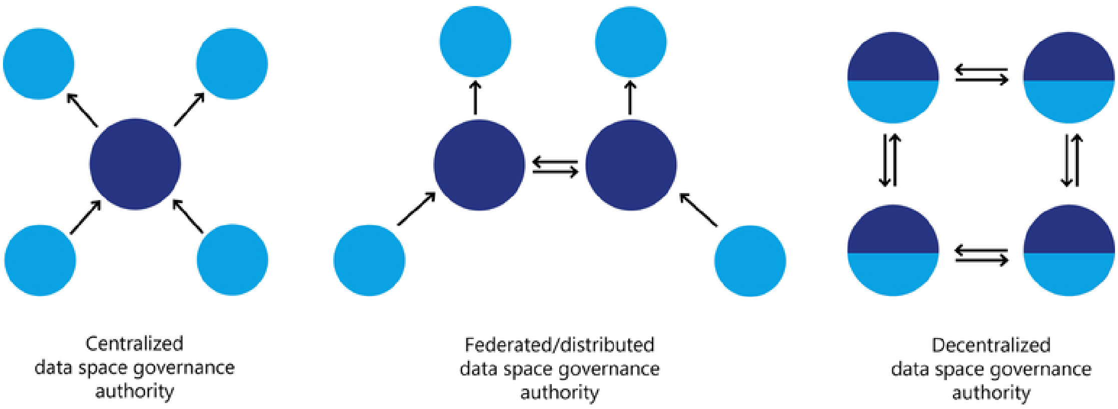


Figure 6.4: Governance models for data spaces [158].

In a centralized governance model, a single entity defines rules, manages onboarding, and ensures compliance, defining standards and interoperability frameworks. A federated governance model relies on a federation authority that establishes common standards and frameworks, collaborating with trusted third-party services to ensure compliance. Under this model, each participant retains control over its data and determines the conditions for sharing them with other members. This approach aligns with the principles of data sovereignty and the standards defined by the IDSRAM, promoting interoperability across organizations while maintaining local autonomy. Finally, decentralized governance models are currently in an experimental phase and leverage peer-to-peer mechanisms such as blockchain and distributed technologies to enable participants to negotiate rules dynamically and form self-governed ecosystems.

6.2.2 Toward Federated Data Spaces

The effort of the European Commission is directed towards the establishment of federated data spaces, which play a pivotal role within the European Data Strategy. The ultimate objective is to establish "a European data space as a genuine single market for data where personal and non-personal data, including sensitive business data, are secure and businesses have easy access to high-quality industrial data, boosting growth and creating value of a common European market for data interconnecting the existing data spaces"[159] Data spaces are currently being developed across 14 strategic sectors, including health, mobility, manufacturing, and public administration [1]. This initiative aims to overcome the current fragmentation of markets and Member States, as well as the limited availability of industrial data, identified as determinant assets for the evolution of business activities. The ambition is to position Europe as a global digital leader, fostering the world's most secure and dynamic data-agile economy [160]. The strategy is based on two pillars: the development of cross-sectorial governance frameworks, to prevent fragmentation of internal and the investment plan for cloud and edge federated infrastructures, which serve as the technological backbone for shared and distributed data platforms.

To support the implementation of common guidelines and ensure consistency in the data economy, the IDSA initiated the OPENDEI project. The initiative defines common design principles for data spaces, aiming to align the conceptual and architectural design of data spaces with the European digital strategy [161]. Since users and organizations typically operate across multiple ecosystems, data spaces are often overlapping or nested and cannot be conceived as isolated silos of data.

[1] The full list is available at digital-strategy.ec.europa.eu

The design principles proposed by OPENDEI include a decentralized soft infrastructure, which is an invisible layer made of a collection of technology-neutral and sector-agnostic agreements. The decentralized soft infrastructure is structured around two categories of building blocks [161]:

- **Technical building blocks:** enable the implementation of a modular architecture within a data space, where additional systems and innovative services (such as data analytics or IoT networks) are implemented plug-and-play to provide new business capabilities to participants tailored to domain-specific needs;
- **Governance building blocks:** define the common guiding principles such as federation, interoperability, collaboration, and scalability, valid across every domain and data space instance.

Among the major initiatives supported by the European Commission fostering federated cloud and data infrastructures is Gaia-X, which aims to "create the de facto standard to enable federated and/or decentralized and trusted data and infrastructure ecosystems, by developing a set of specifications, rules, policies, and a verification framework" [162] The project promotes the development of a trust framework reference architecture to enforce compliance and favor interoperability across ecosystems and data spaces.

Projects fully aligning with Gaia-X standards and principles are referred to as "Lighthouse data spaces." Examples include AgrospAI in agriculture, Catena-X in automotive, and Pontus-X in manufacturing. Beyond data spaces, Gaia-X principles also apply to projects in the European CSP ecosystem, such as Dynamo, a B2B platform that promotes collaboration, trust, and innovation among European CSPs (such as Ionos, Aruba, and OVH). Dynamo has launched the SECA API project to favor interoperability among provider services (see Subsection 5.2.1) and support the creation of a common marketplace.

6.2.3 Multi-Provider Computing Continuum

To foster the development of the infrastructure and data processing services required to support the innovative European data ecosystem, the European Commission approved the Important Project of Common European Interest (IPCEI)-Cloud Infrastructure and Services (CIS) [163]. The initiative encompasses 19 projects involving 12 EU Member States, aiming to develop a competitive cloud-edge ecosystem, a foundational element of the European vision of digital sovereignty. The primary initiative within IPCEI-CIS is 8RA, which aims to develop a Multi-Provider Cloud-Edge Continuum following the principles, standards, and reference architecture defined under the IPCEI-CIS program. Figure 6.5 shows a conceptual framework for the IPCEI-CIS reference architecture.

The architecture is structured into seven components and four cross-cutting domains spanning a layered stack [164].

Components:

- **Application Layer:** provides functionalities required by end-user applications.
- **Data Layer:** includes data collection, processing, and exchange functionalities.
- **AI Layer:** integrates cutting-edge solutions and advanced AI functionalities to support operations within the data layer.
- **Service Orchestration Layer:** integrates and manages multiple cloud and edge services across different providers, crucial for large distributed systems.
- **Cloud Edge Platform Layer:** enables lifecycle management of resources across the cloud-edge continuum, ensuring interoperability and portability.
- **Virtualization Layer:** provides the abstraction necessary to enable sharing and dynamic allocation of hardware infrastructure resources.
- **Physical Layer:** includes the physical hardware and infrastructure-level resources required to implement the continuum.

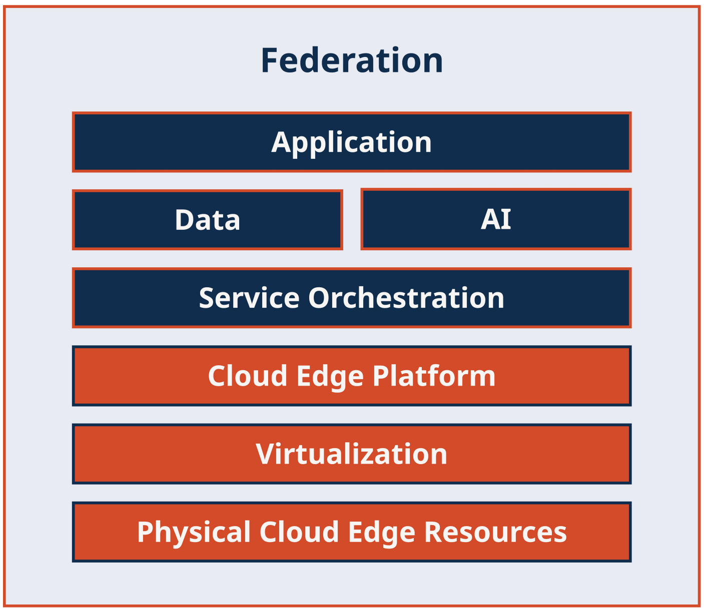


Figure 6.5: IPCEI-CIS reference architecture (adapted from [164])

Cross-cutting domains:

- **Federation Domain:** the fundamental pillar for the multi-provider cloud-edge continuum. It encompasses the components that allow the user to access and manage resources through a unified control plane, which can be either dedicated or provider-specific.

- **Management Domain:** includes capabilities and best practices necessary to operate across heterogeneous cloud infrastructures, ensuring high availability and reliability.

- **Security and Compliance Domain:** ensures a resilient infrastructure that adheres to industry-specific standards. It includes real-time monitoring, security protocols, and compliance frameworks.

- **Sustainability Domain:** focuses on hardware optimization and energy-efficient solutions to mitigate consumption and environmental risks. This is particularly relevant for training and deploying large-scale AIs model, which require substantial computational resources.

Complementary to IPCEI-CIS, the EUCloudEdgeIoT [165] initiative proposed a strategic roadmap for advancing research and innovation in the Cognitive Computing Continuum, a paradigm that integrates cloud, edge, and AI to build an intelligent and unified digital infrastructure. According to the roadmap, the Cognitive Computing Continuum spans across multiple dimensions. The technical dimension enables the integration of AI into the cloud-edge continuum, emphasizing their mutually reinforcing relationship. It also highlights the strategic role of Telcos as both consumers and edge infrastructure providers. Additionally, the paradigm promotes the adoption of open standards, open source software, and cloud-agnostic solutions to favor interoperability, as well as integrating energy consumption as the enabler of a sustainable federated infrastructure. The last dimension addresses market fragmentation, policies, and regulations, aligning with the principles of the European Digital Strategy and maintaining compliance with EU data protection laws (e.g., GDPR, Data Act). The initiative stimulates collaboration between SMEs and cloud providers to strengthen Europe's digital sovereignty.

The framework illustrated in Figure 6.6 summarizes the coordination mechanism within the European data ecosystem and the interrelationship among its key initiatives. Within the framework, the data communication layer encompasses data space projects and embodies the principle of data sovereignty, allowing secure and trusted data exchange among participants. The data processing layer provides infrastructure-level resources, which serve as the computational backbone to ensure security, flexibility, interoperability, and autonomy for all the actors within the ecosystem. The coordination promotes a federated multi-provider continuum, fostering innovation while strengthening Europe's digital sovereignty.

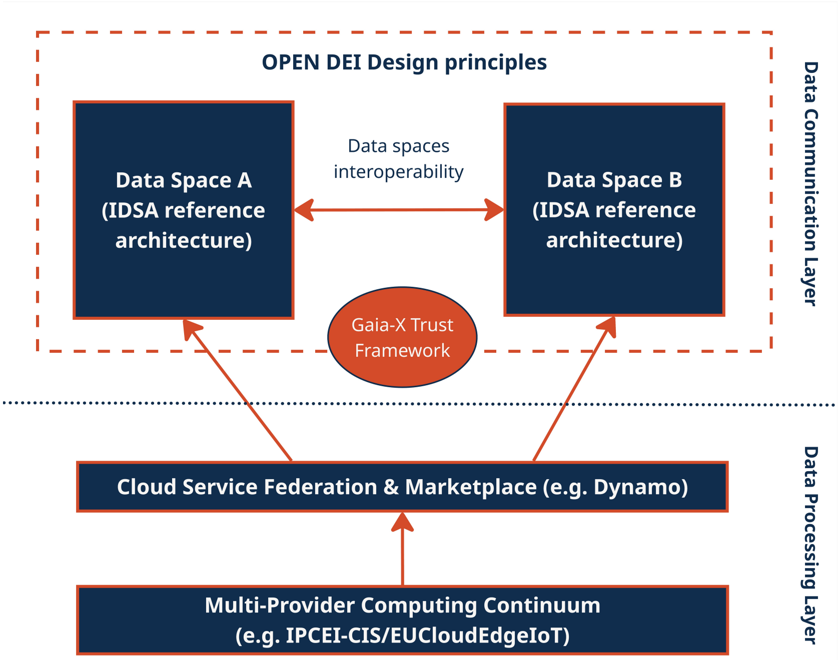


Figure 6.6: Coordination mechanisms within European data initiatives.

6.3 Innovation Management in Multi-Cloud Ecosystems

The development of a distributed data and computing ecosystem, where data and resources are shared across multiple environments and diverse infrastructures, is essential to foster innovation and enable organizations to fully exploit emerging technologies such as IoT and AI. Whether the organization is a startup seeking to scale operations globally or a large-scale enterprise establishing a CCoE or dedicated innovation unit, multi-cloud architectures enable organizations to leverage greater computing capacity, access a broader service portfolio, mitigate the risk of vendor lock-in, and optimize cost-performance trade-offs.

The integration of multi-cloud and edge computing paradigms enables the realization of large-scale and geo-distributed systems, in which edge nodes can store, process, or retrieve data from IoTs devices and transmit information to regional or public cloud providers. In such architectures, cloud infrastructures provide the computational capacity and GPUs resources necessary to deliver cutting-edge AI services. To ensure interoperability and seamless communication across the organization's ecosystem, cloud providers and intermediate computing layers, such as fog and edge infrastructures, must be aggregated in a federated multi-provider continuum, supporting unified governance and orchestration.

6.3.1 AI-Computing Continuum

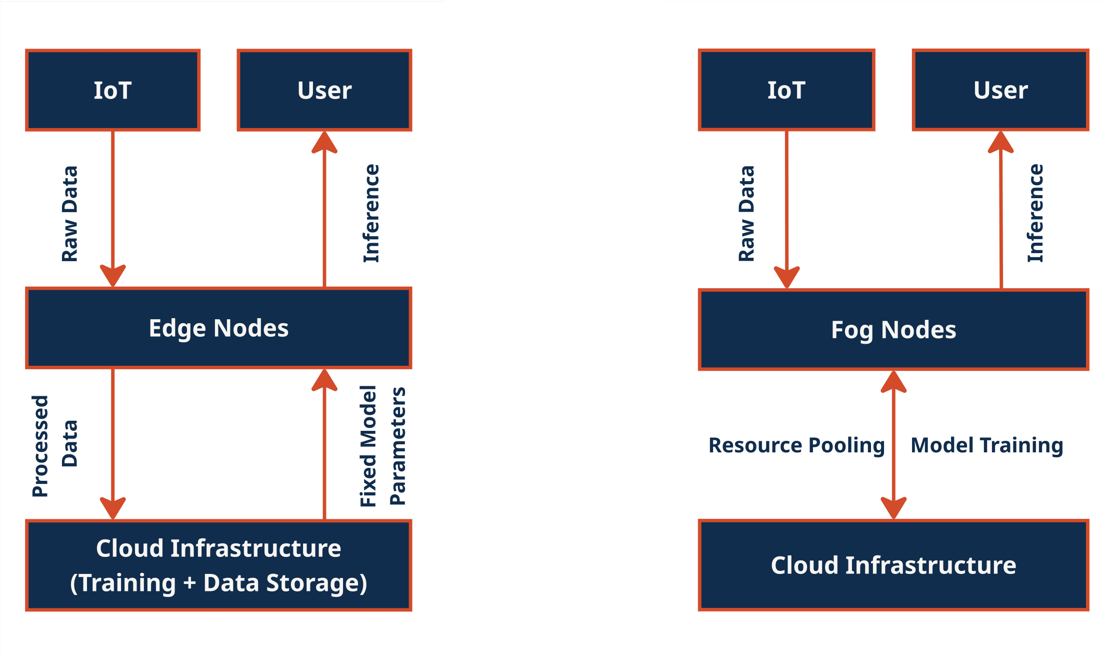
In the context of AI and ML, federated learning represents an innovative approach for implementing the AI-computing continuum, optimizing capacity and resource usage while adhering to the principle of data sovereignty. Federated learning enables AIs models to be trained on distributed datasets across heterogeneous devices and infrastructures, such as mobile devices, local nodes, or regional data centers. This approach enables the organization to share only learned parameters with a central aggregator, ensuring that raw data remains under the owner's control.

Federated learning can be achieved by combining emerging paradigms and ML techniques, particularly through the implementation of Edge-Fog-IoT and Fog-Cloud configurations.

- In the **Edge-Fog-IoT** configuration, the workflow follows a vertical data flow across three layers: IoT devices retrieve raw data from sensors and local devices, edge nodes perform inference or initial data processing closer to the source, and cloud infrastructure executes large-scale model training, global analytics, and long-term storage.

- In the **Fog-Cloud configuration**, fog and cloud layers cooperate dynamically to provide an elastic pool of shared resources, enabling a true computing continuum. The fog layer operates between edge nodes and cloud infrastructure to perform localized learning and feature extraction. This model requires high interoperability and advanced orchestration mechanisms to provide dynamic workload distribution and scheduling between fog and cloud layers.

While both configurations enhance the performance of AIs models, their use case scenarios differ based on application requirements: the Edge-Fog-IoT configuration is ideal for real-time or low-latency inference closer to the data source, while the model parameters are updated in the centralized cloud infrastructure. Conversely, a Fog-Cloud configuration enables decentralized training closer to the user, allowing continuous feedback loops that accelerate training and model responsiveness. In this scenario, a federated multi-provider infrastructure can provide additional and geo-distributed resources for the computing continuum, with Telcos or regional CSPs delivering the fog layer resources, which can provide significant cost and performance benefits.



(a) Edge-Fog-IoT configuration

(b) Fog-Cloud configuration

Figure 6.7: High-level implementations of federated learning workflows.

6.3.2 Liqo Use Cases

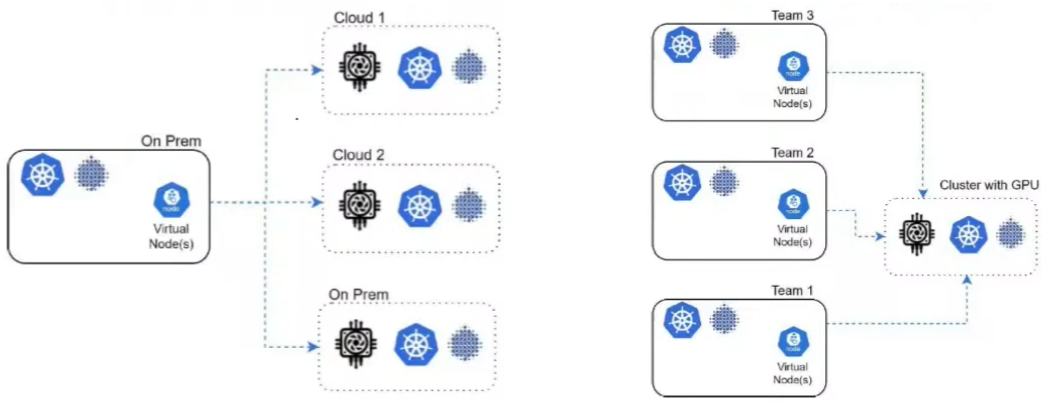
Despite the complexity, K8s multi-cluster architecture represents one of the most promising approaches to implement an infrastructure-level federation. In particular, the dynamic federation model and distributed architecture proposed by Liqo offer a strong fit for realizing the computing continuum. The following section presents three potential scenarios in which Liqo can be implemented to enable innovative solutions within this context.

Scenario 1: GPU Offloading

As discussed in previous chapters, AIs model require significant computational resources during the training phase, particularly in terms of GPUs. However, maintaining dedicated GPU clusters requires substantial capital and operational expenditure. Additionally, GPU workloads are typically subject to demand spikes during intensive training phases, while many GPUs remain in an idle state for extended periods, leading to suboptimal resource utilization. To address this inefficiency, Liqo enables a dynamic and cloud-native GPU offloading mechanism through a cluster federation, allowing resource sharing and on-demand scaling. Within the federation, organizations can leverage cloud bursting, borrowing GPU resources from other clusters, and dynamic onboarding, seamlessly integrating the resources in their workloads. As described in Subsection 5.3.2, Liqo supports flexible federation topologies that can be adapted to specific use cases [166]. Two common configurations are:

- **Multi-cloud GPU offloading:** where an on-premises or private cluster can leverage the GPU capacity from remote federated clusters across multiple cloud providers (Figure 6.8a);
- **Star GPU offloading:** where multiple teams can borrow resources from a centralized cluster or cluster federation while maintaining autonomy, isolation, and enhanced security (Figure 6.8b).

The benefits of these innovative approaches, such as multi-cloud flexibility, enhanced security, operational agility, and cost/resource optimization, can be leveraged by multiple organizations, including corporations, R&D departments, AI or cloud-native business units, or universities. For instance, Politecnico di Torino is currently implementing a full-stack solution based on K8s, Liqo, Jupyter notebooks, and MetaFlow to enable researchers to access virtual GPU capacity without migrating their entire environment.



(a) Multi-cloud GPU offloading. **(b)** Star GPU offloading.

Figure 6.8: Two possible Liqo-enabled topologies [166].

Scenario 2: Infrastructure-level Data Spaces

The second application concerns the development of infrastructure-level data spaces. While traditional data spaces reside at the data and application layers, providing communication and data transfer between participants, infrastructure-level data spaces extend the concept to enable the federation of the underlying cloud-edge infrastructure. To achieve higher interoperability and resource efficiency, organizations increasingly require federated infrastructures that are able to provide a shared pool of data processing resources.

Within this context, Liqo provides an effective solution for building infrastructure-level data spaces based on K8s cluster federation [153]. Through Liqo, organizations can establish a virtual cluster aggregating a pool of shared computing resources, in which each participant is able to retain full autonomy over the infrastructure, adhering to the principle of digital sovereignty.

For instance, a data consumer running an analytics service can offload workloads from the virtual cluster to a data owner's cluster, where the application can be executed seamlessly. To ensure isolation, network policies must be configured to restrict the communication exclusively to the offloaded services. Additionally, K8s namespaces should be defined to provide logical isolation and enforce resource quotas, ensuring that the applications executed within the owner's cluster do not exceed the maximum available capacity.

Figure 6.9 presents the high-level deployment and communication flow.

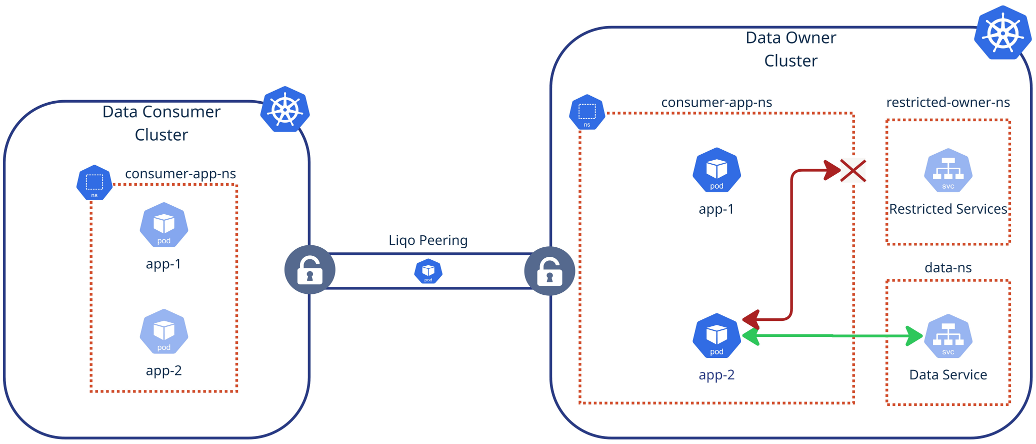


Figure 6.9: Infrastructure-level data space: deployment example and main communication flows (adapted from [153]).

Scenario 3: Myrtus

The third scenario involves the role of Liqo within the Myrtus project, part of the *EUCloudEdgeIoT* initiative [167]. Myrtus aims to develop new technologies to overcome the problem of vendor and platform lock-in, favouring portability and dynamic orchestration across the cloud-edge computing continuum. The project focuses on three main technical objectives:

1. Defining a reference architecture for the computing continuum;
2. Developing MIRTO, an AI-powered cognitive engine enabling dynamic runtime orchestration, optimizing performance and energy efficiency;
3. Designing a unified programming environment to favour cross-layer interoperability.

The Myrtus reference infrastructure (see Figure 6.10) includes three layers:

- **Cloud Layer:** provides long-term storage, subsystem monitoring, and high-performance computing resources through cloud providers' data centers.
- **Fog Layer:** enables interoperability between cloud and edge nodes through fog micro data centers and smart gateways, which provide medium- to long-term data analytics.
- **Edge Layer:** offers localized microanalytics, autonomous processing capabilities, and hardware acceleration on devices.

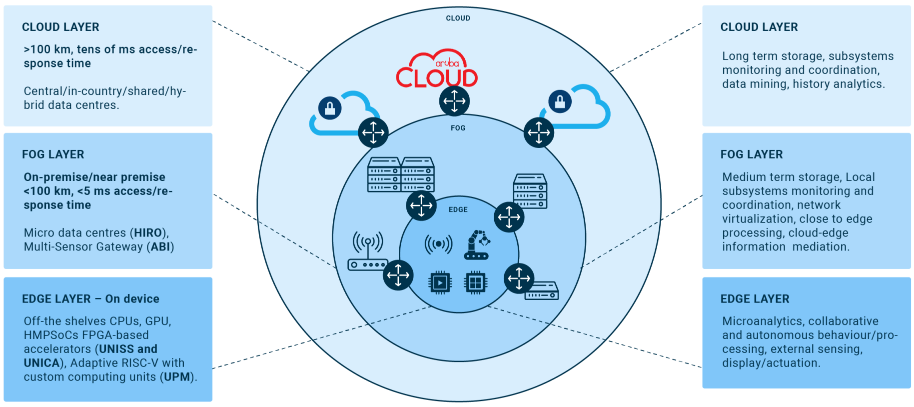


Figure 6.10: MYRTUS Reference Infrastructure [167].

Each layer utilizes Liqo to establish an infrastructure-level federation among its components and K8s clusters, thus enabling a seamless cloud-fog-edge computing continuum. The MIRTO agent enables AI-powered dynamic orchestration, managing both vertical orchestration (across layers) and horizontal orchestration (within the same layer). This combined architecture exemplifies the integration of advanced AI capabilities within the cloud-edge computing continuum enabled by Liqo, providing a high-performance and energy-efficient solution.

6.3.3 Multi-Dimensional Adoption Framework

From the perspective of organizations adopting federated environments, governance of digital infrastructure represents a crucial factor to ensure an effective implementation. In many cases, the transition requires a broader process of digital transformation aiming to the development of a federation-ready infrastructure. The effectiveness of this process depends on the firm's intrinsic characteristics and capabilities, such as size or technological maturity. The framework in Figure 6.11 aims to illustrate the ideal organizational features that enable enterprises to capture the highest value within multi-cloud or federated environments, along with high-level guidelines to support an effective adoption process.

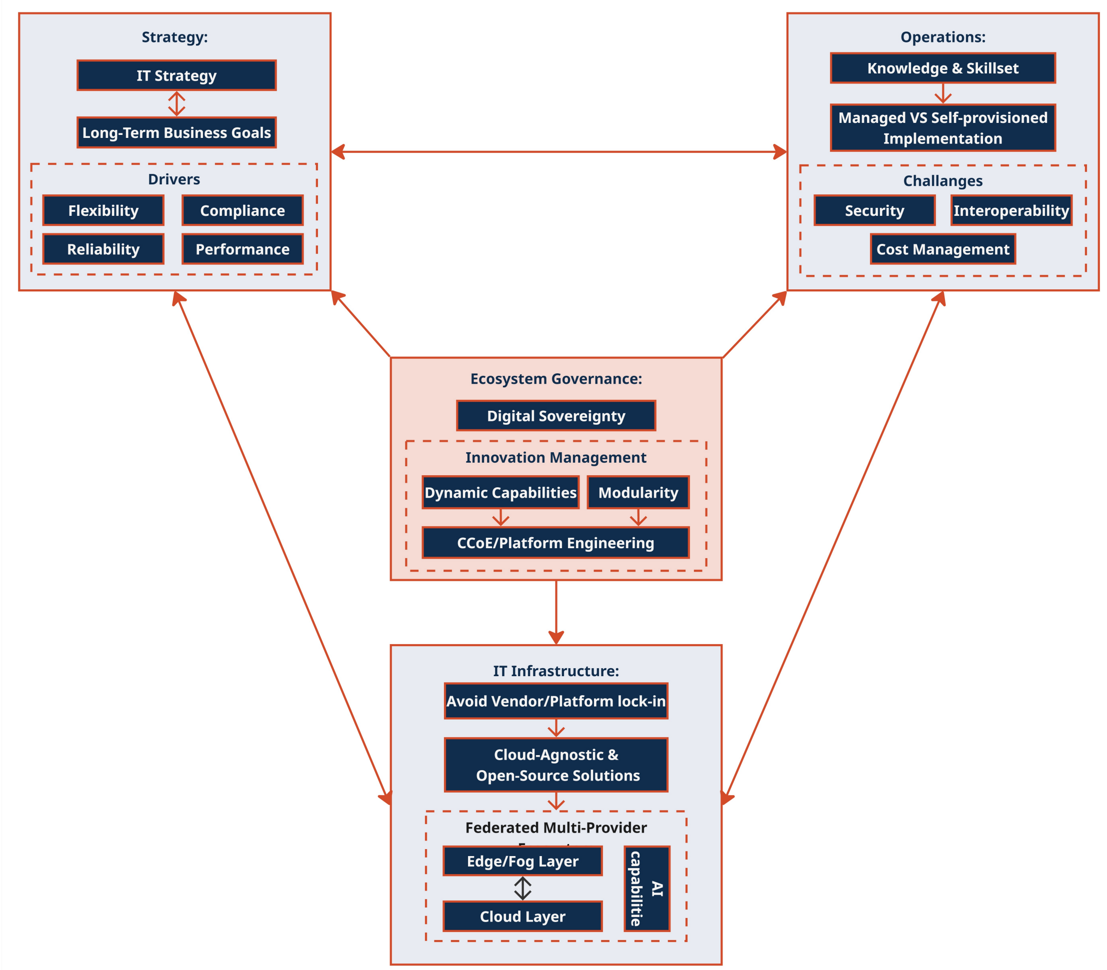


Figure 6.11: Multi-dimensional adoption framework for federated multi-provider ecosystems.

The structure of the proposed framework mirrors the TOE framework (see Figure 4.4), specifically fine-tuned for the internal governance of a federated multi-provider ecosystem. In particular, it provides a holistic perspective on the adoption of multi-cloud and federated infrastructures, integrating cloud-native principles with ecosystem governance and coordination mechanisms.

The framework encompasses three primary dimensions (strategy, operations, and infrastructure) interconnected through a central governance layer that ensures the alignment between business objectives and technological innovation.

- **Strategy:** The strategic dimension focuses on aligning the firm's IT strategy with long-term business objectives. For this purpose, firms must assess the strategic drivers for multi-cloud adoption, such as flexibility, reliability, performance, and compliance discussed in Section 5.1, evaluating how each contributes to the establishment of a competitive advantage. Notably, drivers can significantly vary depending on each organization's activities and industry sector.

- **Operations:** The operational dimension addresses the practical challenges of implementing and managing multi-provider environments. The organization must evaluate the issues, risks, and limitations concerning interoperability between heterogeneous environments and edge/fog layers, security management, and cost optimization. This process requires achieving a high level of observability and transparency across federated infrastructure, supported by internal or external knowledge and skill sets. Depending on a firm's technological capabilities, it may choose to implement the environment internally or outsource configuration and management functions to Managed Service Providers (MSPs).

- **Infrastructure:** The infrastructure dimension concerns the technical implementation of the federated ecosystem. To prevent platform and vendor lock-in, organizations should prioritize open source and cloud-agnostic solutions. By developing a federated multi-provider ecosystem, firms can leverage the computing continuum across cloud, fog, and edge layers, enabling high-performance processing and fully exploiting the benefits of cutting-edge AI solutions.

Governance acts as the central mechanism for coordination across the three dimensions. Effective governance ensures adherence to the principle of digital sovereignty, allowing firms to retain control over their digital assets. Within this context, a federation control plane avoids the fragmentation of resources and unifies management across multiple distributed environments.

Governance encompasses innovation management within the firm ecosystem, allowing the development of dynamic capabilities aligned with infrastructural and organizational modularity (see Sections 3.4 and 3.5). These principles enable the firm to coordinate centralized knowledge and strategic guidelines across decentralized implementations, fostering operational agility. This balance enable the autonomy of business units, which are able to innovate and experiment faster while preserving consistency with the central corporate and IT strategy.

From a pragmatic standpoint, digital leaders are able to achieve this equilibrium through the coexistence of a Cloud Center of Excellence (CCoE) and Internal Developer Platforms (IDPs). The CCoE defines strategic priorities and provides common governance standards to ensure alignment between the firm's technology initiatives and long-term corporate strategies. The IDPs, in turn, operationalize these principles by providing autonomous environments for business units within the governance boundaries, thereby accelerating technological innovation.

Within a similar structure, a unified multi-cloud environment enables the firm to develop digital products or internal platforms more efficiently through the seamless integration of resources and services from multiple sources.

6.4 Empirical Insights: Multi-Cloud Adoption Survey

This section presents an exploratory survey of IT practitioners designed to provide empirical idications on how the frameowk dimensions align with real-world experience. The survey was developed in collaboration with ArubaKube and implemented using a LimeSurvey Community Edition (CE) instance securely hosted on a private Aruba VPS. The survey targeted cloud specialists and managers actively operating in the industry. The distribution was conducted primarily on LinkedIn, leveraging direct engagement by both the author and the company tutor, as well as through specialized DevOps and K8s community groups.

The complete survey design, including question logic to filter respondents based on expertise and organizational cloud maturity, and full statistical outputs, are provided as a digital supplement. Given the specialist nature of the target population, the results are to be interpreted as descriptive and indicative rather than intended for hypotheses testing. Within the analysis, blank responses or those flagged as "other" or "not sure" were excluded from single-question statistics. Moreover, the number of respondents per question varied according to previous selections and branching logic.

All the responses from the survey were analyzed in aggregate, ensuring the non-identifiability of individual participants. Figure 6.12 illustrates the custom PHP template designed for the survey, including the data privacy disclaimer.

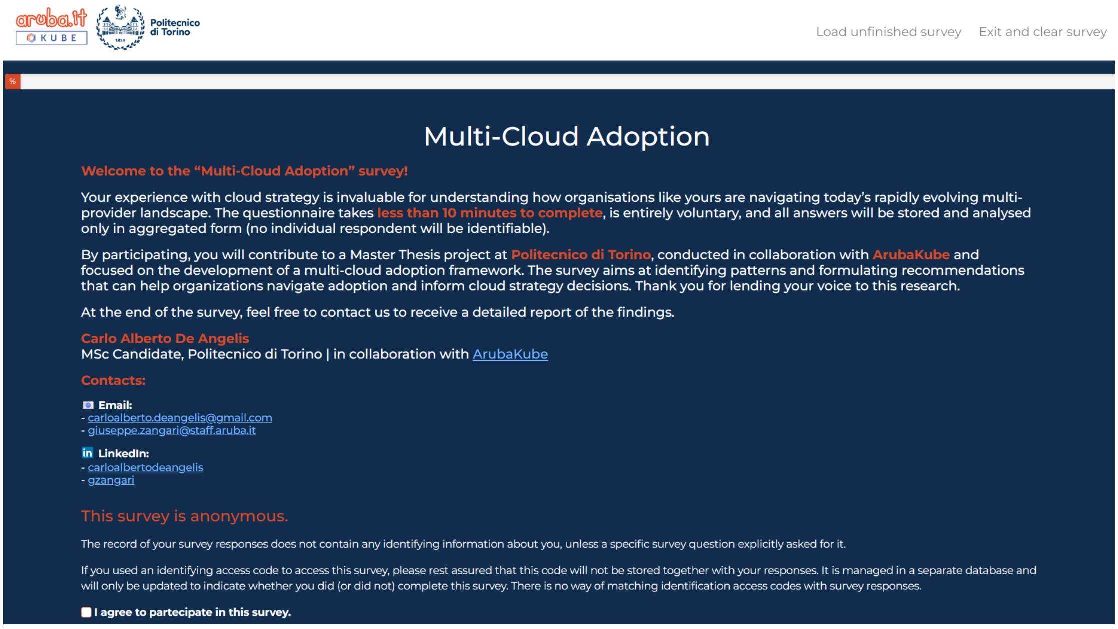


Figure 6.12: Custom PHP survey home page including data privacy disclaimer.

The survey was organized into four sections:

- **Firmographics and Cloud Deployment:** collected organizational characteristics and cloud infrastructure maturity.
- **Multi-Cloud Adoption Patterns:** evaluated through Likert scale items measuring strategic drivers, realized benefits, and organizational challenges. To ensure consistent results, this section was displayed only to respondents using multi-cloud deployment.
- **K8s Usage:** assessed the adoption within organizations of K8s and federation tools, as well as the related operational complexity. The details of this section were displayed only to technical engineers who successfully completed the previous sections.
- **Future Cloud Trends:** explored expectations, trends, and decisions influencing cloud initiatives within a 24-month horizon.

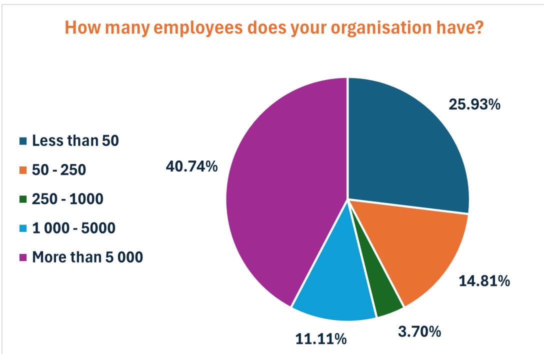
6.4.1 Firmographics and Cloud Deployment

Figure 6.13 shows the industry distribution and respondents' roles, revealing a specialized audience primarily composed of executives and technical engineers. Two dominant clusters emerged: organizations with more than 5,000 employees (40.74%)

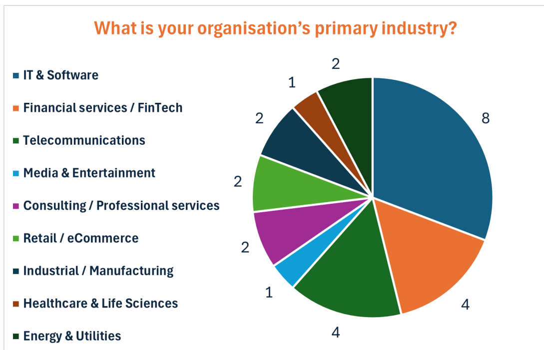
and those with fewer than 50 employees (approximately 26%). As expected, given the focus of the survey, the vast majority of organizations reported operating in a multi-cloud environment. Despite the predominance of multi-cloud deployment among large-scale enterprises, a notable share of small firms (<50 employees) has adopted multi-cloud, suggesting growing accessibility of the model or the presence of cloud-native startups, as highlighted in Table 6.1.

Table 6.1: Primary deployment model by organization size (total n=20).

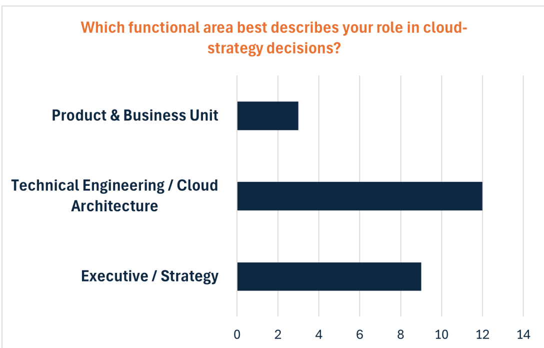
Deployment model	1-50 employees	5,000+ employees
Single public cloud	1	1
Multi-cloud	4	7
Hybrid cloud	0	2



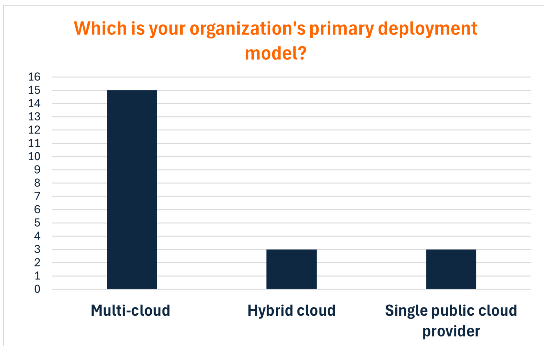
(a) Employee Distribution (n=27).



(b) Industry Sector (n=27).



(c) Professional Role (n=27).



(d) Cloud Deployment Model (n=21).

Figure 6.13: Overview of respondents' characteristics by organization, industry, role, and cloud deployment model.

In the total population, there is a predominance of centralized IT models, where a single IT team manages most systems, tools, and infrastructure, and federated IT models, where a central IT team sets standards or provides platforms and tools, but teams manage their own environments. As highlighted in Table 6.2, the centralized model is typically adopted by small organizations, while larger enterprises tend to prefer federated models, which enable greater agility even in large-scale contexts.

Table 6.2: IT governance model by organization size (total n = 18).

Organization size	Centralized IT	Federated model	Outsourced
Less than 50	1	0	2
50–250	1	1	0
1,000–5,000	0	2	0
More than 5,000	4	4	0

Concerning the satisfaction level, hyperscalers dominate both adoption and customer satisfaction rates (see Figure 6.14). However, a targeted question for European responses (60% of the sample) investigated the strategic relevance of digital sovereignty, in particular regarding the adoption of European providers. The results indicated an average importace score of 4.14 out of 5 based on 14 respondents, signaling a potential growth in the following years.

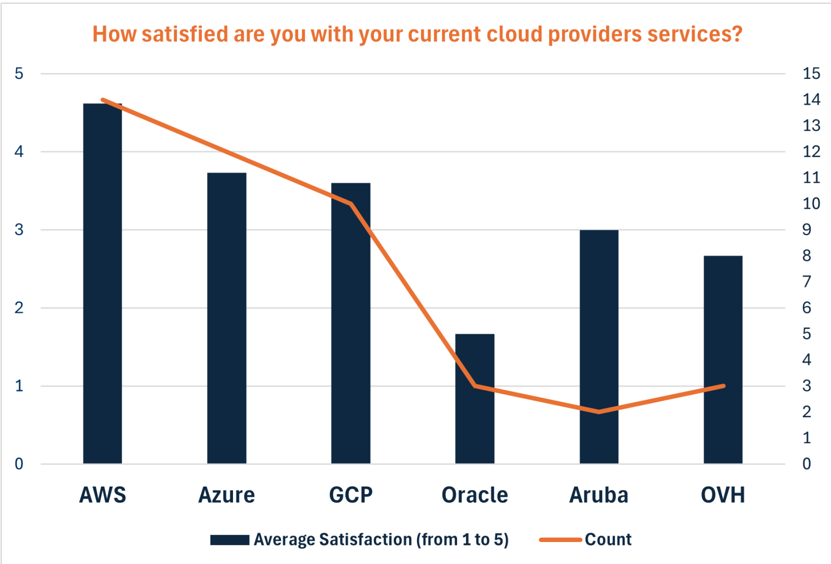
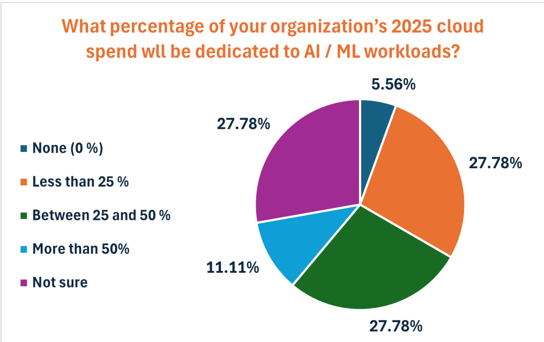
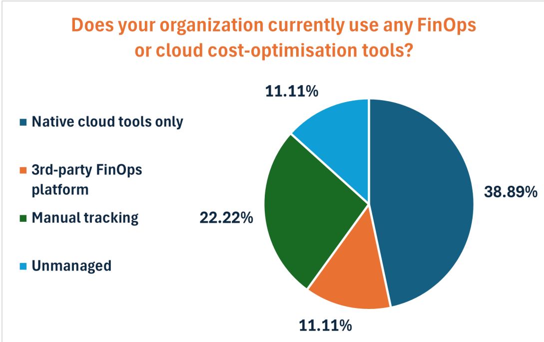


Figure 6.14: Cloud Providers Average Satisfaction and Adoption Rates (n=20).

The financial insights in Figure 6.15 also confirm the current trend of AI/ML workloads among respondents, with over 38% of organizations allocating more than a quarter of their cloud budgets to these activities. Concerning FinOps practices, native cloud tools offered by cloud providers emerged as the dominant model. 22% of organizations still rely on manual tracking (e.g., spreadsheets), while only 11.11% utilized third-party FinOps platforms such as CloudHealth. Surprisingly, a large portion of large-scale enterprises relies on manual methods (33%) and is the only group signaling the use of third-party platforms.



(a) AI/ML workloads budget.



(b) Adoption of FinOps practices and tools.

Figure 6.15: AI/ML workload investment and FinOps practice adoption (n=18).

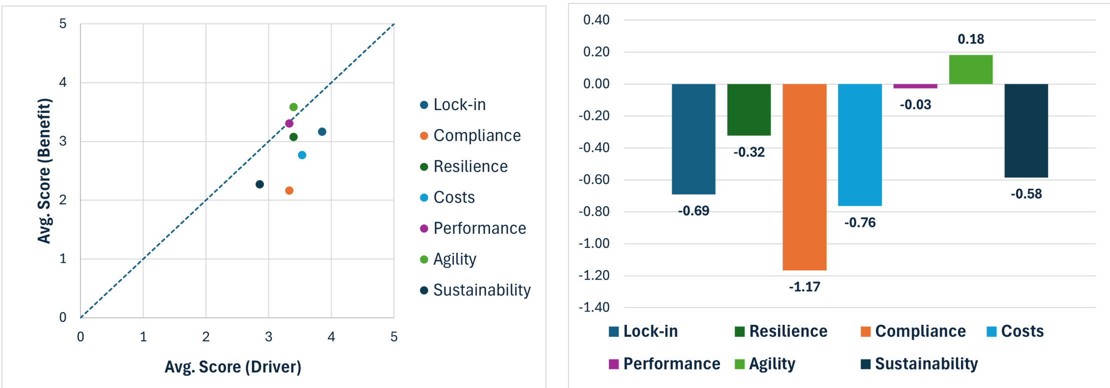
6.4.2 Multi-Cloud Adoption Patterns

This section reports insights from 15 respondents who actively manage multi-cloud deployments, summarized in Table 6.3. Participants evaluated seven key drivers (strategic importance) and benefits (realized extent) on a Likert scale, along with an additional driver concerning access to best-of-breed provider services.

Table 6.3: Average scores of multi-cloud drivers and benefits (n = 15).

Driver/Benefit	Driver Avg.	Benefit Avg.
Lock-in avoidance	3.86	3.17
Resilience	3.40	2.17
Compliance	3.33	3.08
Cost optimization	3.53	2.77
Performance	3.33	3.31
Agility	3.40	3.58
Sustainability	2.86	2.27

Avoiding vendor lock-in emerged as the key strategic driver for multi-cloud adoption, while agility and business unit autonomy were identified as the most tangible realized benefits. The gap analysis in Figure 6.16 compares expected and realized outcomes, signaling agility as the only driver with a positive expectation gap, and service resilience as the largest shortfall between strategic intent and realized value.



(a) Driver importance vs. realized benefits.

(b) Expectation gap by category.

Figure 6.16: Gap analysis comparing strategic intent and realized outcomes across multi-cloud adoption drivers (n=15).

Table 6.4 shows the relationship between service differentiation and AI spending, highlighting how organizations allocating more than 50% of their cloud budgets to AI workloads tend to adopt multi-cloud primarily to access exclusive and best-of-breed provider services. Regarding challenges, interoperability ranked as the most significant barrier to the adoption of multiple providers, followed by switching costs and skills shortage, while the cost observability challenge was particularly severe among organizations relying on manual tracking, as highlighted in table 6.5.

Table 6.4: Average service differentiation driver score, by AI spending (n=12).

AI spending	**Avg. Score**
Less than 25%	2.60
Between 25–50%	2.33
More than 50%	4.50
None (0%)	4.00
Not sure	3.00

Table 6.5: Average cost observability challenge score, by FinOps practice (n=12).

FinOps practice	**Avg. Score**
3rd-party platform	3.00
Manual tracking	4.33
Native tools only	2.50
Not sure	3.00
Unmanaged	2.50

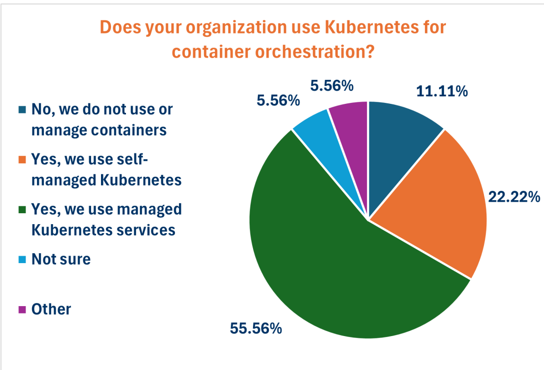
The association between IT management models, agility benefits, and interoperability challenges (see Table 6.6) suggested that fully decentralized teams typically face higher interoperability challenges, whereas centralized and federated models achieve greater agility in managing business unit operations.

Table 6.6: Agility benefits and interoperability challenges, by IT management model (n = 12).

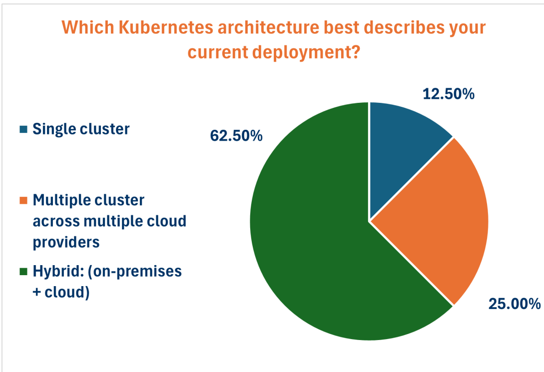
IT management model	Avg. Agility	Avg. Interoperability
Centralized IT department	3.5	3.5
Federated model	3.8	3.8
Fully decentralized teams	2.0	4.0
Outsourced	4.0	3.5

6.4.3 Kubernetes Usage and Future Trends

As illustrated in Figure 6.17, the survey revealed a high adoption rate of K8s among respondents (approximately 90%), with a predominance of managed services (55.56%) and hybrid deployments, where clusters are distributed across a single cloud provider and on-premises infrastructure (62.50%).



(a) Kubernetes adoption rate.



(b) Kubernetes deployment model.

Figure 6.17: Kubernetes usage and deployment models (n=18).

The most common operational challenges were associated with securing K8s clusters and managing upgrades and patching. These issues were particularly pronounced among organizations operating self-managed K8s environments, while users of managed services faced greater difficulties in handling stateful applications and addressing the shortage of skilled staff, as illustrated in Figure 6.18.

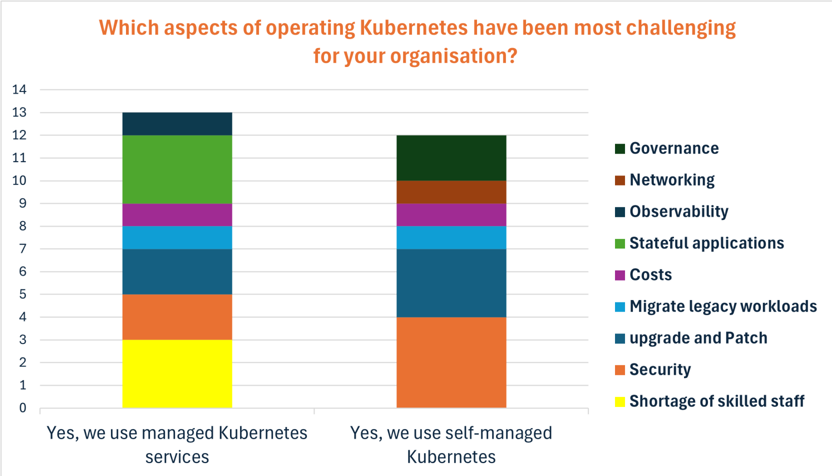


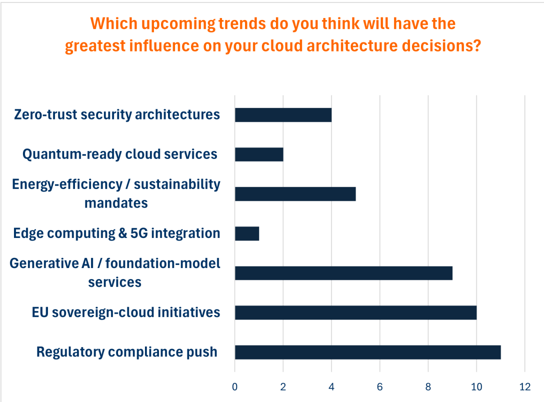
Figure 6.18: Kubernetes management challenges, managed vs self-managed deployments (n = 8).

Additionally, organizations managing multiple clusters across different providers scored higher in terms of interoperability challenges, a trend also observable in organizations relying on self-managed K8s deployments, as shown in Table 6.7.

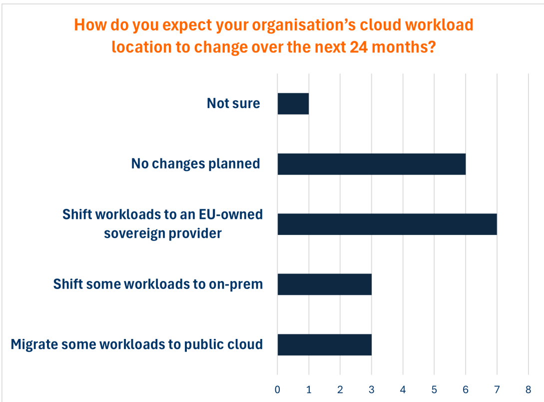
Table 6.7: Interoperability challenge scores by Kubernetes deployment and management model (n=8).

Model	Avg. Score
Kubernetes deployment models	
Hybrid (on-premises + cloud)	3.67
Multiple clusters across multiple providers	4.50
Container orchestration models	
No, we do not use or manage containers	3.00
Yes, we use managed Kubernetes services	3.43
Yes, we use self-managed Kubernetes	4.33

Regarding future cloud strategy decisions (see Figure 6.19), the majority of organizations (62.50%) reported the intention to maintain the current number of cloud providers and their overall spending levels. The results exhibited by future trends influencing cloud strategy decisions were consistent with the discussion of this chapter, highlighting the increasing relevance of European initiatives, regulatory compliance pressures, and generative AI services. Among these, digital sovereignty emerged as the most influential factor expected to shape workload placement over the next 24 months, with most organizations signalling their intention to shift workloads to European providers, emphasizing the growing strategic role of data governance and digital policies in the cloud market landscape.



(a) Upcoming trends expected to influence cloud architectures.



(b) Expected changes in cloud workload placement over 24 months.

Figure 6.19: Strategic cloud trends and expected workload placement decisions among surveyed organizations (n=15).

Chapter 7

Conclusions

This thesis proposed a comprehensive overview of cloud ecosystems from both a technological and strategic perspective. The main objective of the work was to provide guidelines for practitioners through an adoption framework and to highlight the importance and use cases to shift organizations' IT infrastructures toward federated and distributed multi-provider environments. These environments are designed to support emerging technologies and advanced solutions while fostering innovation, reliability, and digital sovereignty.

The work was developed in collaboration with ArubaKube and complements an internship focused on the go-to-market strategy for an enterprise platform based on Liqo and multi-cluster Kubernetes technologies. The platform aims to address common pain points associated with multi-cloud deployments, such as the interoperability across heterogeneous infrastructures, while enabling optimized resource usage and enhancing business unit autonomy through dedicated environments.

Empirical Assessment of the Framework

The empirical insights of the survey are consistent with the proposed multi-dimensional adoption framework, indicating how the strategic, operational, and technical dimensions are translated into concrete business drivers and pain points among surveyed organizations.

At the strategic level, organizations adopting multi-cloud architectures emphasized flexibility as a critical organizational capability, achieved through business unit autonomy and the reduction of vendor-lock. Additionally, the possibility to access exclusive and best-of-breed provider services is particularly relevant for organizations investing in emerging technologies such as AI-based solutions. However, the comparison between strategic intent and realized benefit highlighted that the realized benefits failed to meet some of the key expected outcomes, such as compliance and cost optimization. The most critical operational challenges contributing to this gap were interoperability, security, and skills shortages, as well as the lack of cost observability for organizations relying on manual tracking rather than structured FinOps practices.

At the infrastructure level, Kubernetes confirmed its widespread adoption among respondents, particularly through hybrid and multi-provider deployments. The results also highlighted the importance of choosing between managed and self-managed deployment models, which must align with organizational capabilities to fully exploit the benefit of the implementation.

The survey reinforced the importance of governance as the coordination mechanism across the framework dimensions. Organizations with federated and centralized IT management structures reported higher agility benefits, while fully decentralized teams faced more severe interoperability challenges. In the European landscape, respondents also highlighted the strategic relevance of EU initiatives and sovereignty considerations in shaping their cloud strategies.

Managerial Implications

From a managerial perspective, the thesis offered several implications. First, multi-cloud should not be framed only as a technical architecture but as a strategic decision affecting bargaining power, agility, and innovation capacity. When implemented effectively, it can accelerate development cycles and reduce dependency on a single provider. As technological innovation accelerates, the role of IT becomes increasingly relevant across all industries and organizations. The possibility to leverage providers' resources as infrastructure modules enables organizations to capture greater value from cloud ecosystems by optimizing the use of capacity, retaining control over costs, digital infrastructure, and workloads placement, and choosing the most suitable services or solutions.

The effective implementation of these architectures, particularly Kubernetes multi-cluster environments, strongly relies on the firm's internal knowledge and capabilities. Organizations must balance cloud-agnostic solutions with operational complexity, eventually leveraging managed services based on open-source technologies.

Moreover, European organizations should prioritize digital sovereignty by integrating compliance-by-design principles and sovereign cloud offerings into their multi-cloud strategies.

The proposed modular governance model, combining central IT management structures such as the CCoE with IDPs, can help to avoid fragmentation of knowledge and resources, ensure consistency with central standards, and promote organizational agility.

Limitations and Future Research

The empirical section of the thesis is based on a targeted sample of IT practitioners with a relatively small number of respondents and a focus on organizations which are presumably mature in their level of cloud maturity. For these reasons, the results should be interpreted as descriptive rather than statistically generalizable.

Future research should expand the survey to a larger and more diverse target of respondents across industries and regions. This would allow more robust quantitative analyses, such as clustering techniques or principal component analysis, to test the validity of the framework and categorize organizations based on their size and technological capabilities. Additionally, in-depth interviews and case studies may complement the survey by providing qualitative validation and deeper insights into governance choices and operational practices.

Finally, future studies could explore the role of emerging technologies such as serverless computing or advanced AI techniques to understand the interactions with federated multi-provider ecosystems within adopting organizations, allowing the framework to evolve with ongoing technological developments.

Closing Remarks

Federated multi-provider ecosystems represent a promising direction for organizations seeking to regain control over their digital assets and fully capture the benefits of cloud computing in dynamic and regulated environments. Moreover, the development of these interconnected cloud ecosystems promotes fairer competition and accelerates innovation, reducing the dependency on a single cloud provider. The framework and findings presented in the thesis aim to support practitioners in designing multi-cloud strategies that are resilient, interoperable, and aligned with broader business objectives and sovereignty principles.

Bibliography

[1] Jr. T. Barnett, S. Jain, A. Sumits, U. Andra, and T. Khurana. *Cisco Global Cloud Index: 2015–2020*. Cisco Knowledge Network (CKN) Session. San Jose, CA: Cisco Systems, Inc., Nov. 2016 (cit. on p. 1).

[2] Gartner. *Public cloud services end-user spending worldwide from 2017 to 2025 (in billion U.S. dollars) [Graph]*. [Accessed November 2025, via Polito library]. Statista. Nov. 2024. URL: https://www.statista.com/statistics/273818/global-revenue-generated-with-cloud-computing-since-2009/ (cit. on p. 1).

[3] IDC, Statista and Various sources. *Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2023, with forecasts from 2024 to 2028 (in zettabytes) [Graph]*. [Accessed November 2025, via Polito library]. Statista. May 2024. URL: https://www-statista-com.ezproxy.biblio.polito.it/statistics/871513/worldwide-data-created/ (cit. on p. 2).

[4] Flexera. *Flexera 2025 State of the Cloud Report*. Itasca, IL, USA: Flexera Software LLC, 2025. URL: https://info.flexera.com/CM-REPORT-State-of-the-Cloud-2025-Thanks?revisit (cit. on pp. 2, 61, 63, 76, 86).

[5] IBM. *Cloud migration: Definition, strategies and benefits*. [Accessed November 2025]. IBM Corporation. 2024. URL: https://www.ibm.com/think/topics/cloud-migration (cit. on p. 2).

[6] Statista. *Enterprise spending on cloud and data centers by segment from 2009 to 2024 (in billion U.S. dollars) [Graph]*. [Accessed November 2025, via Polito library]. Statista. Feb. 2025. URL: https://www-statista-com.ezproxy.biblio.polito.it/statistics/1114926/enterprise-spending-cloud-and-data-centers/ (cit. on p. 3).

[7] The Kubernetes Authors. *Kubernetes Documentation*. Accessed November 2025. The Linux Foundation. 2025. URL: https://kubernetes.io/docs/ (cit. on pp. 4, 60).

[8] The Liqo Authors. *Liqo Documentation*. Version 1.0.1. Accessed November 2025. Liqo Project. 2025. URL: https://docs.liqo.io (cit. on pp. 4, 83).

[9] Aruba S.p.A. *Come si costruisce un network cloud?* [Accessed November 2025]. Aruba.it Magazine. July 2025. URL: https://www.aruba.it/magazine/cloud/come-si-costruisce-un-network-cloud.aspx (cit. on pp. 4, 9).

[10] Peter Mell and Timothy Grance. *The NIST Definition of Cloud Computing*. Tech. rep. Special Publication 800-145. Gaithersburg, MD, USA: National Institute of Standards and Technology (NIST), Sept. 2011 (cit. on pp. 6, 11, 90).

[11] Flexera Software. *Enterprise cloud strategy worldwide from 2021 to 2024, by cloud type [Graph]*. [Accessed November 2025, via Polito library]. Statista. June 2024. URL: https://www.statista.com/statistics/817296/worldwide-enterprise-cloud-strategy/ (cit. on p. 10).

[12] BetterCloud, Inc. *State of SaaS 2025*. BetterCloud, Feb. 2025 (cit. on p. 13).

[13] Verified Market Research. *Artificial Intelligence SAAS Market for 2024–2031*. Verified Market Research, Dec. 2024 (cit. on p. 14).

[14] Statista Market Insights. *Public Cloud: Market Data & Analysis*. Statista, Dec. 2024 (cit. on p. 14).

[15] Marty Cagan. *Build vs Buy in the Age of AI*. [Accessed November 2025]. Silicon Valley Product Group (SVPG). Sept. 2024. URL: https://www.svpg.com/article-build-vs-buy-in-the-age-of-ai/ (cit. on p. 15).

[16] Microsoft Corporation. *Azure Pricing Calculator*. Accessed November 2025. Microsoft Corporation. 2025. URL: https://azure.microsoft.com/en-us/pricing/calculator/ (cit. on p. 16).

[17] HashiCorp. *Terraform on GitHub*. Accessed November 2025. HashiCorp, Inc. 2025. URL: https://github.com/hashicorp/terraform (cit. on p. 17).

[18] Michael E. Porter. «What Is Strategy?» In: *Harvard Business Review* 74.6 (1996), pp. 61–78 (cit. on p. 19).

[19] Marco Cantamessa and Francesca Montagna. *Management of Innovation and Product Development*. 2nd. London, UK: Springer, 2023. ISBN: 978-1-4471-7531-5. DOI: https://doi.org/10.1007/978-1-4471-7531-5 (cit. on pp. 19, 31, 56).

[20] Michael E. Porter. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. New York, NY: Free Press, 1980. ISBN: 978-0-684-84148-9 (cit. on pp. 20, 21, 71, 92).

[21] Marty Cagan. *INSPIRED: How to Create Tech Products Customers Love*. 2nd. Hoboken, NJ: Wiley, 2018. ISBN: 978-1119387503 (cit. on pp. 21, 47).

[22] Michael E. Porter. *Competitive Advantage: Creating and Sustaining Superior Performance*. New York, NY: Free Press, 1985. ISBN: 978-0-02-925090-7 (cit. on pp. 21, 22).

[23] Lars de Bruin. *Value Chain Analysis: An Internal Assessment of Competitive Advantage*. Accessed November 2025. Business-to-you.com. Mar. 2018. URL: `https://www.business-to-you.com/value-chain/` (cit. on p. 22).

[24] Joseph A. Schumpeter. *Capitalism, Socialism and Democracy*. New York: Harper and Brothers, 1942. ISBN: 978-0-06-133008-6 (cit. on p. 22).

[25] J. M. Utterback and W. J. Abernathy. «A Dynamic Model of Process and Product Innovation». In: *Omega* 3.6 (1975), pp. 639–656. DOI: `10.1016/0305-0483(75)90068-7` (cit. on p. 23).

[26] W. J. Abernathy and J. M. Utterback. «Patterns of Industrial Innovation». In: *Technology Review* 80.7 (1978), pp. 40–47 (cit. on p. 23).

[27] Roy Rothwell and Walter Zegveld. *Reindustrialization and Technology*. Armonk, NY: M.E. Sharpe, 1985. ISBN: 0873323300 (cit. on p. 24).

[28] Robert Romanowski. «The Nature of Innovation Management». In: *Managing Economic Innovations – Ideas and Institutions*. 2019, pp. 7–38. DOI: `10.12657/9788379862764-1` (cit. on p. 24).

[29] Chunka Mui. *Five Dangerous Lessons to Learn From Steve Jobs*. Forbes. Accessed November 2025. Oct. 2011. URL: `https://www.forbes.com/sites/chunkamui/2011/10/17/five-dangerous-lessons-to-learn-from-steve-jobs/` (cit. on p. 25).

[30] Team Counterpoint. *Apple iPhone Market Share*. Accessed November 2025. Oct. 2025. URL: `https://www.counterpointresearch.com/en/insights/apple-iphone-market-share` (cit. on p. 25).

[31] Lowell W. Steele. *Managing Technology: The Strategic View*. New York: McGraw-Hill, 1989. ISBN: 9780070609365 (cit. on p. 25).

[32] Melvin E. Conway. «How Do Committees Invent?» In: *Datamation* (Apr. 1968) (cit. on p. 25).

[33] Rebecca M. Henderson and Kim B. Clark. «Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms». In: *Administrative Science Quarterly* 35.1 (1990), pp. 9–30. DOI: `10.2307/2393549` (cit. on p. 25).

[34] Everett M. Rogers. *Diffusion of Innovations*. 3rd. New York, NY: Free Press, 1982. ISBN: 978-0-02-926650-2 (cit. on pp. 26, 87).

[35] Clayton M. Christensen. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press, 1997. ISBN: 978-0-87584-585-2 (cit. on pp. 27, 37, 50).

[36] Henry Chesbrough, Wim Vanhaverbeke, and Joel West, eds. *Open Innovation: Researching a New Paradigm*. Accessed November 2025, via ProQuest Ebook Central. Oxford, UK: Oxford University Press, 2006. ISBN: 978-0-19-929072-7 (cit. on p. 27).

[37] Henry William Chesbrough. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Boston, MA: Harvard Business School Press, 2003. ISBN: 978-1578518371 (cit. on p. 28).

[38] Birger Wernerfelt. «A Resource-Based View of the Firm». In: *Strategic Management Journal* 5.2 (1984), pp. 171–180 (cit. on p. 28).

[39] Jay B. Barney. «Firm Resources and Sustained Competitive Advantage». In: *Journal of Management* 17.1 (1991), pp. 99–120. DOI: 10.1177/014920639101700108 (cit. on pp. 28, 29).

[40] Jay B. Barney and William S. Hesterly. «Chapter 3: The VRIO Framework – Internal Analysis». In: *Strategic Management and Competitive Advantage: Concepts*. 6th. Upper Sadd le River, NJ: Pearson Education, 2008 (cit. on p. 29).

[41] Lars de Bruin. *VRIO: From Firm Resources to Competitive Advantage*. Accessed November 2025. Business-to-you.com. Nov. 2016. URL: https://www.business-to-you.com/vrio-from-firm-resources-to-competitive-advantage/ (cit. on p. 29).

[42] C.K. Prahalad and Gary Hamel. «The Core Competence of the Corporation». In: *Harvard Business Review* 68.3 (1990), pp. 79–91 (cit. on p. 29).

[43] Richard R. Nelson and Sidney G. Winter. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Belknap Press of Harvard University Press, 1982. ISBN: 978-0-674-27228-6 (cit. on p. 30).

[44] James G. March. «Exploration and Exploitation in Organizational Learning». In: *Organization Science* 2.1 (1991), pp. 71–87. DOI: 10.1287/orsc.2.1.71 (cit. on pp. 30, 34).

[45] Barbara Levitt and James G. March. «Organizational Learning». In: *Annual Review of Sociology* 14 (1988), pp. 319–340. URL: http://www.jstor.org/stable/2083321 (cit. on p. 30).

[46] Michael L. Tushman and Charles A. O'Reilly III. «Ambidextrous Organizations: Managing Evolutionary and Revolutionary Change». In: *California Management Review* 38.4 (1996), pp. 8–30. DOI: 10.2307/41165852 (cit. on p. 31).

[47] Richard A. D'Aveni. *Hypercompetition: Managing the Dynamics of Strategic Maneuvering*. New York, NY: Free Press, 1994. ISBN: 9781439122631 (cit. on p. 31).

[48] Robert M. Grant. «Toward a Knowledge-Based Theory of the Firm». In: *Strategic Management Journal* 17.S2 (1996), pp. 109–122. DOI: 10.1002/smj.4250171110 (cit. on p. 31).

[49] David J. Teece, Gary Pisano, and Amy Shuen. «Dynamic Capabilities and Strategic Management». In: *Strategic Management Journal* 18.7 (1997), pp. 509–533. DOI: 10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z (cit. on pp. 31, 32).

[50] Annabelle Gawer and Michael A. Cusumano. «Industry Platforms and Ecosystem Innovation». In: *Journal of Product Innovation Management* 31.3 (2014), pp. 417–433. DOI: 10.1111/jpim.12105 (cit. on pp. 32, 37).

[51] Marc H. Meyer and Alvin P. Lehnerd. *The Power of Product Platforms: Building Value and Cost Leadership*. New York, NY: Simon and Schuster, 1997. ISBN: 978-0684825809 (cit. on p. 32).

[52] Carliss Y. Baldwin and Kim B. Clark. *Design Rules, Volume 1: The Power of Modularity*. Cambridge, MA: The MIT Press, 2000. ISBN: 9780262024662. DOI: 10.7551/mitpress/2366.001.0001 (cit. on p. 32).

[53] Constance E. Helfat and Kathleen M. Eisenhardt. «Inter-Temporal Economies of Scope, Organizational Modularity, and the Dynamics of Diversification». In: *Strategic Management Journal* 25.13 (2004), pp. 1217–1232. URL: https://www.jstor.org/stable/20142196 (cit. on p. 33).

[54] Constance E. Helfat and Ruth S. Raubitschek. «Product Sequencing: Co-Evolution of Knowledge, Capabilities and Products». In: *Strategic Management Journal* 21.10-11 (2000). Special Issue: The Evolution of Firm Capabilities, pp. 961–979. URL: https://www.jstor.org/stable/3094422 (cit. on p. 33).

[55] Oliver E. Williamson. «Transaction-Cost Economics: The Governance of Contractual Relations». In: *The Journal of Law & Economics* 22.2 (1979), pp. 233–261. URL: https://www.jstor.org/stable/725118 (cit. on p. 35).

[56] Marshall L. Fisher. «What Is the Right Supply Chain for Your Product? A Simple Framework Can Help You Figure Out the Answer». In: *Harvard Business Review* 75.2 (1997), pp. 105–116 (cit. on p. 35).

[57] Apple Inc. *Apple Supplier List: Fiscal Year 2023*. Tech. rep. Cupertino, CA: Apple Inc., 2024. URL: https://www.apple.com/supplier-responsibility/pdf/Apple-Supplier-List.pdf (cit. on p. 35).

[58] Annabelle Gawer and Michael A. Cusumano. *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Boston, MA: Harvard Business School Press, 2002. ISBN: 9781578515141 (cit. on p. 35).

[59] Jean-Charles Rochet and Jean Tirole. «Platform Competition in Two-Sided Markets». In: *Journal of the European Economic Association* 1.4 (2003), pp. 990–1029. URL: `https://www.jstor.org/stable/40005175` (cit. on pp. 35, 36).

[60] Thomas Eisenmann, Geoffrey Parker, and Marshall W. Van Alstyne. «Platform Envelopment». In: *Strategic Management Journal* 32.12 (2011), pp. 1270–1285. URL: `https://www.jstor.org/stable/41261793` (cit. on p. 36).

[61] Andrei Hagiu and Julian Wright. «Multi-Sided Platforms». In: *International Journal of Industrial Organization* 43 (2015), pp. 162–174. DOI: `10.1016/j.ijindorg.2015.03.003` (cit. on p. 37).

[62] Carmelo Cennamo Michael G. Jacobides and Annabelle Gawer. «Towards a Theory of Ecosystems». In: *Strategic Management Journal* 39.8 (2018), pp. 2255–2276. DOI: `10.1002/smj.2904` (cit. on pp. 37, 38, 52, 53).

[63] Gregory Vial. «Understanding Digital Transformation: A Review and a Research Agenda». In: *The Journal of Strategic Information Systems* 28.2 (2019), pp. 118–144. DOI: `10.1016/j.jsis.2019.01.003` (cit. on pp. 39–41).

[64] Anandhi Bharadwaj, Omar A. El Sawy, Paul A. Pavlou, and N. Venkatraman. «Digital Business Strategy: Toward a Next Generation of Insights». In: *MIS Quarterly* 37.2 (2013), pp. 471–482. URL: `https://www.jstor.org/stable/43825919` (cit. on p. 39).

[65] George Westerman, Didier Bonnet, and Andrew McAfee. *Leading Digital: Turning Technology Into Business Transformation*. Boston, MA: Harvard Business Press, 2014. ISBN: 9781625272478 (cit. on p. 39).

[66] Francesco Paolo Appio, Federico Frattini, Antonio Messeni Petruzzelli, and Paolo Neirotti. «Digital Transformation and Innovation Management: A Synthesis of Existing Research and an Agenda for Future Studies». In: *Journal of Product Innovation Management* 38.1 (2021), pp. 4–20. DOI: `10.1111/jpim.12562` (cit. on p. 40).

[67] IBM Corporation. *Industry 4.0: The Next Industrial Revolution*. Accessed 2025. 2025. URL: `https://www.ibm.com/it-it/think/topics/industry-4-0` (cit. on p. 40).

[68] T. R. Merlo, F. Fard, and S. Hawamdeh. «Cloud Computing's Impact on the Digital Transformation of the Enterprise: A Mixed-Methods Approach». In: *Sustainability* 17.13 (June 2025), p. 5755. DOI: `10.3390/su17135755` (cit. on p. 41).

[69] Aruba S.p.A. *Migrazione al cloud: i criteri essenziali per un'infrastruttura resiliente ed efficiente*. Accessed November 2025. Feb. 2025. URL: https://www.aruba.it/magazine/cloud/migrazione-al-cloud-i-criteri-essenziali-per-un-infrastruttura-resiliente.aspx (cit. on pp. 42, 45).

[70] M. Böhm, S. Leimeister, C. Riedl, and H. Krcmar. «Cloud Computing – Outsourcing 2.0 or a New Business Model for IT Provisioning?» In: *Application Management*. Ed. by F. Keuper, C. Oecking, and A. Degenhardt. Gabler, 2011. ISBN: 978-3-8349-6492-2. DOI: 10.1007/978-3-8349-6492-2_2 (cit. on p. 42).

[71] Jens Dibbern, Tim Goles, Rudy Hirschheim, and Bandula Jayatilaka. «Information systems outsourcing: A survey and analysis of the literature». In: *DATA BASE* 35 (Sept. 2004), pp. 6–102 (cit. on p. 42).

[72] J. B. Quinn. «Strategic Outsourcing: Leveraging Knowledge Capabilities». In: *MIT Sloan Management Review* 40.4 (July 1999), pp. 9–21. URL: https://sloanreview.mit.edu/article/strategic-outsourcing-leveraging-knowledge-capabilities/ (cit. on pp. 42, 43).

[73] A. Benlian, W. J. Kettinger, A. Sunyaev, and T. J. Winkler. «Special Section: The Transformative Value of Cloud Computing: A Decoupling, Platformization, and Recombination Theoretical Framework». In: *Journal of Management Information Systems* 35.3 (July 2018), pp. 719–739. DOI: 10.1080/07421222.2018.1481634 (cit. on p. 43).

[74] Vivek Kundra. *Federal Cloud Computing Strategy*. Office of the Chief Information Officer, The White House. Accessed November 2025. Feb. 2011. URL: https://www.obamawhitehouse.archives.gov/sites/default/files/omb/assets/egov_docs/federal-cloud%E2%80%90computing%E2%80%90strategy.pdf (cit. on p. 43).

[75] European Commission Directorate-General for Digital Services. *European Commission Cloud Strategy*. European Commission. May 2019. URL: https://commission.europa.eu/publications/european-commission-cloud-strategy_en (cit. on p. 43).

[76] Inc. Grand View Research. *Cloud Migration Services Market (2025–2030)*. Tech. rep. Grand View Research, Inc., Mar. 2025. URL: https://www.grandviewresearch.com/industry-analysis/cloud-migration-services-market-report (cit. on p. 44).

[77] Louis G. Tornatzky, Mitchell Fleischer, and Alok K. Chakrabarti. *The Processes of Technological Innovation*. Lexington, MA: Lexington Books, 1990, p. 298. ISBN: 0669203483, 9780669203486 (cit. on p. 44).

[78] K. Katsantonis, P. Mitropoulou, E. Filiopoulou, M. Nikolaidou, and A. Anagnostopoulos. «Cloud Computing and Economic Growth». In: *Proceedings of the 19th Panhellenic Conference on Informatics (PCI 2015)*. Athens, Greece: ACM, Oct. 2015. DOI: 10.1145/2801948.2802000 (cit. on p. 45).

[79] Sebastian Floerecke and Franz Lehner. «Success-Driving Business Model Characteristics of IaaS and PaaS Providers». In: *International Journal on Cloud Computing: Services and Architecture (IJCCSA)* 8.6 (Dec. 2018), pp. 1–22. DOI: 10.5121/ijccsa.2018.8601. URL: https://doi.org/10.5121/ijccsa.2018.8601 (cit. on pp. 46, 50).

[80] Cloudflare, Inc. *How Are Serverless Computing and Platform-as-a-Service Different?* Accessed November 2025. 2024. URL: https://www.cloudflare.com/it-it/learning/serverless/glossary/serverless-vs-paas/ (cit. on p. 46).

[81] Margaret Joy. *How Cloud Computing Companies Created an Oligopoly*. Accessed: 2025-11-08. OnSIP. May 2021. URL: https://www.onsip.com/voip-resources/industry-news-trends/how-cloud-computing-companies-created-an-oligopoly (cit. on p. 47).

[82] Synergy Research Group. *Hyperscale Data Centers Hit the Thousand Mark; Total Capacity is Doubling Every Four Years*. Accessed November 2025. Synergy Research Group. Apr. 2024. URL: https://www.srgresearch.com/articles/hyperscale-data-centers-hit-the-thousand-mark-total-capacity-is-doubling-every-four-years (cit. on p. 47).

[83] Felix Richter. *AWS Stays Ahead as Cloud Market Accelerates*. Accessed November 2025. Statista. Nov. 2025. URL: https://www.statista.com/chart/18819/worldwide-market-share-of-leading-cloud-infrastructure-service-providers/ (cit. on pp. 48, 49).

[84] Matthew Johnston. *How Amazon Makes Money*. Accessed November 2025. Investopedia. Mar. 2025. URL: https://www.investopedia.com/how-amazon-makes-money-4587523 (cit. on p. 49).

[85] Sebastian Floerecke, Franz Lehner, and Sebastian Schweikl. «Cloud Computing Ecosystem Model: Evaluation and Role Clusters». In: *Electronic Markets* 31 (2021), pp. 923–943. DOI: 10.1007/s12525-020-00419-2 (cit. on pp. 51, 53).

[86] Jan Huntgeburth, Michael Blaschke, and Sabrina Karwatzki. «Exploring Value Co-creation in Cloud Ecosystems: A Revelatory Case Study». In: *Proceedings of the European Conference on Information Systems (ECIS 2015)*. Münster, Germany: AIS, May 2015 (cit. on p. 52).

[87] The Linux Foundation. *About the Linux Foundation*. Accessed November 2025. 2025. URL: https://linuxfoundation.org/ (cit. on p. 54).

[88] Javier Luis Cánovas Izquierdo and Jordi Cabot. «The role of foundations in open source projects». In: *Proceedings of the 40th International Conference on Software Engineering: Software Engineering in Society*. ICSE-SEIS '18. Gothenburg, Sweden: Association for Computing Machinery, 2018, pp. 3–12. ISBN: 9781450356619. DOI: 10.1145/3183428.3183438 (cit. on p. 54).

[89] Open Source Initiative. *OSI Approved Licenses*. Accessed November 2025. Opensource.org. 2025. URL: http://opensource.org/licenses (cit. on p. 54).

[90] Moqod. *Understanding Open-Source and Free Software Licensing*. Accessed November 2025. Medium. Sept. 2019. URL: https://moqod-software.medium.com/understanding-open-source-and-free-software-licensing-c0fa600106c9 (cit. on p. 55).

[91] Rowan Trollope. *Redis is Now Available Under the AGPLv3 Open Source License*. Accessed November 2025. Redis Ltd. May 2025. URL: https://redis.io/blog/agplv3/ (cit. on p. 55).

[92] Harsha Dutta and Pankaj Pathak. «Enabling Digital Transformation with Cloud Native Architecture». In: *Proceedings of the First International Conference on Advances in Computational Science and Engineering (ICACSE 2020)*. Vol. 2519. 1. Coimbatore, India: AIP Publishing, Oct. 2022. DOI: 10.1063/5.0109653 (cit. on p. 55).

[93] Giovanni Dosi. «Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change». In: *Research Policy* 11.3 (1982), pp. 147–162. ISSN: 0048-7333. DOI: 10.1016/0048-7333(82)90016-6 (cit. on p. 56).

[94] Christof Ebert, Gorka Gallardo, Josune Hernantes, and Nicolás Serrano. «DevOps». In: *IEEE Software* 33 (May 2016), pp. 94–100. DOI: 10.1109/MS.2016.68 (cit. on p. 56).

[95] Kent Beck et al. *Manifesto for Agile Software Development*. Accessed November 2025. Agile Alliance. Feb. 2001. URL: https://agilemanifesto.org/ (cit. on p. 56).

[96] Amazon Web Services, Inc. *The Difference Between Containers and Virtual Machines*. Accessed November 2025. Amazon Web Services. 2024. URL: https://aws.amazon.com/it/compare/the-difference-between-containers-and-virtual-machines/ (cit. on p. 57).

[97] Saurabh Deochake, Sumit Maheshwari, Ridip De, and Anish Grover. «Comparative Study of Virtual Machines and Containers for DevOps Developers». In: *SSRN Electronic Journal* (Jan. 2023). DOI: 10.2139/ssrn.4404383 (cit. on p. 58).

[98] Stephanie Susnjara. *The History of Kubernetes*. Accessed November 2025. IBM Think. 2024. URL: https://www.ibm.com/think/topics/kubernetes-history (cit. on p. 58).

[99] Cloud Native Computing Foundation (CNCF). *Top 30 Open Source Projects*. Accessed November 2025. The Linux Foundation. 2025. URL: https://insights.linuxfoundation.org/collection/top-open-source-projects (cit. on p. 59).

[100] Valerie Silverthorne and Stephen Hendrick. *Cloud Native 2024: Approaching a Decade of Code, Cloud, and Change*. Accessed: 2025-11-08. Cloud Native Computing Foundation and The Linux Foundation, Mar. 2025. URL: https://www.cncf.io/reports/cncf-annual-survey-2024/ (cit. on pp. 59, 63, 80, 85).

[101] Kubernetes SIG-Security. *Kubernetes Policy Management White Paper*. Tech. rep. Accessed November 2025. Cloud Native Computing Foundation (CNCF), 2021. URL: https://github.com/kubernetes/sig-security/blob/main/sig-security-docs/papers/policy/CNCF_Kubernetes_Policy_Management_WhitePaper_v1.pdf (cit. on p. 61).

[102] KPMG International. *Of the Following Technologies, Which Do You Think Will Be the Most Important in Helping Your Business Achieve Its Short-Term Ambitions (Over the Next 0–3 Years)?* Accessed November 2025. Statista. Sept. 2023. URL: https://www.statista.com/statistics/1446738/important-technologies-business-short-term-ambition/ (cit. on p. 64).

[103] Niklas Kühl, Max Schemmer, Marc Goutier, and Gerhard Satzger. «Artificial intelligence and machine learning». In: *Electronic Markets* 32 (Nov. 2022). DOI: 10.1007/s12525-022-00598-0 (cit. on p. 66).

[104] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. «Attention is all you need». In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NIPS'17. Long Beach, California, USA: Curran Associates Inc., 2017, pp. 6000–6010. ISBN: 9781510860964. DOI: 10.5555/3295222.3295349 (cit. on p. 66).

[105] Aaron Chatterji, Tom Cunningham, David Deming, Zoë Hitzig, Christopher Ong, Carl Shan, and Kevin Wadman. *How People Use ChatGPT*. Tech. rep. Accessed November 2025. OpenAI, Duke University, and Harvard University, Sept. 2025. URL: https://cdn.openai.com/research/how-people-use-chatgpt/how-people-use-chatgpt.pdf (cit. on p. 66).

[106] OpenAI. *Announcing the Stargate Project*. Accessed November 2025. OpenAI. Jan. 2025. URL: https://openai.com/index/announcing-the-stargate-project (cit. on p. 67).

[107] Jordan Novet. *OpenAI Says It Will Use Google Cloud for ChatGPT*. Accessed November 2025. CNBC. July 2025. URL: https://www.cnbc.com/2025/07/16/openai-googles-cloud-chatgpt.html (cit. on p. 67).

[108] Gartner, Inc. *Gartner Hype Cycle Identifies Top AI Innovations in 2025*. Accessed November 2025. Gartner Newsroom. Aug. 2025. URL: https://www.gartner.com/en/newsroom/press-releases/2025-08-05-gartner-hype-cycle-identifies-top-ai-innovations-in-2025 (cit. on p. 67).

[109] Andrej Karpathy. *Software Is Changing (Again)*. Keynote at AI Startup School, San Francisco, June 17, 2025. [Online; accessed 2025-11-08]. Y Combinator. June 2025. URL: https://www.youtube.com/watch?v=LCEmiRjPEtQ (cit. on p. 67).

[110] *What is the Model Context Protocol (MCP)?* Accessed November 2025. Model Context Protocol Project. 2025. URL: https://modelcontextprotocol.io/docs/getting-started/intro (cit. on p. 68).

[111] Aruba S.p.A. *Intelligenza artificiale nel Cloud: le innovazioni che stanno rivoluzionando i servizi*. Accessed November 2025. Aruba Magazine. Aug. 2025. URL: https://www.aruba.it/magazine/cloud/intelligenza-artificiale-nel-cloud-le-innovazioni-che-stanno-rivoluzionando-i-servizi.aspx (cit. on p. 68).

[112] Dinesh Soni and Neetesh Kumar. «Machine learning techniques in emerging cloud computing integrated paradigms: A survey and taxonomy». In: *Journal of Network and Computer Applications* 205 (2022), p. 103419. ISSN: 1084-8045. DOI: https://doi.org/10.1016/j.jnca.2022.103419 (cit. on p. 68).

[113] Nutanix. *Usage and planned usage of IT models worldwide 2024, by type [Graph]*. Statista. Accessed November 2025. Mar. 2024 (cit. on pp. 70, 80).

[114] Google. *Reasons for using a multi-cloud infrastructure 2023 [Graph]*. Statista. Accessed November 2025. Feb. 2023 (cit. on p. 70).

[115] J. Opara-Martins, R. Sahandi, and F. Tian. «Critical Analysis of Vendor Lock-In and Its Impact on Cloud Computing Migration: A Business Perspective». In: *Journal of Cloud Computing* 5.4 (2016). Accessed November 2025. DOI: 10.1186/s13677-016-0054-z (cit. on p. 71).

[116] Aseem Mankotia. «Comparing High Availability and Disaster Recovery in Multi-Cloud Environments». In: *International Journal of Computer Trends and Technology (IJCTT)* 72.7 (2024). Accessed November 2025, pp. 113–121. DOI: 10.14445/22312803/IJCTT-V72I7P115 (cit. on p. 72).

[117] Kentik. *Cloud Latency Map: Latency Changes by City (Torino–Mumbai)*. Kentik Cloud Performance Monitor. Accessed November 2025. Oct. 2025. URL: https://clm.kentik.com (cit. on pp. 73, 74).

[118] *IPCEI-CIS Reference Architecture, Version 1.0*. Tech. rep. Accessed November 2025. 8ra Initiative, May 2025. URL: https://www.8ra.com/wp-content/uploads/IPCEI-CIS_Reference-Architecture_final.pdf (cit. on p. 73).

[119] Kheng-Leong Tan, Chi-Hung Chi, and Kwok-Yan Lam. *Analysis of Digital Sovereignty and Identity: From Digitization to Digitalization*. 2022. arXiv: 2202.10069 [cs.CR]. URL: https://arxiv.org/abs/2202.10069 (cit. on p. 74).

[120] European Union. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation)*. Official Journal of the European Union, L 119, 4 May 2016, pp. 1–88. Accessed November 2025. 2016. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj (cit. on p. 74).

[121] United States Congress. *Clarifying Lawful Overseas Use of Data Act (CLOUD Act), Division V of the Consolidated Appropriations Act, 2018, Pub. L. No. 115–141*. Public Law 115–141, 132 Stat. 1210. Accessed November 2025. Mar. 2018. URL: https://www.justice.gov/d9/pages/attachments/2019/04/09/cloud_act.pdf (cit. on p. 75).

[122] Julie Brill and Erin Chapple. *Microsoft Announces the Phased Rollout of the EU Data Boundary for the Microsoft Cloud Begins January 1, 2023*. Accessed November 2025. Microsoft EU Policy Blog. Dec. 2022. URL: https://blogs.microsoft.com/eupolicy/2022/12/15/eu-data-boundary-cloud-rollout/ (cit. on p. 75).

[123] Andre Bittencourt. *EU Cloud Sovereignty: Four Alternatives to Public Clouds*. Accessed November 2025. Unit8. June 2025. URL: https://unit8.com/resources/eu-cloud-sovereignty-four-alternatives-to-public-clouds/ (cit. on p. 75).

[124] Microsoft Italia and Aruba S.p.A. *Aruba e Microsoft Italia insieme per una nuova offerta Azure Local su infrastrutture dedicate e localizzate in Italia*. Accessed: 2025-11-10. Microsoft News Center Italia. May 2025. URL: https://news.microsoft.com/it-it/2025/05/28/aruba-e-microsoft-italia-insieme-per-una-nuova-offerta-azure-local-su-infrastrutture-dedicate-e-localizzate-in-italia/ (cit. on p. 75).

[125] Stanisław Galij, Grzegorz Pawlak, and Sławomir Grzyb. «Modeling Data Sovereignty in Public Cloud—A Comparison of Existing Solutions». In: *Applied Sciences* 14 (Nov. 2024), p. 10803. DOI: 10.3390/app142310803 (cit. on p. 75).

[126] Stephen Elliot. *Enabling a Sovereign Cloud Using a Multi-Cloud Foundation: Technology Executive Considerations*. Accessed November 2025. Broadcom B-Connected Blog. May 2023. URL: https://www.broadcom.com/blog/enabling-a-sovereign-cloud-using-a-multi-cloud-foundation (cit. on p. 75).

[127] Fernando Almeida, Oliveira José, and Cruz José. «Open Standards And Open Source: Enabling Interoperability». In: *International Journal of Software Engineering & Applications* 2 (Apr. 2011). DOI: 10.5121/ijsea.2011.2101 (cit. on p. 77).

[128] *What Is the OpenAPI Specification?* Accessed November 2025. The Linux Foundation. 2025. URL: https://www.openapis.org/what-is-openapi (cit. on p. 77).

[129] *GraphQL: The Query Language for Modern APIs*. Accessed November 2025. The Linux Foundation. 2025. URL: https://graphql.org (cit. on p. 77).

[130] *gRPC: A High Performance, Open Source Universal RPC Framework*. Accessed November 2025. The Linux Foundation. 2025. URL: https://grpc.io (cit. on p. 77).

[131] *SECA: Sovereign European Cloud API*. Accessed November 2025. Aruba, IONOS, and Dynamo. 2025. URL: https://secapi.cloud (cit. on p. 77).

[132] Sijjad Ali, Dhani Bux Talpur, Adeel Abro, Khulud Salem Alshudukhi, Ghadah Naif Alwakid, Mamoona Humayun, Farhan Bashir, Shuaib Ahmed Wadho, and Asadullah Shah. «Security and privacy in multi-cloud and hybrid cloud environments: Challenges, strategies, and future directions». In: *Computers & Security* 157 (2025), p. 104599. ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2025.104599 (cit. on pp. 77, 78).

[133] Qigui Yao, Qi Wang, Xiaojian Zhang, and Jiaxuan Fei. «Dynamic Access Control and Authorization System based on Zero-trust architecture». In: *Proceedings of the 2020 1st International Conference on Control, Robotics and Intelligent System*. CCRIS '20. Xiamen, China: Association for Computing Machinery, 2021, pp. 123–127. ISBN: 9781450388054. DOI: 10.1145/3437802.3437824 (cit. on p. 78).

[134] Amazon Web Services, Inc. *What is DevSecOps?* Accessed November 2025. Amazon Web Services. 2025. URL: https://aws.amazon.com/it/what-is/devsecops/ (cit. on p. 78).

[135] FinOps Foundation. *FinOps Framework Documentation*. Accessed November 2025. The Linux Foundation. Jan. 2025. URL: https://www.finops.org/ (cit. on p. 78).

[136] FinOps Foundation. *State of FinOps 2025 Report*. Tech. rep. Accessed November 2025. FinOps Foundation, a project of The Linux Foundation, Jan. 2025. URL: https://data.finops.org (cit. on p. 79).

[137] Visak Krishnakumar. *The True Cost of Cloud Data Egress and How to Manage It*. Accessed November 2025. CloudOptimo. July 2025. URL: https://www.cloudoptimo.com/blog/the-true-cost-of-cloud-data-egress-and-how-to-manage-it/ (cit. on p. 79).

[138] European Union. *Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on Harmonised Rules on Fair Access to and Use of Data (Data Act)*. Official Journal of the European Union, L 2023/2854, 22 December 2023. Accessed November 2025. 2023. URL: https://eur-lex.europa.eu/eli/reg/2023/2854/oj (cit. on pp. 79, 94).

[139] Spectro Cloud. *Kubernetes in the AI Era: The Spectro Cloud 2025 State of Production Kubernetes Report*. Tech. rep. Accessed November 2025. Spectro Cloud, in collaboration with Adience Research, 2025. URL: https://www.spectrocloud.com/state-of-kubernetes-2025 (cit. on pp. 80, 86).

[140] Dewan Ahmed and Chris Short. *Spotlight on SIG Multicluster*. Accessed November 2025. Kubernetes Contributors Blog. Feb. 2022. URL: https://www.kubernetes.dev/blog/2022/02/04/sig-multicluster-spotlight-2022/ (cit. on p. 81).

[141] Dennis Heimbigner and Dennis McLeod. «A federated architecture for information management». In: *ACM Trans. Inf. Syst.* 3.3 (July 1985), pp. 253–278. ISSN: 1046-8188. DOI: 10.1145/4229.4233. URL: https://doi.org/10.1145/4229.4233 (cit. on p. 82).

[142] Kubernetes Project. *Kubernetes Cluster Federation (kubefed) [Source Code Repository]*. Archived on April 25, 2023. Accessed November 2025. Github Repository. Apr. 2023. URL: https://github.com/kubernetes-retired/kubefed (cit. on p. 82).

[143] David Eads and Kevin Wang. *Karmada and Open Cluster Management: Two New Approaches to the Multicluster Fleet Management Challenge*. Accessed November 2025. Cloud Native Computing Foundation (CNCF) Blog. Sept. 2022. URL: https://www.cncf.io/blog/2022/09/26/karmada-and-open-cluster-management-two-new-approaches-to-the-multicluster-fleet-management-challenge/ (cit. on p. 82).

[144] Karmada Project. *Karmada Documentation*. Accessed November 2025. Cloud Native Computing Foundation (CNCF). 2025. URL: https://karmada.io/docs/ (cit. on pp. 82, 83).

[145] Marco Iorio, Fulvio Risso, Alex Palesandro, Leonardo Camiciotti, and Antonio Manzalini. «Computing Without Borders: The Way Towards Liquid Computing». In: *IEEE Transactions on Cloud Computing* 11.3 (July 2023), pp. 2820–2838. ISSN: 2372-0018. DOI: 10.1109/tcc.2022.3229163. URL: http://dx.doi.org/10.1109/TCC.2022.3229163 (cit. on p. 83).

[146] Gianluca Arbezzano and Alex Palesandro. *Simplifying Multi-Clusters in Kubernetes*. Accessed November 2025. Cloud Native Computing Foundation (CNCF) Blog. Apr. 2021. URL: https://www.cncf.io/blog/2021/04/12/simplifying-multi-clusters-in-kubernetes/ (cit. on pp. 83, 84).

[147] The Virtual Kubelet authors. *Virtual Kubelet Documentation*. Accessed November 2025. Cloud Native Computing Foundation (CNCF). 2025. URL: https://virtual-kubelet.io/ (cit. on p. 84).

[148] Sergio Moreschini, Fabiano Pecorelli, Xiaozhou Li, Sonia Naz, David Hästbacka, and Davide Taibi. «Cloud Continuum: The Definition». In: *IEEE Access* 10 (2022), pp. 131876–131886. DOI: 10.1109/ACCESS.2022.3229185 (cit. on p. 85).

[149] SkyQuest Technology Consulting Pvt. Ltd. *Kubernetes Market Size, Share, and Growth Analysis: Industry Forecast 2025–2032*. Tech. rep. Accessed November 2025. SkyQuest Technology Consulting Pvt. Ltd., Dec. 2024. URL: https://www.skyquestt.com/report/kubernetes-market (cit. on p. 85).

[150] 6sense Insights Inc. *Kubernetes Market Share and Technology Profile*. Accessed November 2025. 6sense. 2025. URL: https://6sense.com/tech/cluster-management/kubernetes-market-share (cit. on p. 85).

[151] Frank M. Bass. «A New Product Growth for Model Consumer Durables». In: *Management Science* 15.5 (1969), pp. 215–227. ISSN: 00251909, 15265501. URL: http://www.jstor.org/stable/2628128 (visited on 11/09/2025) (cit. on p. 86).

[152] Dave McCrory. *"Data Gravity – in the Clouds"*. Blog post on Data Gravitas. Accessed November 2025. Dec. 2010. URL: https://datagravitas.com/2010/12/07/data-gravity-in-the-clouds/ (cit. on p. 93).

[153] Jacopo Marino, Leonardo Camiciotti, Francesco Cheinasso, Alessandro Olivero, and Fulvio Risso. «Enabling Compute and Data Sovereignty with Infrastructure-Level Data Spaces». In: *Proceedings of the 3rd Eclipse Security, AI, Architecture and Modelling Conference on Cloud to Edge Continuum*. eSAAM '23. Ludwigsburg, Germany: Association for Computing Machinery, 2023, pp. 77–85. ISBN: 9798400708350. DOI: 10.1145/3624486.3624509. URL: https://doi.org/10.1145/3624486.3624509 (cit. on pp. 93, 105).

[154] Federal Ministry for Economic Affairs and Energy (BMWi) and Federal Ministry of Education and Research (BMBF). *Project GAIA-X: A Federated Data Infrastructure as the Cradle of a Vibrant European Ecosystem*. Brochure. Berlin, Germany, Oct. 2019. URL: https://www.bundeswirtschaftsministerium.de/Redaktion/EN/Publikationen/Digitale-Welt/project-gaia-x.pdf?__blob=publicationFile&v=4 (cit. on p. 94).

[155] European Parliament and Council of the European Union. *Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act)*. Official Journal of the European Union, L 277, 27 October 2022, pp. 1–102. Accessed November 2025. 2022. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R2065 (cit. on p. 94).

[156] European Parliament and Council of the European Union. *Regulation (EU) 2022/1925 on Contestable and Fair Markets in the Digital Sector (Digital Markets Act)*. Official Journal of the European Union, L 265, 12 October 2022, pp. 1–66. 2022. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32022R1925 (cit. on p. 94).

[157] European Parliament and Council of the European Union. *Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)*. Official Journal of the European Union, L152, 3 June 2022, pp. 1–44. Accessed November 2025. 2022. URL: https://eur-lex.europa.eu/eli/reg/2022/868/oj (cit. on p. 94).

[158] International Data Spaces Association (IDSA). *International Data Spaces Reference Architecture Model Version 4.0.* Accessed November 2025. International Data Spaces Association. 2022. URL: `https://docs.internationaldataspaces.org/ids-knowledgebase/ids-ram-4` (cit. on pp. 95, 96).

[159] Oscar Corcho and Elena Simperl. *data.europa.eu and the European Common Data Spaces: A Report on Challenges and Opportunities*. Tech. rep. Accessed November 2025. Luxembourg: European Commission, Directorate-General for Communications Networks, Content and Technology (DG CONNECT), Apr. 2022. DOI: `10.2830/91050`. URL: `https://data.europa.eu/sites/default/files/report/data.europa.eu_theRoleofdata.europa.euinthecontextofEuropeancommondataspaces.pdf` (cit. on p. 97).

[160] European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data*. COM(2020) 66 final. Accessed November 2025. Brussels, Belgium, Feb. 2020. URL: `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52020DC0066` (cit. on p. 97).

[161] L. Nagel and D. Lycklama. *Design Principles for Data Spaces – Position Paper*. Tech. rep. Accessed November 2025. OPEN DEI - Aligning Reference Architectures, Open Platforms and Large Scale Pilots in Digitising European Industry, July 2021. DOI: `10.5281/zenodo.5105744`. URL: `https://design-principles-for-data-spaces.org/` (cit. on pp. 97, 98).

[162] Gaia-X European Association for Data and Cloud AISBL. *Gaia-X Architecture Document – 25.05 Release*. Accessed November 2025. Gaia-X European Association for Data and Cloud AISBL. Brussels, Belgium, 2024. URL: `https://gitlab.com/gaia-x/technical-committee/architecture-working-group/architecture-document` (cit. on p. 98).

[163] European Commission. *Commission Staff Working Document on Common European Data Spaces*. SWD(2022) 45 final. Accessed November 2025. Brussels: European Commission, Feb. 2022. URL: `https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces` (cit. on p. 98).

[164] 8ra Cloud-Edge Continuum Consortium. *IPCEI-CIS Reference Architecture*. Tech. rep. Accessed November 2025. Brussels, Belgium: 8ra Cloud-Edge Continuum, May 2025. URL: `https://www.8ra.com/wp-content/uploads/IPCEI-CIS_Reference-Architecture_final.pdf` (cit. on p. 99).

[165] EUCloudEdgeIoT Initiative. *NexusForum Research and Innovation Roadmap*. Tech. rep. Accessed: November 2025. Brussels, Belgium: European Commission, DG CONNECT, Mar. 2025. URL: https://eucloudedgeiot.eu/wp-content/uploads/2025/03/NexusForum_R-I_Roadmap-B.pdf (cit. on p. 100).

[166] Fulvio Risso and Giuseppe Zangari. «GPU Offloading in MLOps: Navigating the Multicloud Ecosystem for Flexible AI/ML Deployments». In: *Proceedings of the GARR Conference 2024*. Accessed: November 2025. 2024. DOI: 10.26314/GARR-Conf24-proceedings-13. URL: https://doi.org/10.26314/GARR-Conf24-proceedings-13 (cit. on p. 104).

[167] MYRTUS Project. *MYRTUS Reference Infrastructure*. Funded by the European Union under grant agreement No. 101135183. Accessed: November 2025. Abinsula srl. 2025. URL: https://myrtus-project.eu/myrtus-infrastructure/ (cit. on pp. 105, 106).