H-CLOUD

CLOUD COMPUTING IN EUROPE

Landscape Analysis, Adoption Challenges and Future Research and Innovation Opportunities

April 2022

h-cloud.eu

Title	Cloud Computing in Europe: Landscape Analysis, Adoption Challenges and Future Research and Innovation Opportunities
Lead Editor	Mark Dietrich and Federico M. Facca
Contributors *(in alphabetical order)*	Carla Arend, Monique Calisti, Gabriella Cattaneo, Massimiliano Claps, Enol Fernández, Tiziana Ferrari, Angele Giuliano, Peter Meadley, Steve Robertshaw
External contributors *(in alphabetical order)*	Jurry de la Mar (T-Systems), Alfonso De La Fuente Ruiz (Prodevelop), Martin Dias d'Ullois (Logius), Ana Juan Ferrer (ATOS), Alain Fiocco (OVHcloud), Geir Horn (University of Oslo), Mark Kuehner (SAP), Geert Machtelinckx (ThreeFold Tech), Ludek Matyska (Masaryk University), Alessandro Perilli (RedHat), Dana Petcu (Western University of Timisoara), Lutz Schubert (University of Ulm), Alban Schmutz (CISPE), Rainer Sträter (IONOS), Andreas Weiss (GAIA-X), Klaas Wierenga (GEANT)
Version	2.0
License	
Date	4th April 2022

ABSTRACT

In the last 2 years the European Commission has worked actively to define priorities and set targets for the new Multiannual Financial Framework (MFF) of the European Union covering the 2021-2027 period and beyond. Among the different identified priorities, "A Europe fit for the digital age" explicitly supports digitalisation. Cloud computing, as a fundamental key of a digital Europe, will play an even stronger role in European economy and society by embracing core European values, spanning fundamental individual rights to market openness and environmental friendliness. Recent events, such as the corona crisis, underscore the essential role of data accessibility and interoperability and digitalisation (processing and analysis) in responding to a growing range of societal challenges, including public health, citizen rights and climate change.
To tackle Europe digitalisation priorities, the European Commission defined "A European strategy for data" and "A new Industrial Strategy for Europe", including a strong focus on data spaces and federated cloud infrastructures and the need for cloud and IT offerings for EU organisations that are both competitive and respectful of European values.
To support the definition of priorities for the upcoming programmes, this report explores the question of how supply and demand can be improved in terms of quality and conditions of use, and increased to boost European innovation in the following cloud computing areas:

1. Effective cloud federation models to stimulate the creation of a European public cloud service market leveraging existing capacities and to enable interoperation of existing data and services;
2. Edge computing to efficiently and sustainably support digital transformation where the data and the users are today and to enable innovative services built on speed, accessibility and responsiveness, while preserving respect for European values;
3. Applying green computing principles to the whole lifecycle of cloud and edge computing to support the transition toward a carbon-neutral (if not carbon-negative) digital society by 2050.

In response to this, the paper summarises an analysis conducted by the H-CLOUD project aiming at identifying the status, challenges, and opportunities that Europe is facing with regards to the adoption and provision of cloud computing with a specific focus on federated cloud, edge computing, and green computing.
The paper explores key challenges and opportunities from the perspective of demand in six key sectors: public administration, healthcare, transport, energy, agriculture and

manufacturing[1]. In addition to these, the paper focuses on the needs of small- and medium-sized enterprises (SMEs). These seven perspectives are referred to as “demand scenarios”.
The demand side analysis is complemented by assessments of the cloud computing supply chain, addressing technical trends, market segmentation and dynamics, and specific examinations of the challenges and opportunities in edge computing, green ICT, and the benefits of federation for cloud infrastructure, edge infrastructure, and data ecosystems.
Building on this analysis, we have positioned our findings in the context of the full value chain supporting the EC’s ambition of creating the Digital Single Market and a Europe fit for the Digital Age, and mapped our preliminary recommendations and conclusions onto this value chain.
The current report (v2.0) has been developed by the H-CLOUD project team and reflects discussions with and feedback from a wide range of experts, and for public consultation in February and March of 2022. This version of the report reflects this community feedback.
Ultimately this will help the EC frame their future funding programmes and help European stakeholders to coordinate key actions to achieve common strategic goals contributing to European competitiveness and ability to innovate in cloud computing.

[1] These sectors are among the key sectors covered in “A European strategy for data”.

EXECUTIVE SUMMARY

Cloud computing, as well as the related technologies addressed by the H-CLOUD project, are essential tools that will enable the digital transformation of, and data-driven innovation by, the EU economy and its roughly 24 million enterprises. This report uses the detailed insights generated by the H-CLOUD project to make recommendations about how these technologies can be used most effectively to achieve these objectives. Since these technologies and tools are necessary but not sufficient for the digital transformation, we also address broader related challenges and opportunities related.

The EU policy context makes it clear that the desired outcome is digital transformation and data-driven innovation, retaining both data and digital sovereignty, while respecting the environment. The *High Impact Project on European data spaces and federated cloud infrastructures,* presented in "A European strategy for data", aims at supporting such data-driven innovation, while keeping carbon impacts under control, by simplifying green cloud services adoption. In addition, a group of European industrial players[2] gathered by Commissioner for Internal Market Thierry Breton has been tasked in December 2020 to deliver a "*European industrial technology roadmap for the next generation cloud-edge offering*" which was released publicly in May 2021[3]. This initiative supports the joint declaration of the 27 EU Member States in October 2020 towards a next generation cloud in Europe[4]. The forthcoming IPCEI for Cloud and Edge developed by several EU Member States with the support of the European Commission may be a key driver of the delivery of such a roadmap, complementing traditional funding schemes by the European Commission.

A recent study by Deloitte[5] highlights how digitalisation is a key opportunity for Europe: the study estimates that if all EU Members States were to reach a Digital Economy and Society Index[6] score of 90 by 2027, this would translate into an increase of the GDP per capita across the EU27 of 7.2% compared to an 'as is' scenario. Based on an EU27 GDP of 16 trillion Euros, this represents an increase of roughly 1.2 trillion Euros. The European Commission's 2012 strategy "Unleashing the Potential of Cloud Computing in Europe"[7] estimated that the proposed actions would add 160 billion Euros to GDP each year.

Despite the support and consideration provided by the EU's policy context, for many organisations, "cloud adoption" is neither simple nor a "one size fits all" process. It is often complex, requiring detailed planning, skilful execution, and careful consideration of return on investment. "Data-driven innovation" is even more difficult for many organisations, and the right conditions and support will be necessary to encourage and enable this essential component of Europe's future. Combining "digital" and "green" makes the context even more complex, especially for small- and medium-sized enterprises (SMEs) – which include many public sector entities (such as health care and education providers, and small municipalities) with important roles to play in the EU's digital journey.

The European Commission tasked the H-CLOUD project with analysing the *status quo* and provide recommendations for future work programmes, developed in conjunction with the European Cloud Community. As stated in the EUSD, "the digital transformation of the EU economy depends on the availability and uptake of secure, energy-efficient, affordable and high-quality data processing capacities, such as those offered by cloud infrastructures and services, both in data centres and at the edge". Consistent with this premise, H-CLOUD focuses on edge computing, cloud federation, and green computing, their role and relevance

[2] Airbus, Amadeus, Aruba, Atos, Cap Gemini, CloudFerro, DE-CIX, Ericsson, Gigas, German Edge Cloud, Indra, IONOS, Irideos, Leaseweb, Magic Cloud, Nabiax, Nokia, Orange, 3DS Outscale, OVHcloud, Retelit, SAP, Schneider Electric, Siemens, Deutsche Telekom, Telefonica, TIM

[3] AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. May 2021.

[4] AA.VV. Building the next generation cloud for businesses and the public sector in the EU. 2020

[5] Deloitte. Digitalisation: an opportunity for Europe. 2021

[6] The Digital Economy and Society Index (DESI) summarises indicators on Europe's digital performance and tracks the progress of EU countries.

[7] EC. Unleashing the Potential of Cloud Computing in Europe. 2012.

in different “use cases”, and recommendations to overcome adoption barriers for different key stakeholders.

This report presents a comprehensive analysis of the roles of cloud, edge, green and federation (as proposed by the EC) in the value chain supporting European digital transformation, and recommendations for relevant stakeholders to meet Europe’s aspirations. It builds on the following analysis components:

- Supply-side analysis: In-depth assessment of the current state of the art in technology and related cloud and edge market dynamics in the EU.
- Demand-side analysis: H-CLOUD’s analysis of deployment challenges faced by representative “demand side” sectors of the economy identified numerous barriers and challenges to adoption. While not every sector was analysed, these analyses collectively exposed many challenges that apply across multiple sectors.
- Analysis of Research and Deployment Landscape. H-CLOUD analysed the research and deployment landscape in European cloud computing, including specific deployment initiatives involving federated cloud, edge computing and green ICT, and identifying best practices that might be adopted more widely.
- Community Engagement and consultation. H-CLOUD engaged extensively with over 500 research and industry experts, in fourteen consultations over a period of two years, to validate the methodology, build consensus for conclusions, and prioritize related recommendations.

This analysis has resulted in a “strategy map,” a conceptual tool that helps separate the many interconnected trends and developments in the European cloud market. Figure 1 illustrates the components of H-CLOUD’s strategy map for European cloud computing.

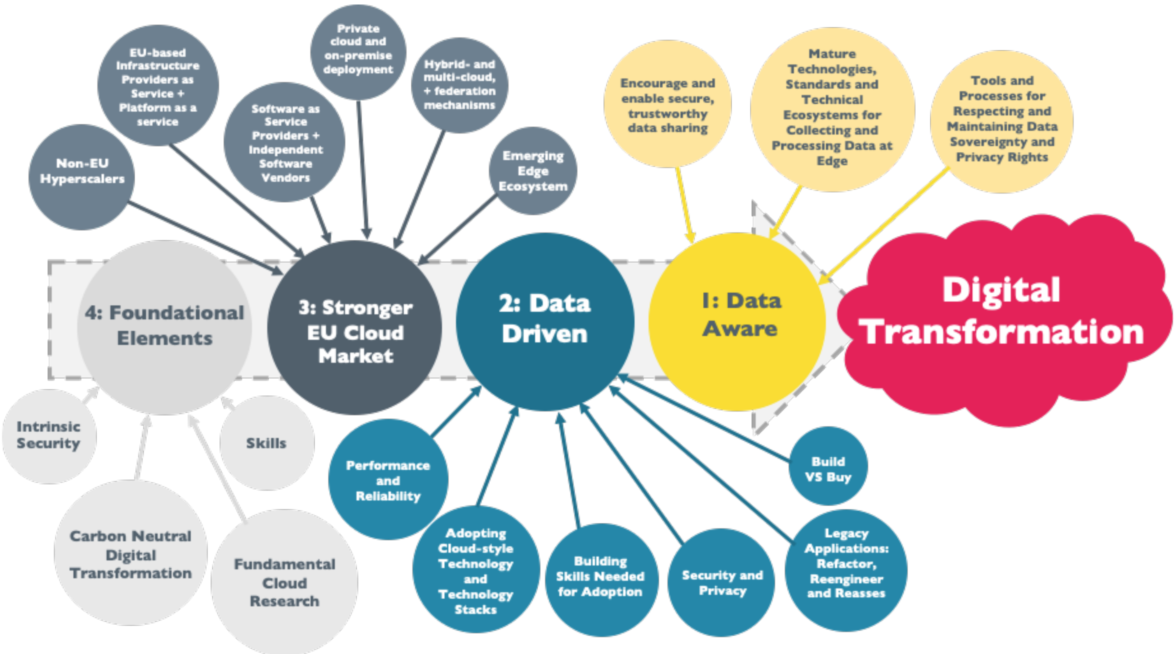


Figure 1: H-CLOUD’s European Cloud Computing Strategy Map

The H-CLOUD Strategy Map is built around the following connected areas, or “pillars”:

Pillar 1: Data Aware EU organisations. There are several ingredients to organisations becoming more “data aware”:

- *Functional, sovereign, affordable, environmentally friendly IT services at the edge*. Edge technologies have the potential to become as important as cloud computing for digital transformation, since they tap into the large volumes of operational, and strategically important, data that is generated at the edge and should remain there.

- *Well-governed business and data ecosystems;* enabling secure, protected data exchange and sharing that respects the rights of data subjects and data owners. A growing number of EU policies highlight the need for organisations to work together, rather than individually, to achieve critical objectives around climate neutrality and social well-being.
- *Accessible, understandable, ethical tools, capabilities, and services to extract knowledge and insights from large amounts of data (big data analytics, artificial intelligence).* This topic is extensively addressed by other initiatives, and outside the scope of H-CLOUD, but has been correctly identified as a strategic enabler for digital transformation.

Pillar 2: Data Driven Innovation, Enabled by New IT Paradigms. Supporting the smooth digitalization organisations need naturally depends on their adopting flexible cloud-style IT paradigms. However, "cloud-style IT" is not the same as "cloud computing" *per se* – it addresses not only technical but also organisational approaches. It is more than just "using the public cloud" – many organisations are making informed decisions to keep their IT infrastructure on-premises, but they are adopting the principles of cloud-style development and operations to enable their own agile and adaptable, yet secure and sovereign, digital futures.

Pillar 3: Strengthening the EU Cloud Market. European organisations need to be able to assemble the right combinations of IT services, with the right attributes and prices, to allow them to meet their strategic objectives. They need clearly defined options from their IT suppliers:

- *EU IT suppliers,* including EU cloud service providers (CSPs), must be encouraged to provide the transparency, flexibility, sovereignty, and sustainability that EU customers need to meet their own objectives and honour their own obligations to society.
- *Non-EU suppliers* must be held to the same standards applied to EU suppliers. The principles of data sovereignty and digital sovereignty should be incorporated in legislation and enforced and tracked with mechanisms such as the ENISA EU Cybersecurity Scheme.
- *As the edge grows in strategic importance*, these quality attributes must also be built from the ground up into the architecture of the edge. Given the diversity and distribution of assets at the edge, and their role handling growing volumes of data, establishing intrinsic, foundational trust at the edge becomes critical.

Pillar 4: Foundational Elements, including Cybersecurity, Environmental Sustainability, Skilled People, and Research and Innovation. Fundamental capabilities are needed in several areas:

- *Carbon Neutral Digital Transformation.* While the EU specifies "energy-efficiency" throughout its policies, the full "stack" of information technologies – from the data centre or hosting site all the way to running software and user-facing services – must be constructed and operated within specific limits on environmental impacts.
- *Intrinsic Security.* Security cannot be added at the last minute – it must be built in from the ground up. This topic is extensively addressed by other initiatives, and outside the scope of H-CLOUD, but will be an essential enabler for digital transformation.
- *Skills.* H-CLOUD identifies several specific skill gaps that must be closed to enable a seamless value chain supporting digital transformation. These gaps must be addressed in the context of the EU's broader workforce challenges, although they are outside the scope of H-CLOUD.
- *Research.* The EU's cloud computing ecosystem must benefit from and exploit the latest research in cloud computing and related digital technologies.

While each pillar is strongly linked with the others, the value of the H-CLOUD Strategy Map is that key issues are separated as much as possible, allowing independent consideration of related challenges and opportunities, and independent development of recommendations, performance measures and implementation plans.

Consultation with the European Cloud Community confirmed that the Strategy Map correctly captures the structure and dynamics of the European cloud sector and its relationship with organisations engaged in their own journeys of digital transformation. H-CLOUD engaged with the community to evaluate the different components of the Strategy Map and develop a consensus prioritization. Table 1 below presents the strategic components, ranked in descending order of community priority. The eight top-ranked strategic issues balance concerns felt by organizations themselves in their digital transformation journey against the ability of Europe's IT market to meet their needs.

Over the course of the project, H-CLOUD has assembled a broad palette of 291 recommendations, spanning every component of the strategy map. 144 recommendations align with the top eight components of the consensus Strategy Map. Another 85 are linked specifically to "Climate Neutral Digital Transformation", although this received comparatively lower ranking. The remaining recommendations (62) are spread across the remaining ten components of the strategy map. Table 1 below details this distribution.

These recommendations can be explored in detail at: https://www.h-cloud.eu/recommendations/. Here we enumerate recommendations in the strategic areas described above.

Strategy Map Component	9-point rating	Total H-CLOUD Recommendations
2.1 Adopt Cloud-Style Technology and Technology Stacks	9	14
2.2 Availability of Cloud Skills	6	10
2.3 Performance and Reliability	6	6
2.4 Security and Privacy	6	9
3.5 Hybrid and multi-cloud, federated cloud	6	38
3.6 Emerging Edge Ecosystem	6	18
3.2 EU-headquartered IaaS and PaaS providers	4	33
3.3 SaaS providers, Independent Software Vendors	4	12
2.6 Legacy Applications	3	4
4.3 Skills	3	1
1.1 Sovereignty and Privacy	2	6
1.2 Data from the Edge	2	14
1.3 Data Sharing and Data Ecosystems	2	16
3.4 Private cloud and on-premises deployment	2	0
4.1 Climate Neutral Digital Transformation	2	85
4.2 Intrinsic Security	2	6
2.5 Management Processes	0	13
3.1 Non-EU-headquartered hyperscalers	0	0

Strategy Map Component	9-point rating	Total H-CLOUD Recommendations
4.4 Foundational Research	0	1

Table 1: Combined Component Ranking, In Descending Order, Fall 2021. Top 8 Areas in bold.

Based on the community's ranking of strategic concerns, H-CLOUD has developed a suite of performance indicators that can be used to track the progress of the European Cloud Community, and of Europe more generally, in addressing these priority challenges. These indicators build on the EC's 2030 Digital Compass indicators, enhancing some existing indicators and offering additional, complementary indicators that might allow for more effective performance measurement and management. The complete set of indicators is presented in Table 2.

<table>
<tr><th>Strategy Map Component</th><th>Revised Digital Compass Targets</th></tr>
<tr><td>2.1 Adopt Cloud-Style Technology and Technology Stacks</td><td>(No Change)
• 75% of European enterprises have taken up cloud computing services, big data and Artificial Intelligence
• More than 90% of European SMEs reach at least a basic level of digital intensity</td></tr>
<tr><td>2.2 Availability of Cloud Skills</td><td>• 20 million employed ICT specialists in the EU
Add:
• Expand cloud-relevant skills and workers specifically.
• Target roughly 50% (10 million) serving the needs of SMEs.</td></tr>
<tr><td>2.3 Performance and Reliability</td><td rowspan="2">New: These strategic challenges are complex and different for every enterprise. We need to empower each enterprise to make informed choices to solve these problems for themselves, making training and support available to any EU enterprise struggling with these challenges.
• Training and consulting supports (on these strategic topics) are available to 100% of EU enterprises – through an estimated 10,000 delivery points.
• 100% of EU enterprises have access to the services and capability of one or more Security Operations Centres.</td></tr>
<tr><td>2.4 Security and Privacy</td></tr>
<tr><td>3.5 Hybrid and multi-cloud, federated cloud</td><td>Revised and Expanded:
• 100% of Europe's large scale data centres, at least 10,000 edge nodes and advanced communications capabilities are</td></tr>
</table>

Strategy Map Component	Revised *Digital Compass* Targets
3.6 Emerging Edge Ecosystem	interconnected, secure, sovereign and climate neutral. • Infrastructure is 100% interoperable, adaptable, and optimized to maximize service quality while minimizing cost and climate impacts. • Data and services are 100% portable, giving organizations flexibility, balanced contractual arrangements, and effective service levels. • 100% of key public services, medical record access, and digital ID services are interoperable, including across Member State borders.
3.2 EU-headquartered IaaS and PaaS providers	**New:** • EU-headquartered firms hold 70% of the IaaS and PaaS market segments in the EU.
3.3 SaaS providers, Independent Software Vendors	**New:** • By 2030, at least 30 EU-headquartered SaaS/software vendors are delivering cloud-native and mobile-native “killer apps” to SMEs in all sectors, in all languages and in all Member States. • At least 10 EU-headquartered SaaS/ software vendors are counted among the innovative scale-ups and unicorns fostered by 2030.

Table 2: Revised Performance Indicators Aligned with Priority Strategy Components

Consistent with the approach described in the 2030 Digital Compass, H-CLOUD proposes that these specific indicators should be incorporated into the planned digital policy programme and included in the monitoring and reporting structure already under consideration.

CONTENTS

1. INTRODUCTION

Cloud computing is a megatrend that is a key enabler for data-driven innovation. It is expected to bring enormous benefits for citizens as stated in the recent EC Communication on Shaping Europe's digital future[8]. It is acknowledged that coordinated efforts are necessary at the European level to make sure that innovation can ultimately make a difference to industry, public administration and eventually society at large.

In the last two years the European Commission has worked actively to define priorities and set targets for the new Multiannual Financial Framework (MFF)[9] of the European Union covering the 2021-2027 period. Among the different priorities, digitalisation is driven by the "A Europe fit for the digital age" priority. In early 2020, EC started to release a set of strategies to support the EU digitalisation, including: "A European strategy for data" (EUSD)[10] and "A New Industrial Strategy for Europe"[11] (NISE). The importance of digitalisation has been further highlighted by the COVID-19 pandemic, resulting in the € 800 billion NextGenerationEU plan[12] and in a "compass" to measure EU progress toward 2030 goals[13]. In these strategies, cloud computing, as a fundamental ingredient of a digital and green Europe, is expected to play an even stronger role in European economy and society by embracing core European values, spanning fundamental individual rights to market openness and to environmental friendliness. Nevertheless, for many organisations, "cloud adoption" is neither simple nor a "one size fits all" process. It is often complex, requiring detailed planning, skilful execution and careful consideration of return on investment. "Data-driven innovation" is even more difficult for many organisations, and the right conditions and support will be necessary to encourage and enable this essential component of Europe's future. Combining "digital" and "green" makes the context even more complex, especially for small- and medium-sized enterprises (SMEs). The *High Impact Project on European data spaces and federated cloud infrastructures,* presented in EUSD, aims at supporting such data-driven innovation, while keeping carbon impacts under control, by simplifying green cloud services adoption. In addition, a group of European industrial players[14] gathered by Commissioner for Internal Market Thierry Breton has been tasked in December 2020 to deliver a "*European industrial technology roadmap for the next generation cloud-edge offering*" which has been released publicly in May 2021[15]. This initiative supports the joint declaration of the 27 EU Member States in October 2020 towards a next generation cloud in Europe[16]. The forthcoming IPCEI for Cloud and Edge developed by several EU Member States with the support of the European Commission may be a key driver of the delivery of such a roadmap, complementing traditional funding schemes by the European Commission.

The European Commission tasked the H-CLOUD project to analyse the *status quo* and provide recommendations for future work programmes with the support of the European Cloud Community. As stated in the EUSD, "the digital transformation of the EU economy depends on the availability and uptake of secure, energy-efficient, affordable and high-quality data processing capacities, such as those offered by cloud infrastructures and services, both in data centres and at the edge". Consistent with this premise, H-CLOUD focuses on edge computing, cloud federation, and green computing, their role and relevance in different "use cases" and the barriers to adoption for different key stakeholders.

This report presents a comprehensive analysis of the roles of cloud, edge, green and

[8] EC. Communication: Shaping Europe's digital future. 2020
[9] EC. Multiannual Financial Framework. 2020.
[10] EC. A European strategy for data. 2020
[11] EC. A new Industrial Strategy for Europe. 2020. Updated in: EC. Updating the 2020 New Industrial Strategy: Building a stronger Single Market for Europe's recovery. 2021
[12] EC. NextGenerationEu. 2020
[13] EC. 2030 Digital Compass: the European way for the Digital Decade. 2021
[14] Airbus, Amadeus, Aruba, Atos, Cap Gemini, CloudFerro, DE-CIX, Ericsson, Gigas, German Edge Cloud, Indra, IONOS, Irideos, Leaseweb, Magic Cloud, Nabiax, Nokia, Orange, 3DS Outscale, OVHcloud, Retelit, SAP, Schneider Electric, Siemens, Deutsche Telekom, Telefonica, TIM
[15] AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. May 2021.
[16] AA.VV. Building the next generation cloud for businesses and the public sector in the EU. 2020

federation (as proposed by the EC) in the value chain supporting European digital transformation, and recommendations for relevant stakeholders to meet Europe's aspirations. It builds on

- an in-depth assessment of the current state of the art in technology and related cloud and edge market dynamics in the EU,
- the needs, challenges and opportunities faced by several key demand sectors in a number of domains prioritised in EU digital strategies,
- a detailed examination of research and deployment initiatives in European cloud computing, and
- extensive feedback from research and industry experts to validate the methodology and the results.

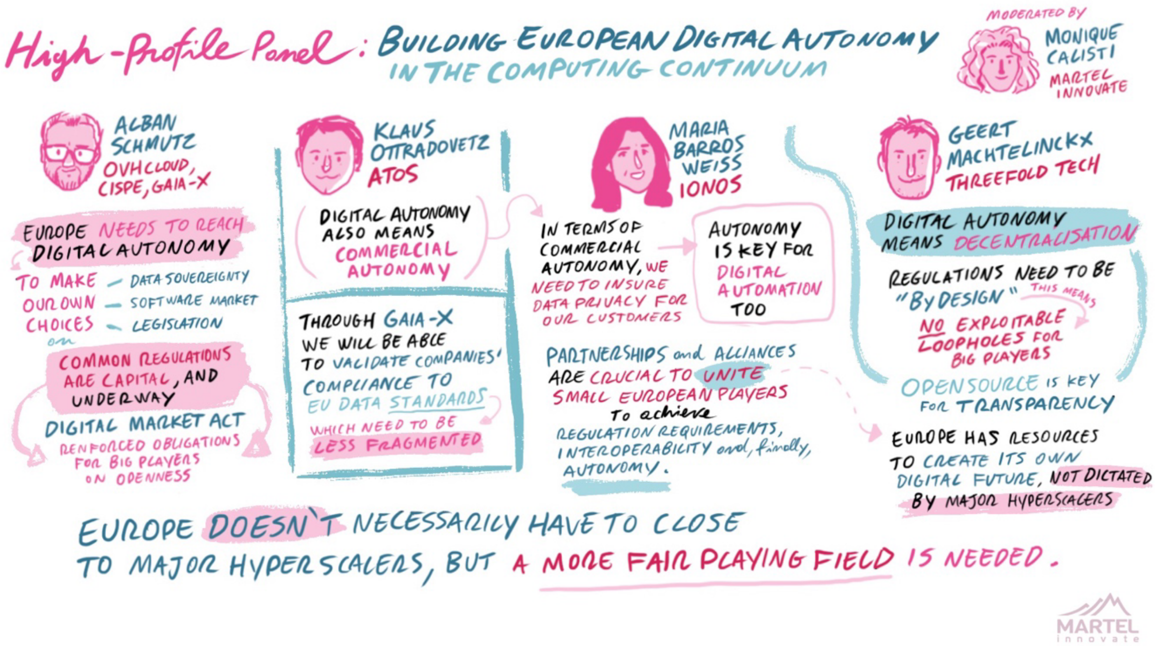


Figure 2. Key outcomes of the "Building European Digital Autonomy in the computing continuum" panel at HORIZON CLOUD SUMMIT 2021 (December).

Finally, this analysis is positioned within the evolving context of EU funding and policy announcements, where numerous new initiatives are translating the EU's aspirations into concrete action. H-CLOUD's analysis has been comprehensive, and this document consolidates this extensive material into common challenges and recommendations to stimulate further discussion with the broader stakeholder community (researchers, SMEs, cloud providers, etc). Following the public consultation, a final report will be released with our final conclusions and recommendations.

1.1. The Horizon Cloud Project

The Horizon Cloud project ("H-CLOUD") was established to consider the future of cloud computing in Europe, and to examine how to address the low rate of adoption of cloud computing by organisations in the European Union (EU) compared to other jurisdictions around the world. A key premise is that cloud computing is essential to digital transformation ("DT") and digitalization. These, in turn, are essential components of multiple EU strategies to improve the EU's economic performance while reducing carbon emissions. Key objectives of the project are to identify obstacles to cloud adoption in the EU and to propose strategic actions to increase adoption, thereby enabling accelerated digital transformation and the benefits associated with DT. This document presents strategic action options identified to address these challenges.

Cloud computing is not the only component or aspect of “digitalization.” The original scope of the H-CLOUD project was expanded at the outset to include other components associated with digitalization: edge computing, green ICT, and the possible role of federation as a mechanism for improving the rate of adoption of cloud computing in Europe. Other important components, including high performance computing (HPC), artificial intelligence (AI) and big data (BD) are out of scope of the H-CLOUD project, nevertheless they are touched on in this analysis since all these technologies are important to digital transformation and build on or complement cloud technologies. Many of the initiatives related to the proposed EU Data Spaces[17], which are fundamentally “big data” initiatives, are considered in this analysis.

To assess the prospects for cloud computing and digital transformation, H-CLOUD used a two-pronged approach, looking at both the demand-side dynamics, as well as the structure and challenges faced by the supply side[18].

- On the demand side we evaluated the suitability of cloud-based products and services to support and enable digital transformation, as well as each sector’s capacity for transformation. This capacity for transformation is partly a function of the sector’s technological readiness, but the H-CLOUD team also identified important business and organisational factors that hinder both cloud adoption and digital transformation, and that might be addressed through additional strategic actions more related to deployment than to research or innovation. To make this examination tractable H-CLOUD selected a set of demand-side scenarios, specifically six of the sectors targeted in the “A European strategy for data”: public administration[19], healthcare[20], transport[21], energy[22], agriculture[23] and manufacturing[24]. To provide a horizontal perspective, the needs of SMEs were also examined[25]. While these demand scenarios are not exhaustive it was expected that they would collectively expose a variety of demand side issues of potentially broader significance.
- On the supply side H-CLOUD considered the role of non-European-based cloud service providers (CSPs), the competitiveness of EU-based CSPs, as well as the changing dynamics of the cloud as the computing becomes increasingly distributed across the “cloud-to-edge continuum”. The H-CLOUD team explores strategies for EU-based CSPs to improve their competitiveness and to gain traction in the growing edge-computing market[26,27,28,29,30].

In addition to H-CLOUD’s own analysis, the H-CLOUD project has considered several valuable analyses from other groups, including

- European Alliance for Industrial Cloud, Data and Edge (EAICDE)[31]

17 EC. A European strategy for data. 2020
18 H-CLOUD. D3.1 – Strategy Analysis Report and Cloud Computing v1.0. 2021
19 H-CLOUD. Demand Analysis: Public Administration (D-PA). 2021
20 H-CLOUD. Demand Analysis: Health (D-H). 2021
21 H-CLOUD. Demand Analysis: Transport (D-T). 2021
22 H-CLOUD. Demand Analysis: Energy (D-E). 2021
23 H-CLOUD. Demand Analysis: Agriculture (D-A). 2021
24 H-CLOUD. Demand Analysis: Manufacturing (D-M). 2021
25 H-CLOUD. Demand Analysis: SMEs (D-S). 2021
26 H-CLOUD. Supply Analysis: Landscape (S-L). 2021
27 H-CLOUD. Supply Analysis: Technology (S-T). 2021
28 H-CLOUD. Supply Analysis: Edge (S-E). 2021
29 H-CLOUD. Supply Analysis: Cloud Federation (S-F). 2021
30 H-CLOUD. Supply Analysis: Green ICT (S-G). 2021
31 AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. 2021

- Future Cloud Cluster (FCC)[32]
- Environment Agency Austria & Borderstep Institute (EAA&BI)[33]

1.2. Structure of this Report

Cloud computing, as well as the related technologies addressed by the H-CLOUD project, are essential tools that will enable the digital transformation of, and data-driven innovation by, the EU economy and its roughly 24 million enterprises. This report uses the detailed insights generated by the H-CLOUD project to make recommendations about how these technologies can be used most effectively to achieve these objectives. Since these technologies, these tools, are necessary, but not sufficient for this transformation, we also address broader challenges and opportunities related to digital transformation.

As we will see in our review of the EU policy context, the desired outcome is digital transformation and data-driven innovation, retaining both data and digital sovereignty, while respecting the environment. H-CLOUD's strategic analysis and recommendations can be organised logically into the following connected areas, or "pillars":

Pillar 1: Data Aware EU organisations. There are several components to this:

- *Functional, sovereign, affordable, environmentally friendly IT services at the edge.* Edge technologies have the potential to become as important as cloud computing for digital transformation, since they tap into the large volumes of operational, and strategically important, data that is generated at the edge and should remain there.
- *Well-governed business and data ecosystems; enabling secure, protected data exchange and sharing that respects the rights of data subjects and data owners.* A growing number of EU policies highlight the need for organisations to work together, rather than individually, to achieve critical objectives around climate neutrality and social well-being.
- *Accessible, understandable, ethical tools, capabilities, and services to extract knowledge and insights from large amounts of data* (big data analytics, artificial intelligence). This topic is extensively addressed by other initiatives, and outside the scope of H-CLOUD, but has been correctly identified as a strategic enabler for digital transformation.

Pillar 2: Data Driven Innovation, Enabled by New IT Paradigms. Supporting the smooth digitalization organisations need naturally depends on their adopting flexible cloud-style IT paradigms. However, "cloud-style IT" is not the same as "cloud computing" *per se* -- it addresses not only technical but also organisational approaches. It is more than just "using the public cloud" -- many organisations are making informed decisions to keep their IT infrastructure on-premises, but they are adopting the principles of cloud-style development and operations to enable their own agile and adaptable, yet secure and sovereign, digital futures.

Pillar 3: Strengthening the EU Cloud Market. EU organisations need to assemble the right combinations of IT services, with the right attributes and prices, to allow them to meet their strategic objectives. They need clearly defined options from their IT suppliers:

- EU IT suppliers, including EU cloud service providers (CSPs), must be encouraged to provide the transparency, flexibility, sovereignty, and sustainability that EU customers need to meet their own objectives and meet their own obligations to society.
- Non-EU suppliers must be held to the same standards applied to EU suppliers. The principles of data sovereignty and digital sovereignty should be incorporated in legislation and enforced and tracked with mechanisms such as the ENISA EU Cybersecurity Scheme.

---

[32] Future Cloud Research Roadmap. 2020

[33] EC. Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market. 2020

- As the edge grows in strategic importance, these quality attributes must also be built into the architecture of the edge from the ground up. Given the diversity and distribution of assets at the edge, and their role handling growing volumes of data, establishing intrinsic, foundational trust at the edge becomes critical.

Pillar 4: Foundational Elements of Cybersecurity, Environmental Sustainability, Skilled People, and Research and Innovation. Fundamental capabilities are needed in several areas:

- *Carbon Neutral Digital Transformation.* While the EU specifies "energy-efficiency" throughout its policies, the full "stack" of information technologies -- from the data centre or hosting site, all the way to running software and user-facing services must be constructed and operated within specific limits on environmental impacts.
- *Intrinsic Security.* Security cannot be added on at the last minute, it must be built in from the ground up. This topic is extensively addressed by other initiatives, and outside the scope of H-CLOUD, but will be an essential enabler for digital transformation.
- *Skills*. H-CLOUD identifies several specific skills gaps that must be closed to enable a seamless value chain supporting digital transformation. These gaps must be addressed in the context of the EU's broader workforce challenges, which are outside the scope of H-CLOUD.
- *Research*. The EU's cloud computing ecosystem must benefit from and exploit the latest research in cloud computing and related digital technologies.

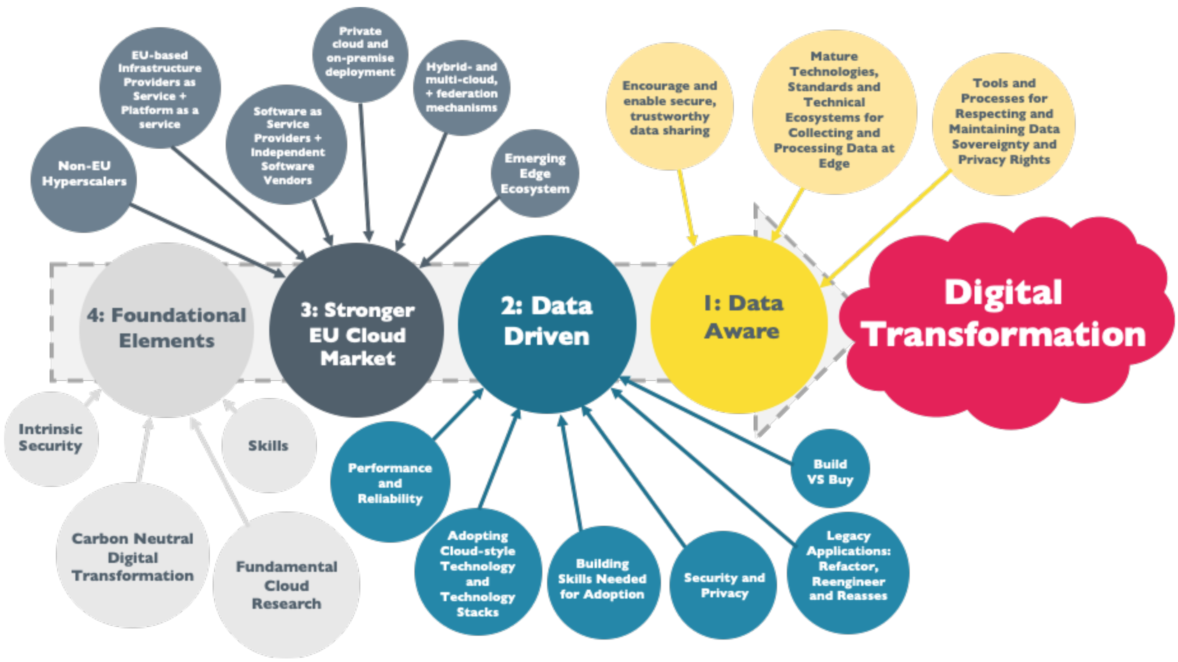


Figure 3. The Logic Model of H-CLOUD's Cloud Computing Strategy

The rest of this report is organised based on this logical structure.

2. CONTEXT

Digitalization or digital transformation, i.e., the adoption of digital technologies to enhance processes, is becoming more and more important in our society as demonstrated by COVID-19 pandemic, which unexpectedly contributed to its rapid acceleration[34]. Digitalization is not only expected to improve our daily life, increase the resilience of society, and innovate different business sectors, it also drives growth through improvements in time and cost efficiency. A recent study by Deloitte[35] highlights how digitalisation is a key opportunity for Europe: the study estimates that if all EU Members States were to reach a Digital Economy and Society Index[36] score of 90 by 2027, this would translate into an increase of the GDP per capita across the EU27 of 7.2% compared to an 'as is' scenario. Accordingly, EC is putting digitalization at the centre of its strategies.

This section describes the current context of digitalization in Europe, including the EU's current bold policy agenda, the benefits expected from Digital Transformation and realisation of a Digital Single Market, as well as the status of cloud adoption in the European Union.

2.1. EU Policy Agenda

In recent years the European Commission (EC) has invested to create the Digital Single Market, including a number of actions specific for cloud computing. In preparation of the Multiannual Financial Framework (MFF) of the European Union covering the 2021-2027 period, the EC set new priorities and directions, that complement and expand strategies defined in the previous MFFs[37]:

- **A European Green Deal**, noting that "becoming the world's first climate-neutral continent by 2050 is the greatest challenge and opportunity of our times".
- **An economy that works for people**, with the aspiration that "The EU's unique social market economy allows economies to grow and to reduce poverty and inequality. With Europe on a stable footing, the economy can fully respond to the needs of the EU's citizens."
- **A Europe fit for the digital age**, empowering people with a new generation of technologies and sustaining the Digital Single Market Strategy to create better and larger opportunities for European companies
- **Protecting our European way of life**, promoting a "vision for a Union of equality, tolerance and social fairness".
- **A stronger Europe in the world**, reinforcing Europe's role as responsible global leader working to ensure the highest standards of climate, environmental and labour protection.
- **A new push for European democracy,** ensuring a stronger role for European citizens in the decision-making process and in the setting of European priorities.

In general, these policies describe pressure from EC for more integration through ecosystems and dismantling of barriers to data sharing. At the same time, there is pressure for increased transparency, data sovereignty protection and security, and energy-efficiency.

Shaping Europe's digital future, provides the overall strategic plan for the European digital agenda in the next decade. In this new agenda, cloud computing will continue to play a key role to tackle the digital challenge in Europe, complementing and extending actions defined by previous strategies.

[34] Pedro Soto-Acosta. COVID-19 Pandemic: Shifting Digital Transformation to a High-Speed Gear, Information Systems Management, 37:4, 260-266. 2020
[35] Deloitte. Digitalisation: an opportunity for Europe. 2021
[36] The Digital Economy and Society Index (DESI) summarises indicators on Europe's digital performance and tracks the progress of EU countries.
[37] EC. 6 Commission priorities for 2019-24. 2019

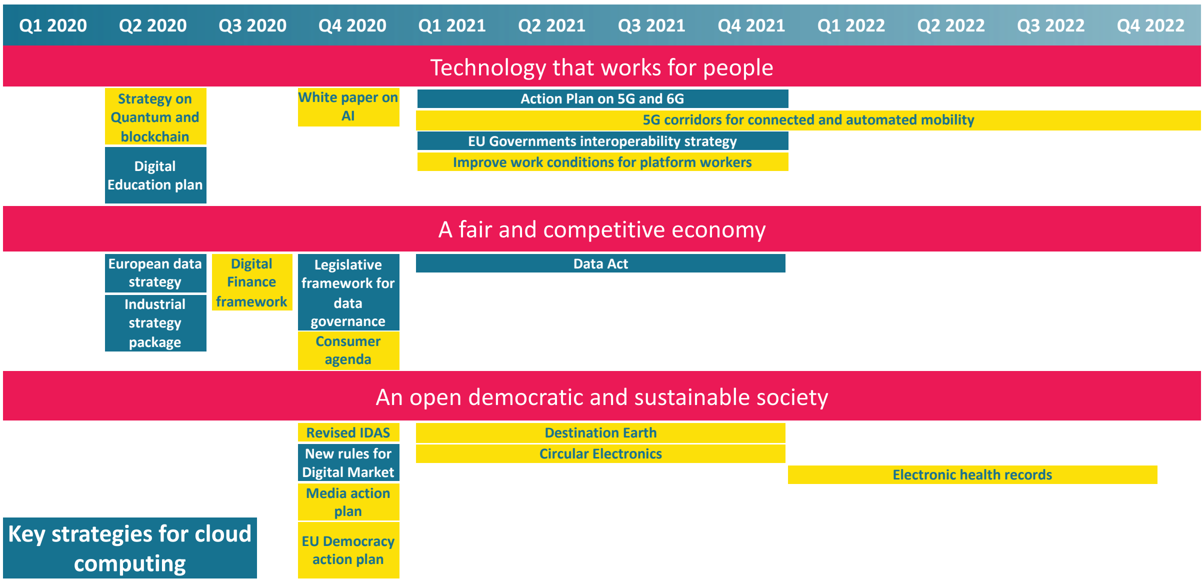


Figure 4. Roadmap of key EC strategies, the ones with blue background are highly relevant to cloud computing industry.

Cloud computing (and digital technologies in general) can play even a stronger role in European economy and society by embracing core European values, from fundamental individual rights (e.g., security and data privacy), to market openness (e.g., interoperability and free flow of data) and environmental friendliness (e.g., reduced carbon footprint and energy consumption).

Following the COVID-19 pandemic, green and digital transition have become more and more the cornerstones for the EU's vision for 2030[38], with President von der Leyen committed to invest at least 20% of Next Generation EU[39] (approximately €150 billion) into digital initiatives. This reinforced focus is central to newly released policies in different sectors, including mobility[40], energy[41], circular economy[42] and agri-food[43]. All these sectoral policies emphasise similar needs: supporting the creation of wider ecosystems, reducing environmental impact, and increasing "understanding" of sector needs by facilitating trusted and secured data sharing and exchange (under fair conditions).

To tackle the digital challenge and enable data sharing in Europe, the EC released several policies, including: *A European strategy for data*[44] and *A new Industrial Strategy for Europe*[45] that describe a *High Impact Project on European data spaces and federated cloud infrastructures*, as a core European project to unleash the potential of the data economy and a digital single market in Europe. The industrial strategy includes actions dedicated to SMEs[46], aiming at supporting their transition toward a green and digital economy, actions cover: digital upskilling, investments on green transitions, and reinforcement of Digital Innovation Hubs. Together with other key policies on data governance, digital services, digital markets, and cybersecurity, the EC has laid out the principles for the next generation of cloud and digital services in Europe. To measure the achievements in relation to EU digitalisation, EC set in the *2030 Digital Compass*[47] several ambitious performance indicators, including "10,000 climate

[38] EC. Stepping up Europe's 2030 climate ambition. 2020
[39] EC. NextGenerationEu. 2020
[40] EC. Sustainable and Smart Mobility Strategy. 2020
[41] EC. Powering a climate-neutral economy: An EU Strategy for Energy System Integration. 2020
[42] EC. Europe's moment: Repair and Prepare for the Next Generation. 2020
[43] EC. EC, Farm to Fork Strategy, 2020
[44] EC. A European strategy for data. 2020
[45] EC. A new Industrial Strategy for Europe. 2020. Updated in Updating the 2020 New Industrial Strategy: Building a stronger Single Market for Europe's recovery. 2021
[46] EC. An SME Strategy for a sustainable and digital Europe. 2021
[47] EC. 2030 Digital Compass: the European way for the Digital Decade. 2021

neutral highly secure edge nodes" and "75% of EU companies using Cloud/AI/Big Data" by 2030.

Fostering an active role by industry in the European digitalization process, while adhering to EU core values and principles, the EC endorsed the establishment of the *European Alliance for Industrial Data, Edge and Cloud* (July 2021). In addition to that, as part of the EC digital strategy, the EC has also committed to deliver a "harmonised set of self-regulatory norms and standards for Cloud services in Europe" (the so-called *EU Cloud Rulebook*) that will incorporate existing EU regulatory frameworks (GDPR[48], FFoD[49], DORA[50], etc.) and reference validated frameworks related to, for example, Data Protection Codes of Conduct validated by EDPB.

Key components of the MFF 2021-2027 that support the take up of digital transformation in Europe include:

- Horizon Europe, the new research and innovation programme.
- Digital Europe, a brand-new programme focusing on building the strategic digital capacities of the EU.
- Connecting Europe Facility 2, focusing on the creation of transnational digital infrastructures.

These three programmes are designed to complement each other and will play a key role in Europe's digitalisation in connection with cloud computing. In particular, the introduction of the Digital Europe programme may play a key role in supporting the deployment of mature research and innovation outcomes, bridging the gap between the research and market penetration phases, and helping to overcome the so-called "valley of death"[51] depicted in Figure 5.

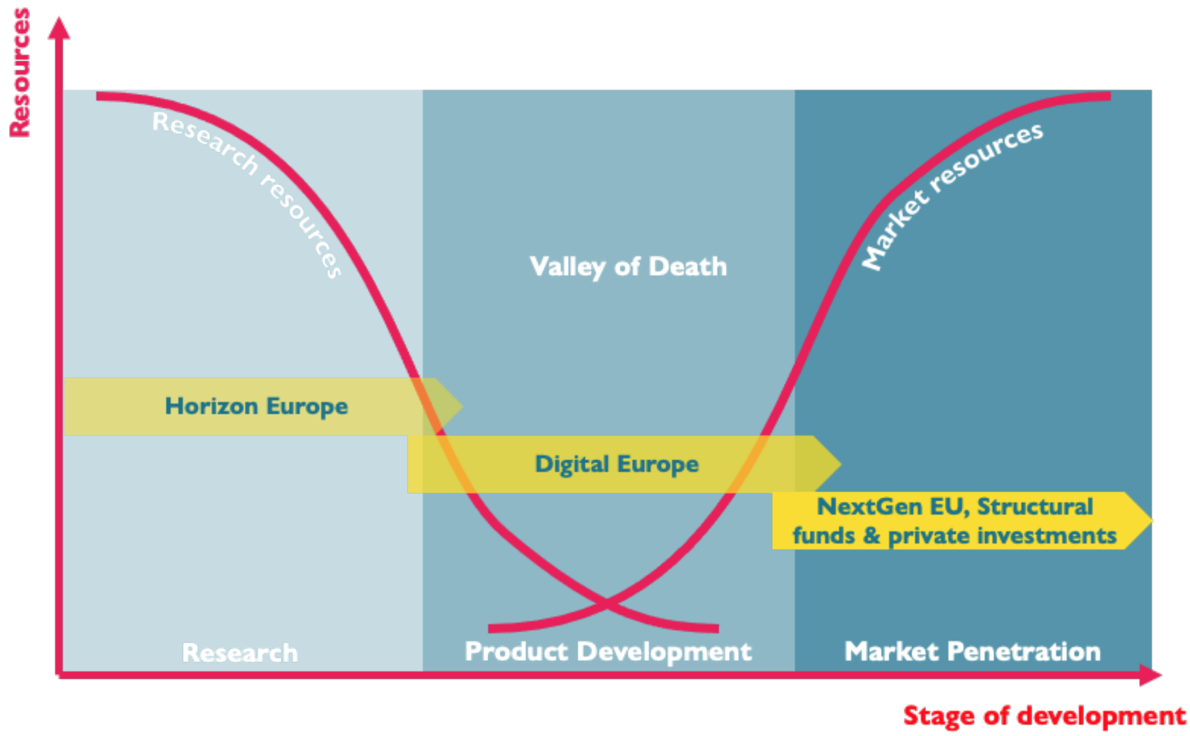


Figure 5. EC vision to support technology roll-out from research to market penetration

The programmes are complemented by the NextGenerationEU[52] plan whose ambition is the creation of a more green, more digital, more resilient, and more equal Europe and that will play

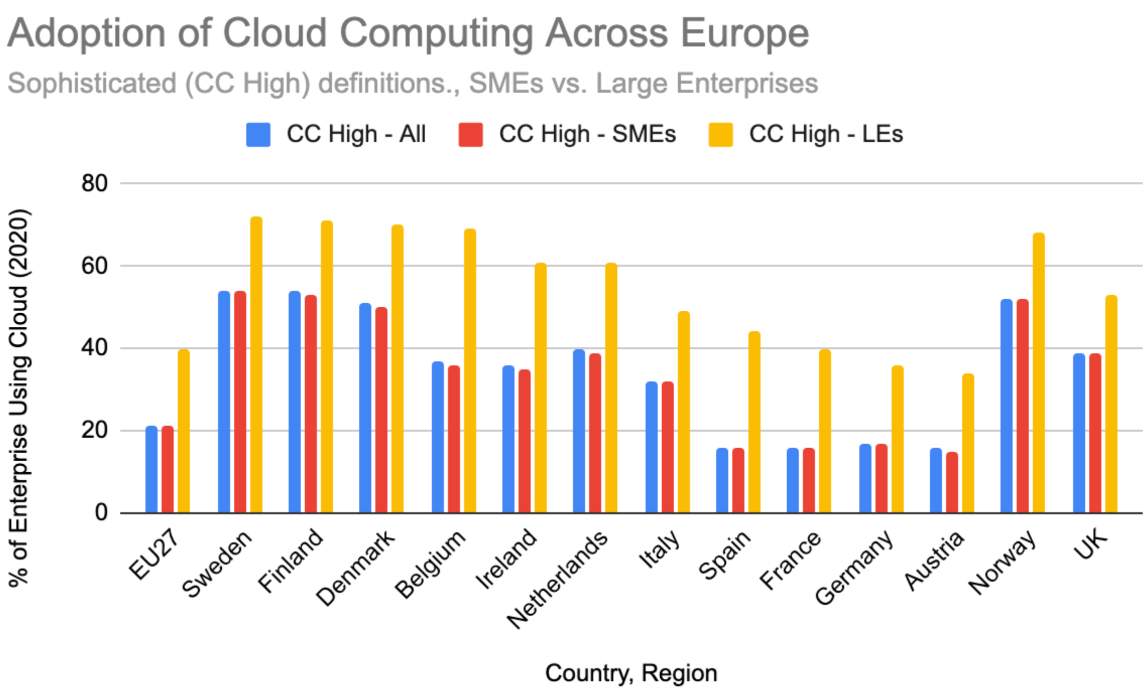
[48] EU. General Data Protection Regulation. 2016
[49] EU. Framework for the free flow of non-personal data in the European Union. 2018
[50] EU. Digital Operational Resilience Act. 2020
[51] Sensenbrenner, F.J. "Unlocking Our Future: Toward a New National Science Policy", Committee Report, 105-B, p.40, 2008
[52] EC. NextGenerationEu. 2020

a key role in the market penetration of digital technologies, including cloud and edge computing.
The regulatory role of EC is complemented by industrial initiatives in the different sectors, that on one side builds and contributes on the strategy promoted by EC -- as the Industrial Roadmap[53] promoted by a group of EU CEOs - and on the other side influence the EC Agenda and become corner stones in it -- as the GAIA-X initiative that is clearly behind the push for cloud federations in EC strategy and that was referred as the cornerstone of EU Digital strategy by the President of the European Commission Ursula Van der Leyen in her State of the Union Address[54] in September 2020.

2.2. Status of Cloud Computing Adoption in Europe

Statistics on the adoption of cloud computing by EU enterprises present a complicated picture. According to Eurostat, using a range of definitions of cloud computing, the proportion of enterprises adopting cloud computing varies significantly depending on both enterprise size and location.

Figure 6 and Figure 7 present 2020 adoption statistics for two definitions of cloud computing, for the entire EU27 as well as a selection of larger and higher adoption member states (including the former EU member UK, as well as Norway), distinguishing between small- and medium-sized enterprise (SMEs, up to 250 employees) and large enterprises (LEs, more than 250 employees). Figure 8 and Figure 9 show the growth of these adoption statistics from 2018 to 2020. In these figures, “CC Low” equates to Eurostat’s definition of cloud computing to be “buy cloud computing services used over the internet”, which would include activities as simple as using web-based email applications or office productivity software (e.g., word processing). “CC High” equates to Eurostat’s most rigorous definition of cloud computing: “Buy high CC services (accounting software applications, CRM software, computing power)”, which requires an enterprise to be using the cloud for its accounting, customer relationship management, and for general computing in support of other applications.



[53] European industrial technology roadmap for the next generation cloud-edge offering. 2021
[54] EC. State of the Union Address by President von der Leyen at the European Parliament Plenary. 2020

Figure 6: Cloud Computing Adoption – "CC High" – Across the EU[55]

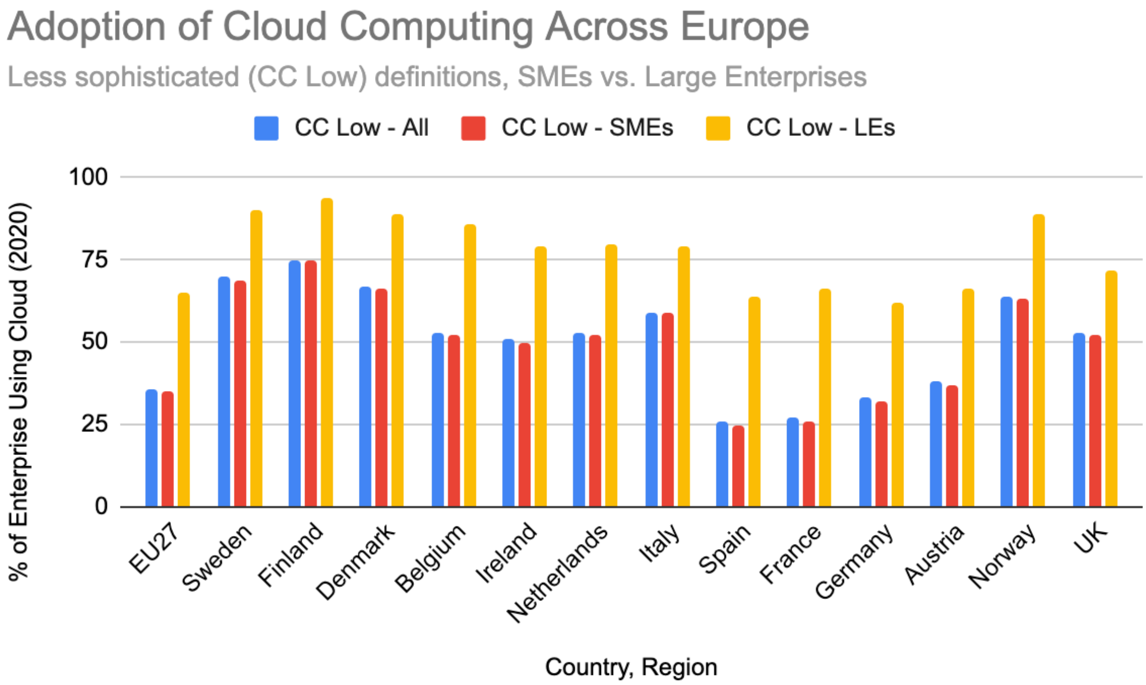


Figure 7: Cloud Computing Adoption – "CC Low" – Across the EU[56]

Figure 6 and Figure 7 illustrate the diversity of adoption levels across the EU. Statistics for only a few selected countries are shown individually, and in most cases adoption levels for these few countries exceeds that of the EU27 overall, and there are many countries with lower levels of adoption, but in most cases these are comparatively smaller countries (in terms of GDP).

Note that for the EU Member States shown in Figures 4 and 5, they are presented in order of declining adoption of the highest form of cloud computing defined by Eurostat ("CC High") for each country's large enterprises. This statistic probably provides the best indication of economically significant cloud adoption for each country. This figure aligns with survey-based estimates of cloud adoption for multiple countries and appears to correlate with estimated levels of public cloud spending.

Figure 8 and Figure 9 show that the pace of adoption varies widely across member states and enterprise sizes -- measured by the increase in adoption from 2018 to 2020. For example, in Italy, adoption of the cloud at the more mature level appears to have increased by 191% from 2018 to 2020, but there do not appear to be any fundamental drivers for this rapid increase. Nevertheless, at this rate of increased adoption, Italy might reach 75% adoption for all enterprises by 2024. The COVID-19 pandemic is expected to further accelerate the pace of adoption beyond 2020 levels.

[55] Eurostat. Cloud computing services adoption. 2021
[56] Eurostat. Cloud computing services adoption. 2021

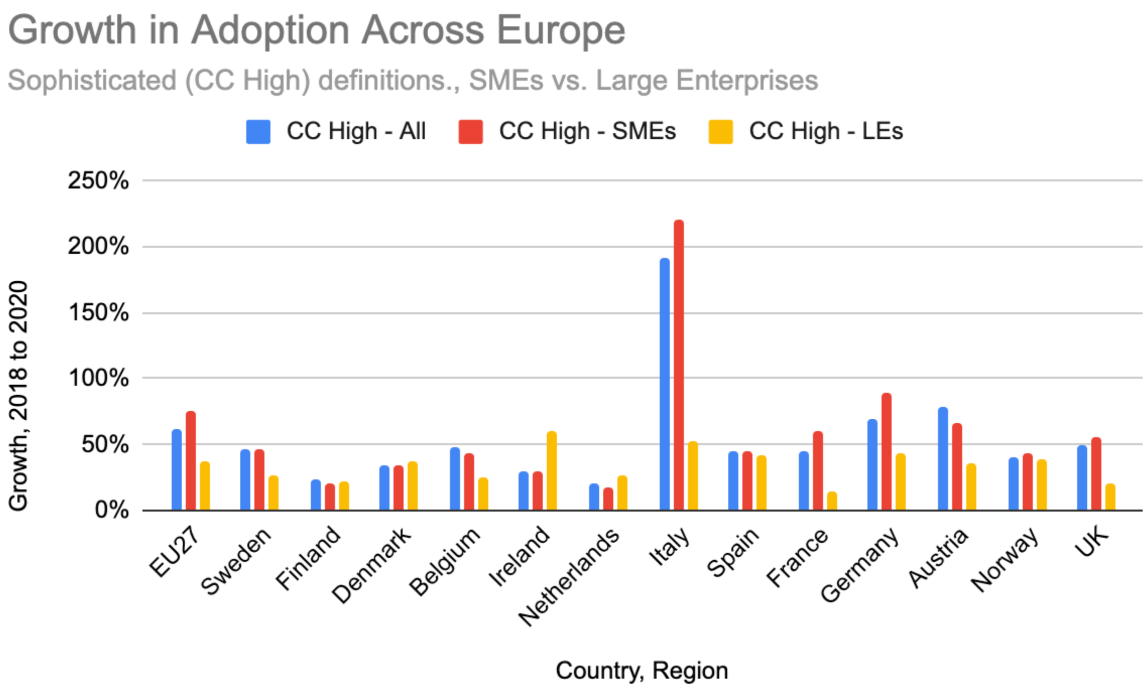


Figure 8: Cloud Computing Growth (2018 to 2020) – "CC High" – Across the EU[57]

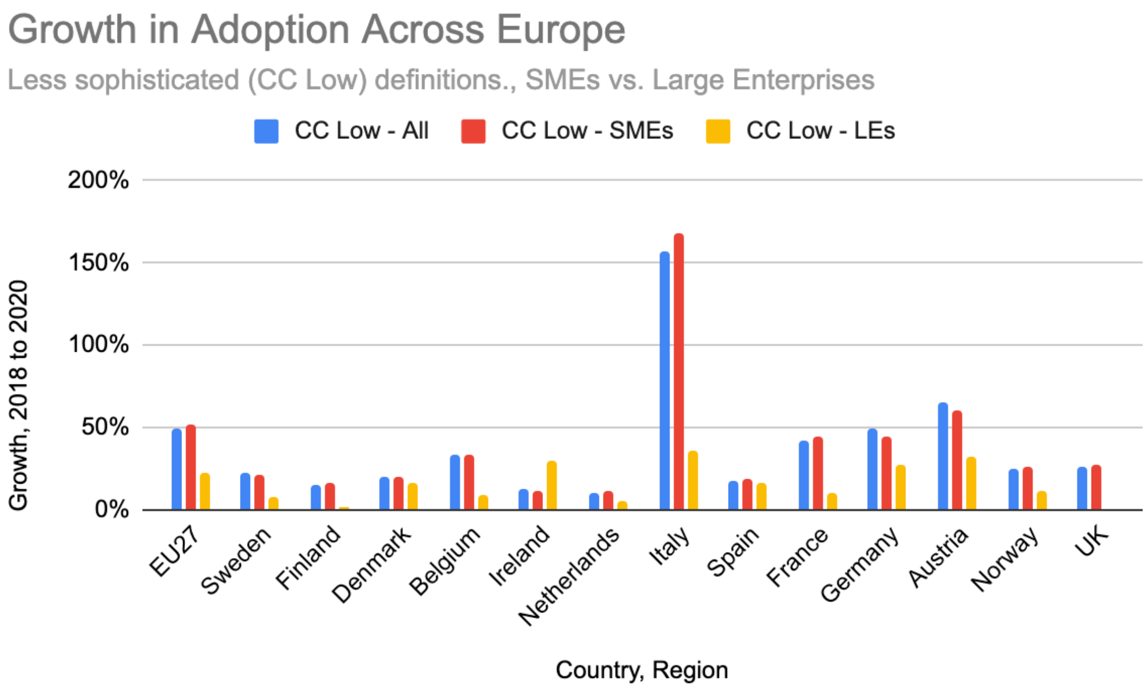


Figure 9: Cloud Computing Growth (2018 to 2020) – "CC Low" – Across the EU[58]

Outside the EU, most countries, including the US, do not collect comparable data that is statistically significant for their entire economies. One exception is Canada, where the use of "cloud computing" (not defined in detail) has been measured for over 15,000 private sector

[57] Eurostat. Cloud computing services adoption. 2021
[58] Eurostat. Cloud computing services adoption. 2021

enterprises[59]. Rates of use of 35.3% were found for enterprises with up to 19 employees, 52.1% for enterprises with 20-99 employees, and 73.3% for enterprises of 100 employees or more. The overall adoption of cloud computing was 38.9%.

Most other cloud adoption statistics report the results of enterprise surveys, with smaller samples, typically of larger enterprises, which provide insight into cloud adoption but can only be compared to some of the figures above. For example, IDC's May 2020 *Worldwide Industry CloudPath Survey* showed relatively high adoption among the enterprises surveyed in each of the following countries (see Figure 10).

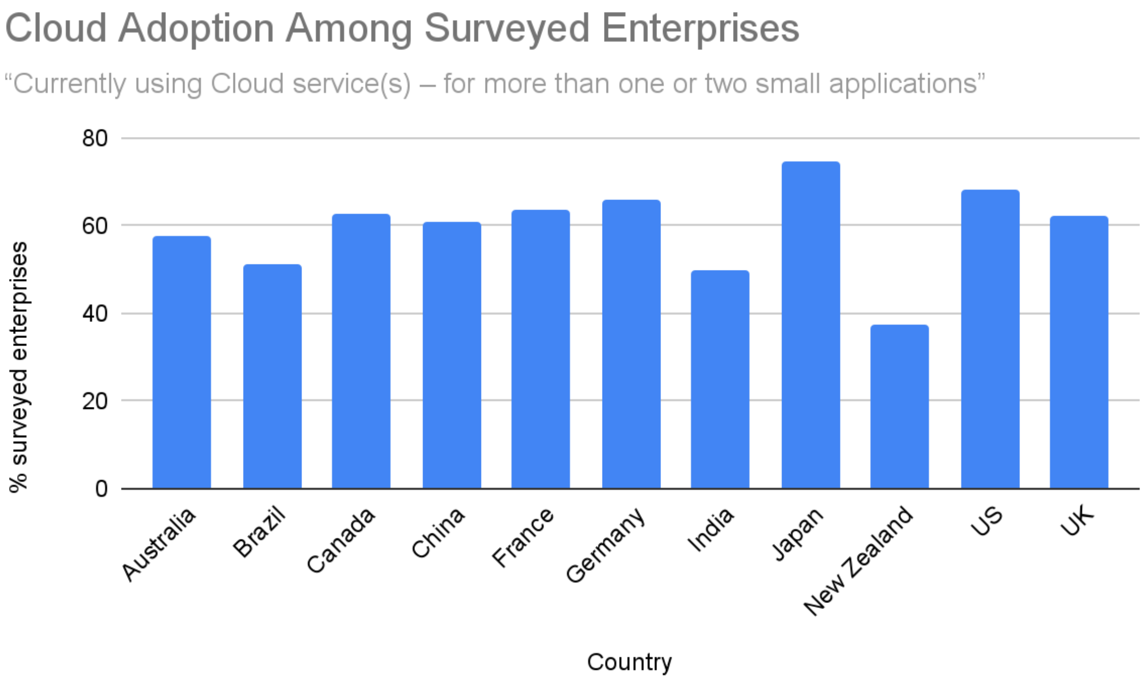


Figure 10: IDC Survey of Enterprise Cloud Adoption Across Multiple Countries[60,61]

Statista reports that 68.7% of over 2,200 Japanese enterprises had either partially or fully implemented "cloud computing services" as of September 2020[62].
Although the data is not completely consistent, these statistics and survey results suggest that, overall, the largest European enterprises are adopting the cloud at roughly the same levels as enterprises in other jurisdictions around the world. The differences between official statistics and survey-based figures may reflect a wider use of cloud computing in the context of "Shadow IT", where individual business units and departments decide for themselves to take advantage of cloud-based IT solutions rather than working through their enterprises' IT departments, which might be adopting cloud-based solutions more slowly. Regardless of the reasons, there are still EU member states where large enterprise cloud adoption appears to lag.

SMEs represent over 99% of enterprises in almost every country around the world, so SME adoption rates dominate economy-wide cloud adoption rates everywhere. Figure 4 (above) shows that SME cloud adoption lags adoption by larger enterprises in all EU member states, so the target set by the EC's 2030 Digital Compass of 75% cloud adoption by all EU enterprises will be the most challenging for SMEs. However some member states will reach the desired

[59] Statistics Canada. Table 22-10-0117-01 Information and communication technologies used by industry and size of enterprises. 2020.
[60] IDC. Industry CloudPath. 2020.
[61] IDC. Industry CloudPath: Industry Application and Workload Cloud Deployment. May 2020
[62] Statista. Cloud service penetration rate among business enterprises Japan. 2020

level of adoption within a few years, given the current pace of adoption. Other member states may need targeted support for cloud adoption, regardless of the size of enterprise.

These adoption statistics measure whether or not an enterprise has reached various defined levels of adoption (e.g., "CC Low" and "CC High" as detailed above). The extent of cloud adoption beyond that level is also important, but harder to measure. A rough indicator could be the level of spending on public cloud services in different countries and regions, normalized as a proportion of country/region GDP. Figure 11 presents this indicator for several countries as well as for the EU27.

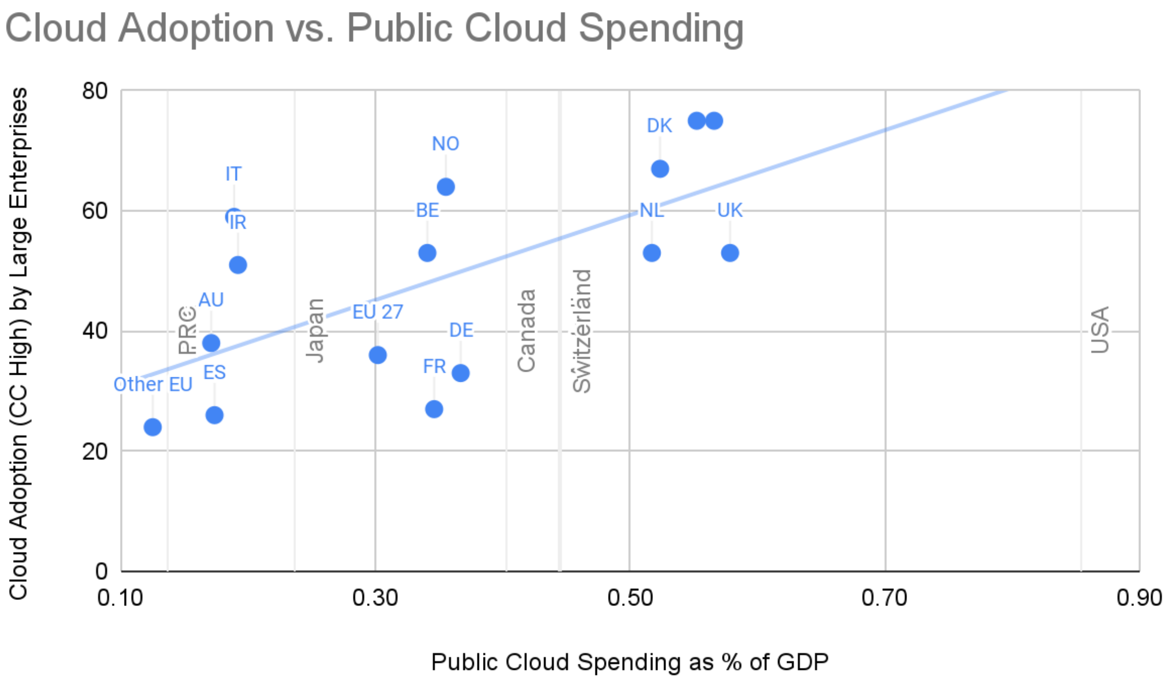


Figure 11: Cloud Adoption vs. Public Cloud Spending Across Multiple Countries[63]

This measurement correlates to some extent (R^2=0.43) with the cloud adoption statistics from Eurostat, but there are significant outliers such as Ireland, Italy and Japan, where cloud adoption (using the stronger definition) is relatively high among large enterprises, but cloud spending (as a proportion of GDP) seems very low. Conversely, in France and Germany, cloud adoption even among large enterprises appears low, while cloud spending is higher than might be expected.

Given the importance placed on the adoption of cloud computing by the EC, there should be more robust measurement of the current state of cloud adoption and spending across the EU.

2.3. Overview of EU Public Cloud Market

Appendix 1: "Public Cloud Market Segments in Selected Markets" provides IDC data[64] that characterises the EU public cloud market, along with comparisons with this market in other parts of the world. Some of this data is illustrated in Figure 12.

[63] H-CLOUD analysis of IDC Public Cloud's Contribution to the European Economy: A Macroeconomic Approach
[64] IDC. Worldwide Semiannual Public Cloud Services Tracker, 2021

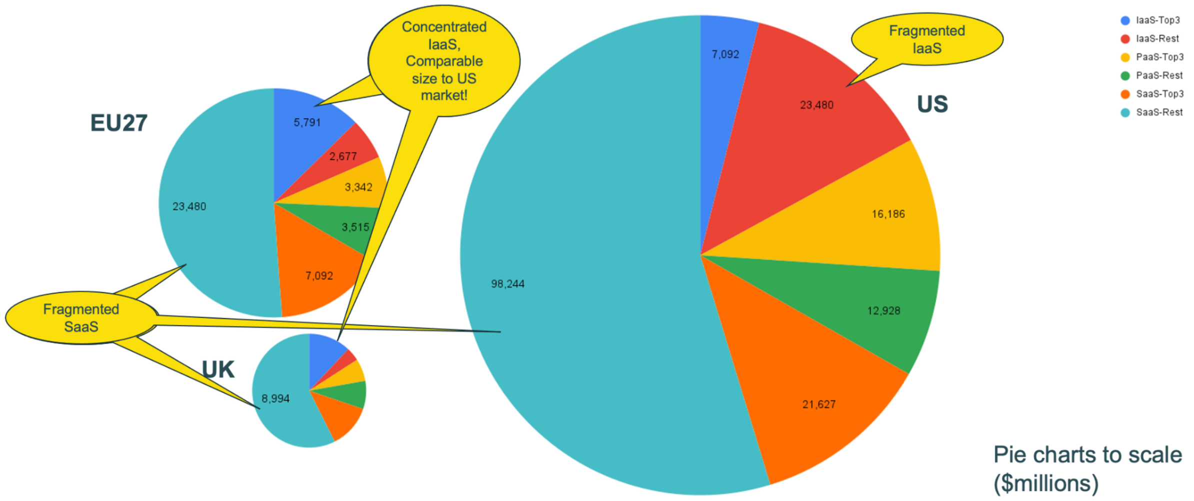


Figure 12: 2020 Market Shares of Top 3 Providers for I/P/SaaS Market Segments in EU27, UK and US[65]

Several important observations can be made:

Spending Intensity: As noted in Section 2.2, the "intensity" of public cloud spending varies widely across countries and regions (as a proportion of GDP), and overall cloud intensity in the EU27 is low compared to the US and UK.

- While the US economy is only about 50% larger than that of the EU27 (based on GDP -- $21B vs. $15B in 2020), the US public cloud market was 4 times larger than that of the EU27 in 2020.
- The EU27 public cloud market is only 3 times larger than that of the UK, while EU27 GDP is over 5 times that of the UK.
- It is unclear the extent to which the large number of US and UK based data centres and cloud-based software businesses may skew these figures.

Across these markets, IDC projects between 18-22% compound annual growth in the total public cloud market from 2021-2025, except for Japan, where they forecast roughly a 14% CAGR.

Software-as-a-Service Market Segment: Across all these countries and regions, the software-as-a-service (SaaS) market segment dominates the total public cloud market, accounting in 2020 for 67-70% of the total market in the EU27, US and UK, and 48-50% of the total market in Japan and Korea.

- The importance of this segment will fall in all these countries and regions since the SaaS segment is projected to grow between 2-7% more slowly than the total market. By 2025, the SaaS market segment will still dominate public cloud spending but will account for between 50-60% of the total market, except in Japan and Korea, where SaaS will account for only 35% and 46%, respectively, of the total market.

The SaaS market segment is the most diverse of all the public cloud market segments, with the top 3 providers in each country and region controlling small portions of the SaaS segment, together only accounting for between 18% and 29% of the total SaaS market in that location.

- In Europe the top 3 providers are Microsoft, Salesforce, and the German SAP. Visma is another notable EU-headquartered SaaS provider.

Platform-as-a-Service Market Segment: Platform-as-a-service (PaaS) currently represents the smallest segment of the total public cloud market, accounting in 2020 for only 15-17% of the total market, except in Korea where PaaS accounts for only 9% of the total market.

- The importance of this segment will grow in all these countries and regions, with the PaaS segment projected to grow between 9-11% more rapidly than the total market (except in Japan where PaaS is projected to grow only 4% more rapidly than the overall

[65] IDC. Worldwide Semiannual Public Cloud Services Tracker, 2021

market). By 2025, the PaaS market segment will have grown to account for between 20-24% of the total market, except in Korea, where PaaS will only reach 15% of the total market.

The PaaS market segment, in all the countries and regions presented above, is somewhat concentrated, with the top 3 providers in each country and region controlling between 43% and 56% of the PaaS segment.

- In Europe the top 3 providers are Microsoft, AWS and Google. SAP is the only notable EU-headquartered PaaS provider.
- Large US-headquartered providers dominate this market segment in all the other countries presented here, with Microsoft, AWS, Google, Salesforce and Oracle collectively holding the top 3 ranks in each of the other countries.
- Despite this market concentration across so many countries, the PaaS market segment in the US still represents the larger opportunity for the top US PaaS providers.

Infrastructure-as-a-Service Market Segment: Infrastructure-as-a-service (IaaS) segments in each country and region are slightly larger than the corresponding PaaS segments, accounting in 2020 for 16-18% of the total market in the EU27, US and UK, and 35-39% of the total market in Japan and Korea

- The importance of this segment will grow in all these countries and regions, except for Korea. The IaaS segment is projected to grow between 6-8% faster than the total market in all jurisdictions, except Korea where IaaS spending is not projected to grow significantly. By 2025, the IaaS market segment will have grown to account for between 22-24% of the total market in the EU27, US and UK, and 39-45% of the total market in Japan and Korea.

The IaaS market segment, in all the countries and regions except the US, is highly concentrated, with the top 3 providers in each country and region controlling between 68% and 85% of the IaaS segment. By contrast, in the US, the IaaS market is very diverse, with the top 3 providers accounting for only 23% of the US IaaS market segment.

- In Europe the top 3 providers are AWS, Microsoft, and Google. OVHcloud and Orange are the only notable EU-headquartered IaaS providers in the Top 10, based on IDC data.
- Large US-headquartered providers dominate the IaaS market segment in all the other countries presented here, with AWS, Microsoft and Google collectively holding the top 3 ranks in each of the other countries, except in Korea and Japan where Google is edged out by Korea Telecom and Fujitsu, respectively.
- In contrast to the PaaS market segment, for AWS, Microsoft and Google, the EU27 and UK together represent a larger market for IaaS services than does the US. These US-headquartered providers should be expected to strongly defend their European market shares.

Based on these observations several key conclusions can be drawn:

- Achieving the EU's objectives for digitalization could easily translate into increased intensity of public cloud spending in the EU27, moving from the current 0.3% of GDP toward the 0.8% of GDP being devoted to public cloud spending in the US.
- If this increased level of spending follows current patterns, it will benefit US-headquartered providers rather than EU-headquartered providers, particularly in the IaaS and PaaS market segments. However, there is an opportunity to create a level playing field that will "lift all boats", giving US-headquartered providers proportional access to the massive market growth that the EU needs to reach its goals.
- If this increased level of spending is increasingly devoted to services delivered from the edge, this creates a new market that must be centred in the EU. However, supporting

a new market of such size will require significant investment in new infrastructure at the edge.

- The worldwide diversity of the SaaS market segment represents both an opportunity and a challenge. The opportunity is for EU-headquartered SaaS providers, and independent software vendors in general, to improve their ability to create value for their customers and enable the digital transformation that is needed across all sectors of the EU economy, and thereby take a principal role in delivering on the EU's priorities. However, this segment faces a two-fold challenge -- competing in the fragmented EU market of 27 sovereign Member States and finding capable IaaS and PaaS partners and platforms on which to build secure, sovereign, and sustainable services.

2.4. Success Stories and Good Practices in the European Cloud Computing Landscape

To support the analysis and the definition of recommendations in the context of cloud, federation, edge, and green IT, H-CLOUD catalogued[66] more than 240 relevant public and/or private initiatives, out of which 56 have been further analysed and identified as Good Practices[67] with the aim of understanding successful patterns for the deployment of cloud federation, edge, and green IT solutions and feeding these success patterns into the H-CLOUD strategic agenda.
Key findings can be summarised as follows:

- **Federated Cloud.** So far, federation projects are more successful in the public sector than in the private sector because the strategic incentive is strong to have full control over and sovereignty of IT infrastructure and to share data and insights across departments and agencies in the public sector, whereas the business incentives to create a federation are less strong in the private sector. As discussed also in H-CLOUD briefing paper on Cloud Federation, to support cloud federation adoption, challenges include: aligning the focus of federation model with objectives of its stakeholders; reaching agreement on technical and not technical standards among the different stakeholders; prioritizing identity and access control management (especially for data); capitalizing on the growing needs for secure private sharing of data. Valuable good practises to support these challenges have been identified in different success stories. For example, Cloud28+ created a community of service providers with a shared business interest. These providers publish their services using a joint service catalogue on the Cloud28+ platform; City Network has adopted OpenStack as its underlying technology to enable federation at the technology architecture level; and Aquacloud, Polymore, and Gaia-X are working to provide a standard data model to create value for participants in their ecosystems. Academic research networks had to find effective IAM solutions to manage users across dynamic communities.

- **Edge Computing.** Despite the early-stage deployment of edge solutions, the business case is often quite clear, as edge is seen as the enabler of use cases that could not be developed in other ways, thus diminishing doubt regarding ROI for edge solutions. Challenges include: maturity of edge solutions; complexity of regulation compliance. Good practises highlight: importance of investing in building the skills needed to sustain the next wave of innovation; centrality of 5G investments and availability; easing and rationalising regulations and governance concerning cloud-to-edge interoperability.

- **Green IT.** Green IT practices are often a by-product of the cloud infrastructure renewal rather than a goal of companies per se. To drive awareness and accountability in this area, it is important to create a set of KPIs on which projects, initiatives, and private companies need to report.

[66] https://catalogue.h-cloud.eu/
[67] H-CLOUD. D1.4 – Success Stories and Good Practices Guide. 2021

Based on the analysis of good practices, H-CLOUD presents several recommendations for policy makers and cloud providers that are included in our analysis.

2.5. Current technology trends

In the digitalisation arena, the technology landscape is evolving very rapidly, and analysts expect that multiple technologies will be instrumental in the future of digital infrastructure. Although many scenarios are possible, Figure 13 presents the outcomes of an analysis by IDC[68] and categorises technologies that are expected to bring “transformational”, “incremental” or “opportunistic” changes to the market, positioned in terms of adoption rates and time frames.

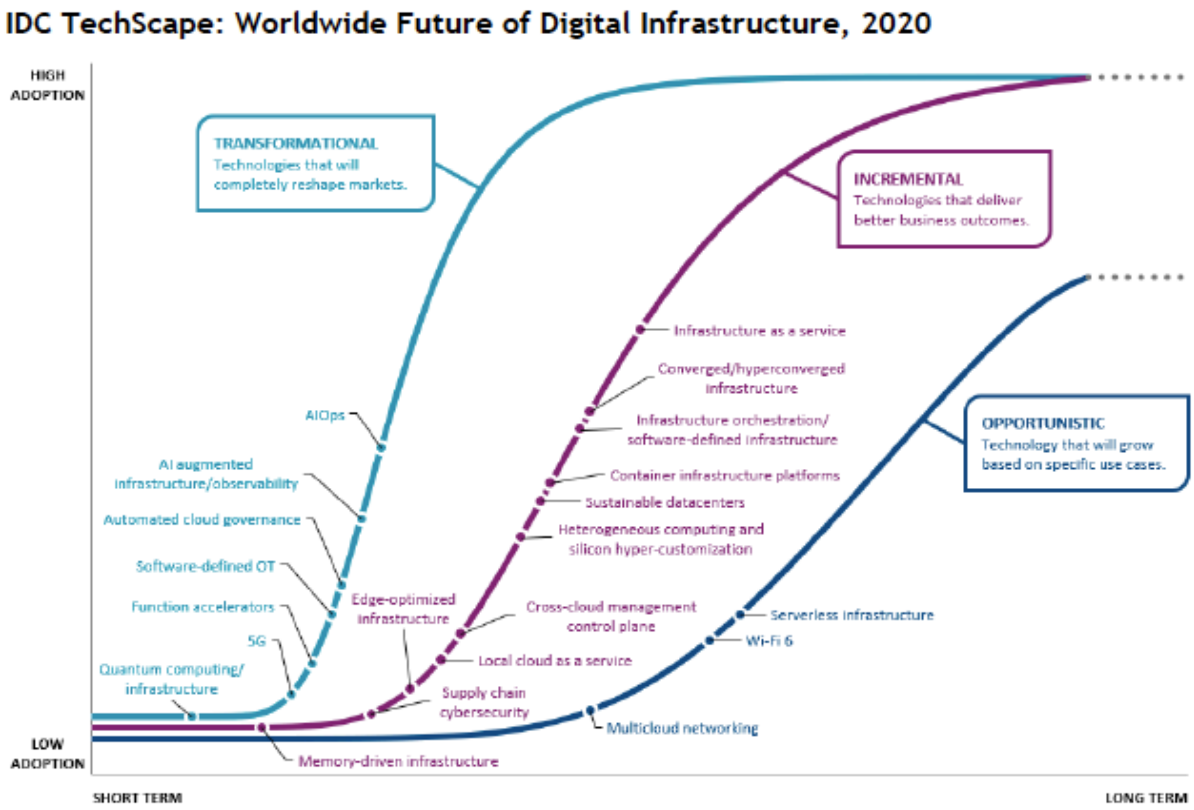


Figure 13: Worldwide Future of Digital Infrastructure[69]

In the first category, IDC places:

- Automated cloud governance, AI augmented infrastructure, and AI operations concepts relate to “Cognitive Cloud” and other similar terms used today in the research arena and aim at leveraging AI/ML to increase automation of the digital infrastructures.
- Software-defined OT (linked to the IT and OT convergence phenomena and supported by edge computing adoption) focuses on applying Software-defined infrastructure principles, largely adopted in the cloud, to Operational Technologies.
- Function accelerators extend the concept of “function” specific computing units (e.g. GPUs, ASICS) and make them accessible to hypervisors and containers to provide high performance execution of specific tasks.
- 5G encloses the last generation of fixed and mobile communication infrastructure, that not only adopts a new generation of radio antenna, but strongly leverage on cloud

[68] IDC. IDC TechScape: Worldwide Future of Digital Infrastructure. 2020
[69] IDC. IDC TechScape: Worldwide Future of Digital Infrastructure. 2020

native approaches to deploy and govern the communication infrastructure (building as well on “function accelerators”).

- Quantum Computing introduces infinite states between 0 and 1, thus enabling a new model of computing and ability to solve problems that classical computing (that accept a bit to be only 0 or 1) is not able to solve. This technology, after years of nice research, seems to be almost ready to be on the market and accessible via modern cloud-based platforms.

In the second category, key technologies include:

- Edge-optimised infrastructure and cross-cloud management plan, provide on the one side hardware that is designed for running at the edge (i.e. low power) and providing specific functions (e.g. GPUs), on the other side resource management solutions that are able to orchestrate seamlessly public cloud, private clouds and edges.
- Sustainable data centres aiming at improving energy efficiency and operational efficiency.
- Converged infrastructure, i.e. solutions packaging in a single unit multiple capacities of a digital infrastructure, spanning from computation, to memory, and from storage to network. Thus providing faster deployment of infrastructures and simplifying their management.
- Container infrastructure platform and software-defined infrastructure, i.e. solutions introducing “containers” as a modern and efficient way to deploy and manage applications at scale in the cloud, and managing the reproducible deployment of infrastructures via “code” and API calls.
- Supply chain cybersecurity, i.e. the end-to-end and integrated security of all the elements composing a digital infrastructure, spanning from heterogeneous hardware to heterogeneous services.

Beyond these infrastructure technologies, other trends that are facing on the market, or have already showed early success, are key instruments to digitalisation, this includes, for example:

- Digital Twin, AR and IoT platforms, bringing new ways to digitally represent the physical world.
- Distributed Ledgers and Smart contracts, enabling mechanisms for trusted exchange of information, its validation, and business agreement verification and enforcement.
- AI advancements, toward model compression and composite AI, thus increasing suitability of edge to run AI algorithms.

2.6. Research Landscape for Cloud Computing

The European Commission has actively invested in research on topics related to cloud computing and edge computing. There are currently 23 projects in progress, funded in the Horizon 2020 work programme under the supervision of the E.2 Cloud & Software unit. Among these, six projects were funded in the “Cloud Computing”[70] objective, focused on composition of resources across heterogeneous clouds, enablement of edge computing and automatic discovery and composition of cloud services, and five projects were funded in the follow-up “Towards a smart cloud computing continuum”[71] objective, which targeted complete solutions encompassing network, computing and data services at the core or the edge, advanced data privacy and security techniques, and novel programming models adapting services to different resources and usage contexts.

[70] EC. ICT-15-2019-2020: Cloud Computing. 2018
[71] EC. ICT-40-2020: Towards a smart cloud computing continuum. 2019

In addition, 11 software technology-focused projects also addressed topics relevant to cloud and edge computing, funded under the objective “Software Technologies” (five projects) in 2018[72] and “Software Technologies” (six projects) in 2020[73].
Based on project descriptions and outcomes, this portfolio of projects can be mapped to four key areas: adaptive cloud infrastructure and services[74], cloud-edge orchestration[75], DevOps[76] for cloud services and security & trust[77] for cloud infrastructure, services and data.

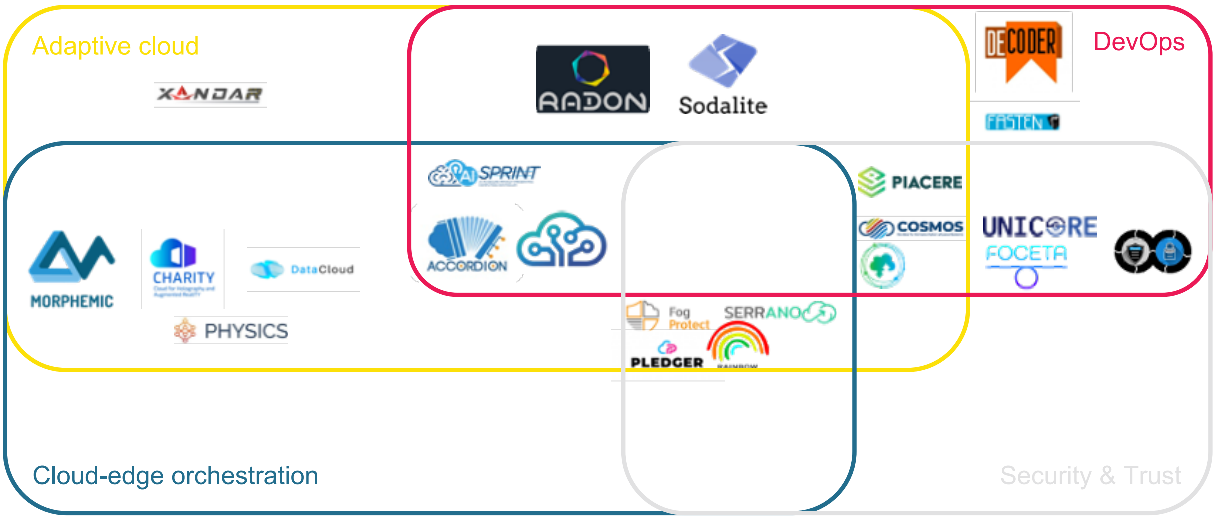


Figure 14: Macro topics for H2020 projects in E.2 Cloud & Software unit

As the figure shows, the projects, reflecting priorities identified in the work programme, focus primarily on cloud-edge orchestration. The second topic for number of projects, is self-adaptive cloud. A first analysis, shows that there are a variety of research topics that, following the recent evolution of technology toward edge computing, require more investigation (e.g. cloud energy optimisation, distributed data processing, delivery and business models for cloud-edge) and where the lack of available data constitute a limit (e.g. real time data on energy consumption for transport and computing in the cloud). However, these areas are generally well covered in the new Horizon Europe work programmes that promote the inclusion of cloud energy consumption as a first class topic and address the importance of new programming models to enable distributed data processing (under the name of “swarm computing programming”).
In terms of application areas, the projects cover a wide variety of domains, being healthcare, manufacturing, and mobility the most represented ones.

[72] EC. ICT-16-2018: Software Technologies. 2018
[73] EC. ICT-50-2020: Software Technologies. 2020
[74] Self-adaptive clouds are the application of autonomic computing principles to Cloud, i.e. the capacity of a cloud-based system to autonomously adapt its configuration and behavior based on the context (e.g. user’s behavior, energy available, …). In case the autonomous capacity is driven by Artificial Intelligence or Machine Learning, the term Cognitive Cloud is also used. See Endo et al. Autonomic Cloud Computing: Giving Intelligence to Simpleton Nodes. 2011.
[75] Cloud-edge orchestration research refers to the capacity of deploy, configure and manage resources and services both in the cloud data centers that in devices at the edge of the network. See Petri et al. Edge-Cloud Orchestration: Strategies for Service Placement and Enactment. 2019.
[76] DevOps is a set of practices that combines software development (Dev) and IT operations (Ops). The application of DevOps practices, originally linked to the put development and delivery of software (in cloud), is constantly growing, embracing new areas. Including, for example, Machine Learning algorithms, Security policies, ... See Ebert et al. DevOps. IEEE Software, vol. 33, no. 3, pp. 94-100, 2016.
[77] Security and trust are key concerns spanning across different layers of cloud computing architectures, from the lowest (e.g. hardware), across the transport segment (i.e. network and transport protocols), up to the layer where applications, services and data are accessed. For an overview of the topic, see Pearson, S.. Privacy, Security and Trust in Cloud Computing. In: Privacy and Security for Cloud Computing. Computer Communications and Networks. Springer, London. 2013.

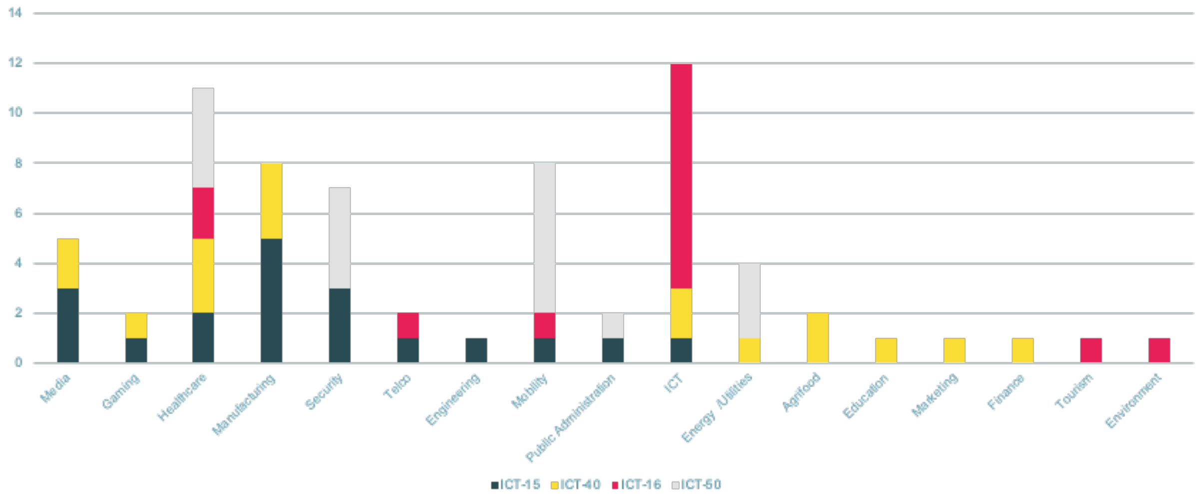


Figure 15: Application domain distribution across H2020 projects in E.2 Cloud & Software unit

3. DIGITAL TRANSFORMATION AND DIGITALIZATION

H-CLOUD's examination of cloud adoption in several sectors highlights the relationship between digital transformation, digitalization, and cloud adoption. Cloud adoption is often the fastest route to digitalization and the fastest means to enable digital transformation. Four general scenarios for digital transformation (DT) were found in almost all the sectors examined by H-CLOUD, and they can be seen to apply across the economy. (Section 3.1 describes these scenarios at a high level.) These are not the only DT scenarios possible -- but they illustrate recurring business requirements that in turn highlight gaps in both deployment and R&D.

All these DT scenarios involve two primary digitalization actions:

- *Becoming Data Aware*: Bringing input data from a variety of sources into the organisation. This data ranges from transactional and "system of record" data that already resides in the organisation, to operational data being generated and stored in the field or on the factory floor, to external data from government, research, and ecosystem partners. Data can be both static and real-time streaming data.
- *Becoming Data Driven*: Assembling, processing, and analysing all the available data to find new relationships and generate actionable insights, that can be used for better decision-making, to create beneficial outcomes and enable transformation.

Executing these two key actions requires both business/organisational readiness and a commitment to build or buy the infrastructures and capabilities needed to support the desired actions. In many cases the fastest route to digitalization relies on adoption of the cloud.

Actual transformation depends on the readiness of an organisation for such transformation, which in turn is a function of its culture. Major cloud suppliers, and major system integrators, recognize the need for culture change and agility for their customers to benefit fully from cloud adoption, and many suppliers (including Accenture, AWS, Capgemini, EY, Google, Microsoft, PWC and SAP) are actively engaged in exporting their own innovation cultures to clients by offering innovation workshops and similar change management services.

For example, improving operational efficiency in manufacturing requires:

- the manufacturer to collect real-time data from factory equipment and from its supply chain partners, to identify underutilised equipment, triggering changes to production scheduling and calls to suppliers to accelerate deliveries of out-of-stock parts, all of which requires
- real-time data collection from the factory floor, smooth communication and data exchange with suppliers, data analytics software that can be flexibly and securely executed in the manufacturer's data centre or in edge computing infrastructure, production scheduling systems that can respond in real-time to new circumstances, factory staff and procedures that can accommodate changes to plans, suppliers that can respond to requests for accelerated product delivery, and
- executive-level commitment to lead the organisation through the cultural changes needed to implement these new techniques and embrace a new way of working.

The manufacturer's adoption of cloud computing, either through direct use of public cloud service providers or the adoption of cloud-style IT processes, is necessary, but not sufficient, to enable the desired outcome of increased factory utilisation and faster product delivery. This section describes many other factors where action can be taken to expand digital transformation in the EU.

3.1. Common Scenarios for Digital Transformation

H-CLOUD's analysis found that organisations in different sectors are pursuing several common scenarios for digital transformation (DT):

- Improving Operational Efficiency and Sustainability, Reducing Risks.
- Engaging with Ecosystems of Related Organisations.
- Improving Engagement with Customers/Clients/Stakeholders, Increasing Revenues, Co-creation of New Products and Services.
- Improving Competitiveness by changing business models/ adopting new business roles.

We concentrate on these four scenarios since they were identified as the most common in H-CLOUD's demand side analysis.

These scenarios are described in more detail below.

1. **Improving Operational Efficiency and Sustainability, Reducing Risks.** There are numerous case studies illustrating the importance of digital transformation as a necessary enabler for organisational efficiency and risk reduction (including reduction of both cyber-risks and environmental impact). H-CLOUD found this driver to varying degrees in all the sectors examined. Specific examples:
    - *Public administration*: improved land use planning, data-driven improvements in public safety
    - *Transport/Mobility*: data-driven asset maintenance, digital parking space management, improving detection of and resilience against digital attacks, decarbonization while improving both operations and customer satisfaction, optimising pricing and payment mechanisms.
    - *Healthcare:* remote health monitoring and telemedicine, electronic health records
    - *Energy and utilities*: predictive operations and maintenance, decarbonization through integration, improved system integrity (e.g., leak detection)
    - *Agriculture*: precision farming, Agriculture 4.0
    - *Manufacturing*: Intelligent monitoring, optimization, and remediation; automation and orchestration of resources.

2. **Engaging with Ecosystems of Related Organisations.** In contrast with the driver above, which focuses on internal efficiency, more and more organisations in the studied sectors need to work across ecosystems of related organisations to achieve their objectives.

    For example, manufacturers have historically managed their supply chain ecosystems. Manufacturers are being driven into tighter alliances both with their upstream suppliers, and with both their direct customers and their ultimate users. Specific points of engagement include:

    - Data-driven integration of and collaboration with suppliers in strategic planning and execution
    - Collaborative condition monitoring
    - Shared production: cross-factory and cross-company production, orchestration of distributed manufacturing resources

    According to IDC, by 2022, 25% of manufacturers will be engaged in cross-industry collaboration, driven by rising customer expectations and competition from the platform economy, resulting in a 10% revenue increase. Among those collaborations, manufacturing companies are embracing new platform-enabled business opportunities through the creation of virtual buyer-and-seller communities, thus brokering interactions of makers and users with diverse but complementary interests. Cloud-based B2B platforms, such as International Data Space, BMW's Open Manufacturing Platform, Siemens Additive Manufacturing Network, Dassault Systèmes 3DEXPERIENCE Marketplace, fictiv, Krauss-Maffei Polymore, Ericsson Connected Vehicle Cloud, Volkswagen Automotive Cloud and Kloecner's XOM Materials marketplace, are flourishing across European industrial sectors.

This level of engagement is not limited to manufacturing, but in other sectors, ecosystem links tend to be more horizontal than vertical, and less "proprietary":

- *Public administration*: coordinating with other levels of government and with neighbouring jurisdictions for seamless services to citizens
- *Transport/Mobility*: multi-modal passenger and logistics services (responding to the EU's "Sustainable and Smart Mobility Strategy"[78] particularly the federated system proposed by Digital Transport and Logistics Forum (sub group 2)).
- *Energy*: Operational integration with other energy ecosystem players including nearby energy providers (responding to "An EU Strategy for Energy System Integration"[79]), distributed energy management.
- *Healthcare*: improving communication, collaboration and coordination among providers to improve patient health outcomes
- *SMEs*: Ecosystem links, engagement and coordination are critical to all SMEs, but particularly to "traded" enterprises, as defined by Michael Porter[80], which have customers outside their local region, which historically have a greater impact on competitiveness than non-traded enterprises.
- *Agriculture*: value chain coordination, collaborative efforts, e.g. through farm cooperatives (responding to the EU's "Farm-to-Fork Agricultural Strategy"[81]).

3. **Improving Engagement with Customers/Clients/Stakeholders, Increasing Revenues, Co-creation of New Products and Services.** We group these scenarios together into a single scenario because they are tightly linked and hard to consider separately. Collectively, this is a key driver of many digital transformation efforts, present across all the sectors studied. Specific examples:

   - *Public administration*: data-driven improvement in citizen experience, openness and transparency; creating new and more inclusive citizen service channels, as well as increasing engagement and participation in governance. The 2017 Ministerial Declaration on eGovernment (Tallinn Declaration) committed to creating "high quality, user-centric digital public services for citizens as well as seamless cross-border public services for businesses".
   - *Transport/Mobility*: deploying new services channels, offering multi-modal purchase capabilities to customers
   - *Energy*: Providing tools for customers to manage their energy usage and optimise their choice of energy supplier.

4. **Increasing Competitiveness by Changing business models/ adopting new business roles.** True transformation entails new business models and/or the transformation of the organisation's role in its sector, e.g. shifting from manufacturer to service provider to orchestrator. While many such transformations have occurred without the benefit of digitalization or cloud computing, they make such transformations easier and more likely to succeed. Examples of such transformations include:

[78] EC. Sustainable and Smart Mobility Strategy. 2020
[79] EC. Powering a climate-neutral economy: An EU Strategy for Energy System Integration. 2020
[80] Porter, M.E. The Economic Performance of Regions. Regional Studies, 37, pp. 549-578. 2003
[81] EC. EC, Farm to Fork Strategy, 2020

- *Transport:* Vehicle manufacturers become service providers, new entrants provide orchestration services, all private players see themselves as being in the data business (viewing data as a competitive differentiator and an asset to be monetized)
- *Agriculture:* Orchestrators restructure the traditional chain of relationships between producers, wholesalers, distributors and retailers to optimise the flow of food products
- *Energy/Utilities:* Emergence of resellers, aggregators, orchestrators, supply flexibility services
- *Manufacturing:* Manufacturers shift from making products to selling the services produced by those products (e.g. jet engines are transformed into propulsion services).

3.2. Becoming Data Aware and Data Driven

Each of these transformation scenarios, regardless of sector or organisation, requires organisations to first become data aware and then to become data driven:

Becoming Data Aware: Capture and access new and expanded data. This includes internal data, which increasingly must be collected from remote or distributed facilities through edge infrastructure and Internet-of-Things devices. This increasingly also includes sensitive data, for example data about customers or stakeholders, that must be handled carefully. This also includes data from other organisations, through business ecosystems that depend on data exchange and sharing to create value for ecosystem partners. As organisations transform to become digitally enabled and agile, these new collections of data will increasingly include real time data, collected from operations or from online customer interactions, and processed in real time to trigger real time improvements in efficiency and customer engagement.

Challenges to becoming data aware include:

- Managing sensitive data, including personal data,
- Adopting edge and IoT-based technologies to collect data from distributed sources and facilities,
- Participating in business ecosystems to access data from, and share date with, ecosystem partners in a secure and trustworthy manner.

In all cases, organisations must balance the potential benefits of accessing and storing “more” data with the costs, risks and responsibilities of collecting, holding and retaining this data.

Recommendations to support EU organisations becoming more data aware are presented in section 4.

Becoming Data Driven: Understand the Data, Predict and Optimise. Use data analytics tools to explore and analyse gathered data to discover actionable relationships. Use these relationships to predict performance and optimise processes. These activities often start with the creation of “single versions of the truth” which collect all relevant data about a process or customer so that they can be analysed together.

Becoming more data-driven requires the adoption of agile cloud-style IT technologies and requires organisations themselves to become more agile. Recommendations to support EU organisations adopt cloud-style IT technologies are presented in section 5.

These two “phases” of digitalization can be explored to identify challenges and opportunities for improvement across the EU.

4. BECOMING MORE DATA AWARE: RECOMMENDATIONS TO IMPROVE DATA ACCESS AND COLLECTION

Here we identify a variety of challenges to the first digitalization step described above and offer recommendations for actions to address those challenges. Many of these challenges are being tackled by the efforts of Gaia-X, IDSA, BDVA, FIWARE and others. These efforts are promising, creating many useful approaches, such as

- IDSA's Reference Architecture (RAM) v3.0[82]
- FIWARE's data models in several sectors[83]
- OASC's Minimum Interoperability Mechanisms (MIM)[84]
- OpenDEI's Design Principles for Data Spaces[85].

These approaches build the foundation for the complete range of solutions that will be required to enable the level of "data awareness" that is needed.

4.1. Tools and Processes for Respecting and Maintaining Data Sovereignty and Privacy Rights

A key step in becoming "data aware" is to "capture internal data", and this increasingly involves the capture, storage and manipulation of sensitive data, particularly personal data. The General Data Protection Regulation[86] specifies how organisations should handle personal data in order to respect and maintain the data sovereignty of individual data subjects, placing significant legal and operational obligations on, as well as introducing significant risks to, organisations working with personal data. These obligations and risks make it more difficult for organisations to embrace digitalization. Gaia-X is explicitly focussed on meeting such requirements, since they will clearly enable a range of important business cases across multiple sectors[87].

This type of requirement is not limited to personal data: for example, the EU Code of Conduct on Agricultural Data Sharing[88] defines the rights of farmers, and of other stakeholders in agricultural value chains, to control how their data is used. The Nagoya Protocol[89] places restrictions on data about endangered species. The CARE protocol[90] recommends processes to protect data about indigenous people and culture. Confidential data, trade secrets and intellectual property have historically been categories as sensitive data that needs special treatment. The recently released EU Data Act recognizes this and similar rights by a number of protected subjects, as well as recognizing that certain data may refer to multiple protected subjects, requiring mechanisms to respect shared sovereignty rights. Despite the extent of these requirements, tools and processes to implement such controls are not mature.

These rights are clear and fundamental to the societal values of the European Union. Nevertheless the processes required are complicated and evolving to reflect expanded definitions of protected subjects. Technical solutions are not mature enough for confident

---

[82] https://internationaldataspaces.org/download/16630/
[83] https://www.fiware.org/smart-data-models/
[84] https://geoportais.pt/minimal-interoperability-mechanisms
[85] https://design-principles-for-data-spaces.org/
[86] EC. General Data Protection Regulation, 2016
[87] Gaia-X. Data Spaces Business Committee - Position Paper, August 2021.
[88] EU Code of conduct on agricultural data sharing. 2018
[89] United Nations. The Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity. 2011
[90] Global Indigenous Data Alliance. CARE Principles for Indigenous Data Governance. 2019

adoption by EU organisations. The EC, through its Horizon Europe work programme[91], has highlighted the importance of this challenge and identified the need to support additional research and innovation in this area.

The H-CLOUD team identified this challenge[92] in multiple demand sectors, including public administration, healthcare, transportation, and agriculture, as well as in its supply side analysis.

4.2. Mature Technologies, Standards and Technical Ecosystems for Collecting and Processing Data at the "Edge"

Significant amounts of data have always been generated at the "edge" of the network, since the devices that generate much of this data (e.g. sensors, mobile phones, personal computers) are located at the edge. The conception of edge computing can be traced back to the 1990s, when Akamai introduced a content delivery network (CDN) service to reduce latency of content distribution to end users[93]. With the rise of the Internet-of-Things (IoT) and mobile internet, increasing amounts of data are generated at "edge" of the network - where a network connects with specific devices, such as smartphones, wearable devices, and IoT devices - and then stored and processed in the cloud. Evolving technical and policy requirements, such as latency, environmental impacts, data sovereignty, and accelerating data growth, suggest that more and more data generated at the edge should also be processed and stored there. This is the concept of edge computing: a distributed computing paradigm that brings computation and data storage capacities closer to where data is generated. Modern edge computing solutions support self-provisioning of services, and API-based service management and monitoring. The trend is accelerated by new generations of embedded devices that are now capable of advanced processing tasks such as machine learning. As IoT and Industrial IoT devices become more reliable, less expensive and easier to connect, the new challenge becomes the growing flow of data. Edge computing provides an alternative to transferring all this data to a central data centre or to the cloud, allowing optimization of data transfer, storage and processing, reductions in network latency and increased responsiveness.

Thanks to this, edge computing is becoming a key ingredient to bridge the gap between Information Technology (IT) and Operational Technology (OT), thus becoming a key enabler for digitising OT solutions. According to IDC[94], by 2023 73% of the data will be created outside the core, i.e., outside a traditional data centre or the cloud. IDC[95] highlights that this trend will be driven by IoT: "By 2022, 50% of the initial analysis of IoT data will occur at the Edge".

The EC has recognized the important role of edge computing as a digitalization enabler and refers to it both in the context of EU data spaces and EU cloud federations[96] and in the 2030 Digital Compass[97]. The Digital Compass specifically defines an objective that 10,000 climate-neutral highly-secure edge nodes should be created by 2030. This vision is also echoed by the recently established European Alliance for Industrial Cloud Data Edge and in the roadmap[98] presented by a group of EU CEOs in the area of digital technologies.

As of today, European companies are spending less than 10% of their ICT infrastructure budget on edge capacities, and, as in the case of cloud technologies, SMEs' adoption is far behind[99]. In term of global deployments, the research of Linux Foundation[100], highlights how

[91] EC. HORIZON-CL4-2021-DATA-01-01: Technologies and solutions for compliance, privacy preservation, green and responsible data operations. 2021
[92] Originally this was labelled "Major Challenge M1" in H-CLOUD "D3.1 – Strategy Analysis Report and Cloud Computing v1.0".
[93] BOSH. Cloud and edge computing for IoT: a short history. 2021
[94] IDC. Revelations in the 2020 Global DataSphere (idc.com). 2020
[95] IDC. European FutureScape. 2018
[96] EC. A European strategy for data. 2020
[97] EC. 2030 Digital Compass: the European way for the Digital Decade. 2021
[98] AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. 2021
[99] IDC, European Tech and Industry Pulse Survey 2019-2020
[100] Linux Foundation. State of the edge. 2021

edge deployment is expanding in all global regions. By 2028, Asia is forecasted to have 37.7% of the global edge infrastructure, Europe 29% (with a major role by Germany, France and the United Kingdom), North America 20.5%, Latin America 7% and the remaining 5.8% will be deployed in the Middle East and African regions (see Figure 16).

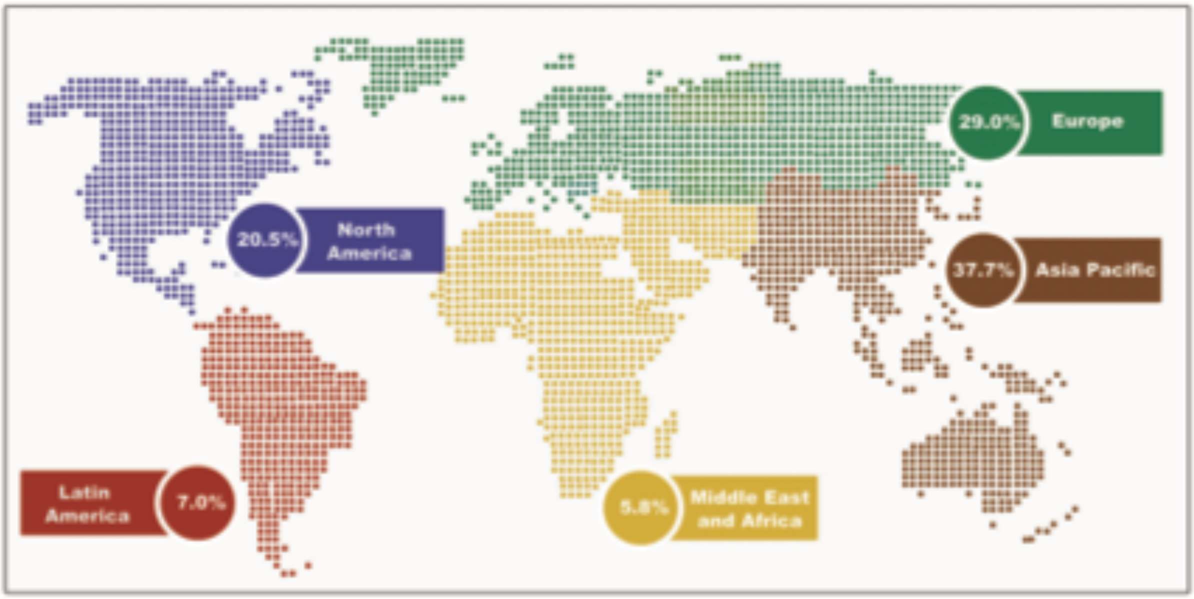


Figure 16: Regional distribution of edge deployments[101]

While broad adoption of edge computing faces several challenges, solving these challenges will enable richer and more timely data capture, and potentially digital transformation, in almost all the sectors analysed by H-CLOUD:

- Public Administration: government infrastructure, notably at the urban level
- Transport: vehicles and rolling stock, roads and rail lines, ports, airports
- Energy: electricity and fossil fuel generation, processing, storage, transmission, and distribution
- Healthcare: connecting isolated medical devices and monitoring equipment throughout hospitals and clinics
- Manufacturing: discrete and continuous manufacturing equipment, plus the products themselves (smart products, tagging of commodity products and packaging).
- Agriculture: tagging agricultural products and packaging (e.g. livestock and food safety tracking and surveillance), instrumenting agricultural production systems (field monitoring, farm implements, milking machines, etc.).

Edge computing deployment requires investment in several related areas:

- hardware (nodes capable of computing at the edge),
- connectivity (network infrastructure connecting the nodes and eventually the cloud),
- software (particularly for optimising data flows and the placement of processing workloads against latency, sustainability and cost factors, and orchestration of complex distributed data processing), and
- skills (competencies across the new hardware and software stack as well as total end-to-end architecture).

[101] Linux Foundation. State of the edge. 2021

While in certain high-value use cases, such as in manufacturing, organisations are investing in edge computing with the expectation of a quick return on investment (ROI), this is not the case for all sectors. For example, smart cities and smart mobility are exploring edge technologies, supported by a strong innovation culture, but adoption at scale is still limited, beyond initial pilots. This may be also related to the fact that network effects are needed to generate an adequate ROI. Network effects require use of common solutions and/or the interoperability of different technologies, but the market potential of edge makes vendors hesitant to standardise or enable interoperability. The landscape of edge technologies is also evolving very rapidly, both at the hardware and software level; in this market a major role is still taken by large cloud providers that want to entrench their own solutions, by moving data from the client's edge to their cloud, i.e., delaying the growth of edge computing as a central data processing and storage location. Moreover, it is also unclear how existing solutions can enable a widely distributed infrastructure to be as cost effective as today's cloud computing solutions. Today, most stakeholders hesitate to invest in edge technology because it is not seen as either mature or consolidated enough, limiting broader adoption, and leaving investment to less risk-averse pioneers who can see how early investments might create competitive advantage.

4.3. Support for Data Sharing and Data Ecosystems

A significant component of "becoming data aware" involves accessing data from partners. This access depends on those partners' willingness to share their data -- and their ability to share that data in a secure and trustworthy way. Supports are needed both to encourage, and to enable, secure, trustworthy data sharing.

Data sharing is a critical aspect of a growing organisational opportunity: participating in business ecosystems. Business ecosystems are increasingly central to modern business and competitive economies, from supply chain networks to multi-side platforms. The creation and strengthening of business ecosystems are central to the European Union's recent strategies for the green deal[102], mobility and transport[103], energy[104] and agriculture[105]. Enabling data sharing through data ecosystems and data spaces is a central objective of Gaia-X.

Business ecosystems build on the exchange of data across the ecosystem and the creation of corresponding Data Exchange Approaches[106] that enable business ecosystems to function. organisations need the ability to access, share and process data across organisational boundaries[107]. organisations from many sectors need secure data access and sharing capabilities to enable their businesses to grow, rather than just complying with regulations. This business growth (or mission effectiveness for public good organisations such as healthcare providers) requires managing data that is distributed across organisational boundaries. Data governance requires coordinated approaches among the different stakeholders to ensure tracking and verification of the ways data are used and processed. Solutions supporting such scenarios need to be robust and affordable in order to encourage wide adoption.

This challenge is discussed in five demand scenarios (transport, *D-T Challenge 1*; energy, *D-E Challenge 4*; public administration, *D-PA Challenge 8;* agriculture, *D-A Challenge 4*; manufacturing, *D-M Challenge 3*; and healthcare, *D-H Challenge 1 & 2*) as well as in one supply side analysis (technology landscape, *S-T Challenge 3*).

[102] EC. Stepping up Europe's 2030 climate ambition. 2020
[103] EC. Sustainable and Smart Mobility Strategy. 2020
[104] EC. Powering a climate-neutral economy: An EU Strategy for Energy System Integration. 2020
[105] EC. EC, Farm to Fork Strategy, 2020
[106] "Data Exchange Approach" is proposed as a generic structure that encompasses "data ecosystems," federated data infrastructures and "data spaces". See EGI Working Document on Common Approaches to Data Exchange for definitions of both data ecosystems and data spaces (forthcoming).
[107] Originally Major Challenge M3 from the H-CLOUD "D3.1 – Strategy Analysis Report and Cloud Computing v1.0".

This challenge is implicit in the EUSD's planned creation of common European data spaces, both cross-sector and in nine specific sectors, intended to foster data-driven innovation across Europe[108]. The need is very clear in each of the sectors identified, but solutions, both technological and organisational, must become more mature to address this challenge. Technical and organisational interoperability is needed not only within individual business ecosystems, but also between specific ecosystems to reduce barriers between them and avoid the creation of new silos.
Addressing this major challenge requires action in six different areas:

1. Create Model Value Propositions for Data Sharing and encourage Participation in Broader Data Exchange Initiatives.
2. Support Assessment of the "Shareability" of Data; Harmonize and Clarify Relevant Legislation and Regulation
3. Support the Preparation of Data for Sharing
4. Develop Reusable Models for Governance of Data Exchange Initiatives
5. Create Secure and Trusted IT Solutions for Data Exchange Initiatives.
6. Develop Tools and Platforms that Enable Secure Sharing of Large or Very Sensitive Data

4.3.1. Create Model Value Propositions and Usage Conditions for Data Sharing; Encourage Participation in Broader Data Sharing Initiatives

Organisations already have significant amounts of data that could be used to solve many problems[109]. However even organisations that are prepared to share their data still need a clear idea of how this will create value for them. Difficulty assessing the value of sharing data not only hampers data sharing but also makes organisations hesitate even to participate in data ecosystems. Gaia-X and other initiatives are working both to make data sharing feasible and to operationalise value propositions for data sharing to encourage that sharing.

For example, in the agricultural sector, farmers remain sceptical of the return on investment (ROI) for data sharing, which is necessary to enable the adoption of precision agriculture (an important digital transformation scenario for agriculture). For larger farms, and the broader "farm to fork" value chain, business models and ROI do not support complex new solutions. For the smallest farms, of which there are millions across Europe, embracing precision agriculture is difficult given limited resources for this kind of activity. (See D-A Challenge 1: Value Proposition for Farmers for Precision Agriculture and Data Sharing.)

For SMEs, defining value propositions is especially critical since SMEs see their data as a competitive advantage. Usage-rights must be defined to ensure fair use and eliminate any possible disadvantages for SMEs.

In general there are several types of data sharing value proposition:

- Exchange to enable indirect benefits: Data may be shared with others as a contribution to broader efforts by ecosystem partners to create shared value, for example identifying trends and insights in patient data to enable better health outcomes. Data is shared without the expectation of returns specifically for the benefit of the data holder. The value proposition can be defined by developing a common understanding among ecosystem participants regarding the purpose of the ecosystem.
- Exchange to enable specific benefits: Data can be shared with ecosystem partners to enable the creation of knowledge or insights that might possibly benefit the data holder. For example, the compilation of operational data on components installed by many industrial customers of a component-manufacturer can enable the component

[108] EC. A European strategy for data. 2020
[109] EC. Experts say privately held data available in the European Union should be used better and more. 2021

manufacturer to perform data analysis or machine learning on that data. The broader the sharing, the better the insights (e.g., through reduced bias). This might yield recommendations for improved component operation by the current component owners, or it may benefit the component manufacturer by enabling value-added services, which can be offered to those same component owners, or by contributing to improved designs for future components. This is a more common type of value proposition for data sharing, but work is needed to capture these relationships and set reasonable expectations, in advance, among participants. Templates and models are needed to guide the development of these types of value proposition.

- Monetization: the creation of a commercial transaction involving the data, generating revenue for the data holder in exchange for the data consumer being able to access the data. Here the value proposition is well-defined, but data is seldom monetized today. Realistic expectations about data "market values", as well as the creation of more data "marketplaces" for the trading of data, would help define the monetization value proposition.

Possible value propositions for data sharing are tightly linked to the conditions on the use of that data specified by the data holder.

When participants do see the value in sharing data, overlapping or competitive business ecosystems can develop, creating possibly "lock-in" and obstacles to participation in wider data sharing initiatives, such as the data spaces proposed by the EC. Again, looking at the agricultural sector, the EC is perceived by some ecosystem participants as creating a new platform that, if not "competitive", at least disrupts their current business plans. Operators of existing ecosystems/platforms are concerned about how an agricultural data space would operate, how it would interrelate with those existing ecosystems/platforms, and how it would create value for both existing ecosystems and their participants. (See D-A Challenge 3: Value Proposition for Service Providers/FMS Vendors in the context of a Common European Agricultural Data Space).

4.3.2. Support Assessment of the "Shareability" of Data, Harmonise and Clarify Relevant Legislation and Regulation

Before data can be shared it must be assessed to ensure it does not contain sensitive data (including data about individuals or other protected data subjects), confidential data, or data which the data holder may not have the right to share (e.g. textual material or images covered by copyrights held by others). Guidelines and tools should be developed and made easily available to EU-based organisations to assist them in the assessment of their data for potential sharing.

This assessment is complicated in different sectors by conflicting, overlapping or vague legislation or regulation, particularly across Member States. Two recommendations on this aspect were identified in H-CLOUD's sectoral analysis:

- **Healthcare:** National differences in the regulation of aggregated analysis of sensitive medical data have been recognised as a problem by the EU Parliament. The EU should take legislative action to address these issues, perhaps through the actions contemplated by the EUSD.
- **Agriculture:** Sovereignty and confidentiality for farm-based data. Although the current Code of Conduct for Agricultural Data takes a self-regulatory approach and is linked to the need for clear and balanced contractual arrangements, the EU is encouraged to formalise these rights in legislation.

4.3.3. Support the Preparation of Data for Sharing, including Data Interoperability

When the value proposition is clear enough for a data holder to consider sharing, data holders must still prepare that data for sharing. Effort is needed in several areas, where tools, frameworks and guidance could facilitate increased data sharing:

- Tools to "sanitise" sensitive data, for example through anonymization of personal data, or data with uncertain IP status, for example using automated scanning utilities. This is not a trivial problem, since most data today is not properly annotated with metadata to identify such issues.
- Guidance on, and models for, the "terms of use" for the data, i.e. the usage conditions to be agreed by a prospective data consumer. These can range from simple agreement to a standard licence (e.g. "CC-BY"), to payment schemes, or more specific restrictions on use (e.g. not sharing certain data with competitors, restricting research data to non-commercial use). Ideally software that "consumes" data (e.g. ML training algorithms) should also "consume" and embed any licence restrictions on source data into its output[110]
- Utilities to aid packaging data for sharing, including reformatting into shareable file formats (e.g. JSON-LD, XML, PNG), assignment of identifiers, combination with standardised metadata to assist in finding and using the data, reformatting to align with standard data and metadata formats (appropriate to the domain), copying the data to externally accessible storage infrastructure (operated by the data holder or an external data repository), and posting metadata so that ecosystem partners can find and potentially access the data being offered.
- Guidance and best practises for processes to manage requests to access offered data, negotiate and record terms of use, track actual exchanges, monitor usage and scan for inappropriate use.
- A single repository or marketplace where these tools, frameworks and guidance can be found. The diversity of data types, licensing and terms of use represent barriers to data sharing and an easy-to-use source of solutions in this area would contribute to increased data sharing.

Many of the same tools are needed to prepare open public data for use in digital transformation efforts and for inclusion in the proposed pan-European data spaces. This challenge was specifically identified in the agriculture sector (see D-A Challenge 5: Accessibility and Interoperability of data across the data space).

Effective data sharing requires harmonising data definitions and metadata so that the data can be discovered, accessed and shared as appropriate for meaningful analysis. H-CLOUD identified this challenge in almost every sector, with specific discussion in the transport sector. The term "interoperability" can refer to different kinds of interoperability, ranging from technical interoperability at the lowest level, to syntactic, semantic, organisational and legal interoperability at successively higher levels. Interoperability at a lower level may not allow meaningful interoperation when higher levels of interoperability are required by the intended use case. Potentially useful data (e.g. cumulative distance travelled for vehicles, collected for fleet maintenance) may not actually be usable by other ecosystem partners (e.g. utilities seeking to offer timely and affordable electric vehicle charging services). Higher levels of interoperability may only be possible in more limited situations or specific use cases, such as shipping and logistics or human genetics.

Given this challenge, H-CLOUD recommends the creation of a clearinghouse and master taxonomy service to help organisations that are preparing data for sharing (as well as for better internal exploitation) to identify and apply the most appropriate and meaningful metadata

[110] Peng, K., Mathur, A., Narayanan, Arvind. Mitigating Dataset Harms Requires Stewardship: Lessons from 1000 Papers. 2021.

standards. Supports should be available to assist organisations prepared to share their data to increase the semantic interoperability of that data within and across sectors.

4.3.4. Develop Reusable Models for the Governance of Data Exchange Approaches

H-CLOUD's analysis of data exchange approaches[102] identifies the need for both organisational policies and roles, as well as technical specifications, to enable the trustworthy exchange of data between two or more parties. This recommendation addresses the first requirement through development of reusable governance models for data exchange initiatives. The EC itself, in its European Strategy for Data[111], highlights the need to create governance frameworks for the proposed pan-European data spaces, and the draft Work Programme for the Digital Europe Programme includes several planned calls for funding that include activities targeting the creation of governance frameworks. The OpenDEI project has also identified the need for "soft infrastructure" as an essential building block for any data sharing initiative.[112]

The importance of governance is highlighted by the experience of Public Administrations in establishing, but often failing to sustain, various smart city projects across Europe (See D-PA Challenge 13: "Many smart cities programs have failed to scale beyond pilot projects because they encountered governance, technical and regulatory challenges"). The result of those investments was often a plethora of fragmented pilot projects that did not scale from a corridor or neighbourhood to the entire city. Or segments of the resident population were excluded from the intended benefits. The cities that succeeded in orchestrating sustainable ecosystems took a realistic approach to funding and budgeting. They set up programs to make sure that all residents were included in the benefits. They delivered quick wins in specific use cases, and then re-use the modular solutions they had built to extend the capabilities across the whole community. These experiences represent good practises that should be identified, catalogued and re-used.

Data exchange initiatives are examples of federated infrastructure. H-CLOUD's analysis of federation[113] identifies challenges for the governance of federated infrastructure (see S-F Challenge 1: "Coordinated/federated approaches must be structured around the objectives of their stakeholders, balancing community focussed initiatives with pan-European solutions", as well as S-F Challenge 3: "Federated data has great potential to support secure, private sharing of data held by many different organisations") and provides a number of recommendations for developing robust federated governance frameworks:

- Community: Define the stakeholders and community to be served by the data exchange initiative. This should distinguish data providers, data consumers, and other possible roles within the initiative, and identify any criteria for taking on each role.
- Visibility: Identify which stakeholders and community members will be allowed to "see" metadata about the data within a data exchange initiative, as well as those that will be allowed to access and use the data itself, once related usage conditions have been agreed.
- Purpose: Define the common purpose to which the whole community is committed. It is important to avoid mismatches in expectations between, e.g., providers and consumers. Identify scenarios for access, use and benefit, ideally in the form of detailed business cases that itemize participants' gains, costs and risks, broader community gains, costs and risks.

[111] EC. A European strategy for data. 2020
[112] OpenDEI. Design Principles for Data Spaces. 2021
[113] H-CLOUD. Supply Analysis: Cloud Federation (S-F). 2021

- Value Propositions and Usage Conditions: Develop value propositions for the data exchange initiative and for each type of participant (distinguishing value propositions for different roles), along with the related conditions for use. The effectiveness of any initiative will depend strongly on the clarity of its value proposition and how it is constituted to realize that value proposition.
- Governance: Define the business and operating model for the initiative as a whole -- including all the business rules and functions necessary to deliver the expected value propositions.
- Data Sovereignty and Interactions of Different Data Exchange Initiatives: Possible interactions of overlapping or intersecting data exchange initiatives must be considered when designing a data exchange initiative. For example, how would an existing agricultural data ecosystem interact with a new “pan-European” data space, as proposed by the EC? Recognizing that the principle of data sovereignty applies to a data holder’s decision about the data exchange initiatives in which it will participate, how can the needs of different stakeholder communities be balanced against the need for broader solutions?
- Oversight Mechanisms: What mechanisms are needed to ensure compliance with the many components of a data exchange initiative’s governance model? Are there mechanisms for recording transactions, auditing the actions of participants, investigating possible infringements of agreed policies, enforcing penalties for violations? Will the community “self-regulate,” or do stakeholders need recourse to an external evaluation, audit and enforcement capability?

4.3.5. Create Secure and Trusted IT Solutions for Data Exchange Initiatives

Once data is ready to be shared, business ecosystems need a reliable, trustworthy “operating system” to support the activities of the ecosystem. Required functions range from membership management and catalogue/marketplace functions to secure data exchange and tools that can ensure that data consumers use the data they receive in compliance with the agreed usage conditions.

Various custom solutions have been implemented by specific business ecosystems, but it is unclear if these solutions can be extended to reliably meet the technical requirements of various business ecosystems, while also flexibly supporting the different governance and operating models selected by each ecosystem’s stakeholders and community members.

The experience of Public Administrations is again relevant, where selected technology solutions were not re-usable across cities, did not allow for efficient cross-border best practice exchange and did not enable tech suppliers to generate solid revenue growth to sustain the business, much less enabling re-investment in innovation. (See D-PA Challenge 13: “Many smart cities programs have failed to scale beyond pilot projects because they encountered governance, technical and regulatory challenges”.)

The need for secure and trustworthy solutions to support data exchange initiatives, that could also accommodate different governance structures as well as common technical and semantic interoperability, was found in the public administration (see D-PA Recommendation 13), transport (see D-T Recommendation 1), agriculture, and SME sectors (see D-S Recommendation 1.4). SMEs are seeking to enable sharing with public administrations and with business ecosystems that extend beyond single sectors, so inter-ecosystem exchange will need to be carefully defined and supported.

Note that many existing IT platforms for business ecosystems have been built on digital ledger technologies (DLT or "blockchain") to enable secure and controlled data exchange[114], so this approach should be formally evaluated and tested alongside other privacy-preserving technologies.

4.3.6. Develop Tools and Platforms that Enable Secure Sharing of Large or Very Sensitive Data

In some cases, the data to be exchanged is very large, making data transfer time-consuming. In other cases, the data holder is only prepared to give access to the data using the data holder's own infrastructure, for reasons of either size or security. Both situations require a "bring the compute to the data" solution (vs. the more common "bring the data to the compute" approach), which introduces several additional requirements.

H-CLOUD identified this requirement in healthcare and in agriculture:

- In healthcare (see D-H Challenge 5) data volumes are magnified by clinical imaging as well genomics. Initiatives to aggregate and process large healthcare datasets (in both national healthcare systems as well as for the common European health data space proposed by the EUSD[115]) highlight difficulties in efficiently accessing, sharing, and analysing distributed healthcare data while maintaining the protection, security and privacy of that data and avoiding unnecessary data movement and duplication.
- Agriculture relies significantly on earth-sensing data gathered by satellites such as the European Copernicus fleet (see D-A Challenge 6: "High performance, *in situ,* analysis of distributed big data"). In addition, the primary users of this data, namely the farmers themselves, must rely on remote processing services to take advantage of this data.

These challenges are not unique to these two sectors.

[114] Woods, J, Iyengar, R.. *Enterprise Blockchain Has Arrived: Real Deployments. Real Value.* 2019. ISBN: 978-1-753308-0-9.
[115] EC. A European strategy for data. 2020

5. BECOMING MORE DATA-DRIVEN: ENABLING ADOPTION OF CLOUD-STYLE IT METHODOLOGIES AND CLOUD INFRASTRUCTURE

Organisations can become more data driven first by adopting cloud-style information technology (IT) methodologies, and then deploying those technologies onto suitable cloud infrastructure[116].

- Many of the benefits of being more data driven flow from the first step and parallel the benefits from digital transformation: increased IT agility, faster time to solution, uncoupling functional improvement from limits on underlying infrastructure, etc.
- Once an organisation adopts cloud-style IT methodologies, it then has the flexibility to consider different forms of cloud deployment, including hybrid and "multi-cloud" combinations: use of EU cloud providers, use of foreign hyperscalers, use of distributed edge compute infrastructure, use of on-premises cloud-style infrastructure.

Separating these two steps is essential for organisations to maintain control and sovereignty over their digital infrastructure -- ensuring that digital activities reflect the priorities of the organisation rather than the latest marketing efforts of suppliers, avoiding lock-in at any level, and enabling the use of the most cost-effective cloud infrastructure.

H-CLOUD Technology analysis[117] provides a snapshot of cloud-style technologies and the related "stack" of technologies and explores the challenges faced by organisations in their journey to the cloud. Cloud-style IT starts with the adoption of cloud-native technologies and software, such as virtualization and containers, and related technology stacks, but also includes several other dimensions:

- the explicit design of security, privacy and reliability capabilities,
- new software development skills, approaches, and deployment methods,
- new perspectives on "build vs. buy" questions, procurement, and financial management, and
- new ways to manage legacy applications.

Adopting a cloud-style approach to information technology usually requires significant changes for established IT organisations, as well as for small- and medium-sized enterprises (SMEs).

The European Commissions and EU-based cloud service providers can take several complementary actions to support EU organisations becoming more data driven.

5.1. Adopting Cloud-Style Technologies and Technology Stack

As described in H-CLOUD's Technology Analysis[118], the "cloud" is based on a fundamental separation of software from the infrastructure on which it runs. In the past, in-house software developers would design their software to take advantage of the characteristics of the systems that were installed and running in their organisations' on-premises data centres. Changes to those systems, such as updates to operating systems or disk storage hierarchies, would sometimes trigger the need for changes to running software, since that software had been designed with dependencies on operating system services, file storage arrangements or other infrastructure details.
Moving to cloud-style technologies abstracts these details through "virtualization", "containerization" as well as the recent development of "serverless computing":

[116] Andreessen Horowitz. The Cost of Cloud, a Trillion Dollar Paradox. 2021
[117] H-CLOUD. Supply Analysis: Technology (S-T). 2021
[118] H-CLOUD. Supply Analysis: Technology (S-T). 2021

- Virtualization logically packages physical compute resources into virtual machines that can be assigned and reassigned to software jobs as required. OpenStack and Open Nebula are two open-source systems, and VMware is a commercially available system, all of which can be run on physical hardware to provide virtual machine services. (The functional scopes of OpenStack, Open Nebula, and VMware are different, and all three systems have been complemented by a range of other offerings including commercial complements to OpenStack and Open Nebula, and open-source complements to VMware.) Virtualization is usually associated with infrastructure-as-a-service (IaaS) cloud computing, although the underlying physical hardware might also often be made available to customers through “bare metal” service offerings.
- Containerization involves grouping a software program itself with the libraries and subroutines it requires into a self-contained file or “container” that can be loaded onto physical or virtual compute resources to run the program. Docker is a popular form of open-source container, but other container implementations are possible including Container, LXC, Podman, Singularity and Udocker. Kubernetes is another popular open-source tool associated with software containers, although Kubernetes is actually used to orchestrate, rather than directly implement, containers. In the same way, several companies offer platforms that build on top of Kubernetes and Docker adding additional high-level functionalities or managed platforms, like OpenShift and Rancher and various integrated implementations of Kubernetes created by certain commercial cloud providers (Amazon EKS, Azure AKS, Google GKE, Open Telekom Cloud CCE, SAP Gardener).
- In serverless computing, the cloud provider allocates resources on demand in response to granular requests for service, taking care of the servers on behalf of its customers. Operation is like a function call or subroutine in programming, and serverless is sometimes referred to as “function-as-a-service” (FaaS). Serverless offerings include AWS Lambda, Google App Engine, Google Cloud Functions, and Microsoft Azure Functions.

These abstractions free developers from the constraints of physical infrastructure, allowing demanding processes to be flexibly supported with on demand compute and storage capacity as required.

Although virtual machines and containers introduce their own dependencies (for example, the choice of container implementation), these cloud-based approaches are more standardised than the variety of infrastructure that might be found in on-premises data centres. This higher level of hardware abstraction, combined with standardisation of the software stack within organisations and increased use of automation for “DevOps”, enables developers to focus on the application rather than deployment, reducing the need for internal data centre operations teams and the overhead of physical machine management. This makes many cloud-based development skills highly transferable and, once those skills are in place, also makes development easier and more productive.

As organisations expand their activities to include big data analytics, machine learning and artificial intelligence, as well as edge-based data and computing, in their digital transformation efforts, adopting cloud-style methodologies will be essential to working with and integrating these capabilities.

Although the benefits of taking this step are clear, there are also several risks and challenges to be managed:

- The cloud technology landscape is complex and continuously evolving. Most organisations, even larger organisations, cannot afford to track this landscape on their own, find it hard to evaluate competing claims from vendors and/or open-source communities, and must make fundamental decisions that will affect their own IT investments across multiple cycles of innovation.
- EU organisations need to take a more prominent role in the direction of and contributions to the open-source software technologies that are so important to cloud

and edge-related solutions. Many of the components of the current cloud technology landscape, while open source, are backed primarily by non-EU technology and innovation companies. For example, out of the 43 gold and platinum members of the Cloud Native Computing Foundation[119], which plays a key role in incubating and maturing new cloud-style tools and solutions, only one is based in the EU. Another measure of engagement is the comparative numbers of software updates, or "commits," being made by different groups of developers. Of the ten companies making the most contributions to OpenStack as measured by commits, eight are from the US, while none are EU-based[120]. This means that relatively few EU-based individuals and organisations are contributing to the mainstream of cloud-style innovation.

- Conversely, despite active support by the EU of a range of promising cloud-based innovations through the EC's research and innovation actions, very few of these innovations have gained more than limited acceptance in the market.
- Proprietary software ecosystems are complex combinations of software developed and controlled primarily by non-EU players. Promising new proprietary services such as Snowflake[121], combine proprietary software whose integrity and security cannot be directly verified, and which is currently available only on US hyperscaler-owned infrastructure (including various EU facilities). Even with open-source solutions, such as the OpenStack virtualization platform, customers often need complementary components to meet their full requirements, and customers frequently choose proprietary "versions" of OpenStack with the required features, introducing concerns over technical lock-in and possible loss of control over data flow and sovereignty.
- System integrators play a key role in the cloud adoption value chain, and most major system integrators focus their own businesses on implementations using services from the major US hyperscalers. In some sense, EU organisations look to these system integrators to solve problems of security, sovereignty and data protection using what the customers believe to be the technically more complete or superior offerings of these hyperscalers. For their part, system integrators do not spend time trying to convince customers of the value of using EU IT suppliers, creating a "vicious circle" that benefits US hyperscalers at the expense of EU suppliers
- The cloud technology stack is not just software – it extends to the underlying hardware infrastructure and ancillary services offered by cloud service providers. Even if an organisation builds its software stack from software components that are guaranteed to respect EU laws, regulations and values, this software must run on hardware whose ownership and location should be similarly documented and assured.

5.2. Skills

According to IDC's European Multicloud survey, 2021 (N=925),[122] "lack of skills" is identified as one of the four leading causes of cloud implementation failure. The cloud skills problem has several dimensions:

- Existing IT professionals need to learn about the new technologies and gain experience working with them. Cloud-style IT represents a significant change in approach that cannot be mastered by taking a single course. It is more complex than simply learning a new programming language. Components of a cloud "curriculum" might include:
    - Fundamentals of virtualization and containerization
    - Architectural patterns for virtualized/containerized IT systems[123]

[119] www.cncf.io
[120] https://www.stackalytics.com/?release=all&metric=commits
[121] "A single, global platform for your data and your essential workloads, with seamless data collaboration" www.snowflake.com
[122] IDC, European Multicloud survey. 2021
[123] Bill Wilder. Cloud Architecture Patterns. 2012.

  - Working with open-source technologies
  - Agile development methods
  - Continuous Integration/Continuous Deployment (CI/CD)
  - DevOps, Service Reliability Engineering (SRE)
  - Automation of CI/CD and DevOps, as well as many other common tasks
  - Designing for high-availability, geodiversity, scalable solutions.
  - Strategies for migrating on-premises information systems to cloud-style and cloud-based infrastructure.
- Established organisations and their managers need to learn about these same technologies and approaches, as well as how to integrate them into existing organisational structures and how to manage them effectively. For example, managing software development using traditional "waterfall" methods (exemplified by the Software Development Life Cycle, SDLC) is very different from managing agile software development. Business continuity and disaster recovery for cloud-based deployments are handled very differently from dedicated server deployments. New HR skills and approaches are needed to attract, qualify, and retain personnel with the right IT skills. The new IT skill set needed by larger organisations may in fact drive a need for increased in-house training and development, since new IT staff may not have the complete set of skills required.
- Open-source software plays a significant role in cloud-style IT deployment, and effectively using open-source software often demands a corresponding commitment to make contributions to those software projects. However such contributions assume a level of expertise in the software, as well as an organisational support for such a diversion of resources. As noted above, the EU's contributions to open-source initiatives lag those of the US and Asia[124], which points both to a lack of expertise and a lack of support for such contributions.
- Small- and medium-sized enterprises (SMEs), including start-up companies, in all sectors, face special challenges because of their size. They all need to become more data-driven, but the typical path to cloud readiness requires human and financial resources that they do not have and entails risks that loom large for smaller organisations (recognizing that the possibility of using public cloud has reduced both the costs and risks of such transformations). Even the use of "open-source" software, while usually free to use, requires initial and ongoing investment of time to understand and use the software correctly. (These skills challenges were identified in several demand sectors, including Public Administration (D-PA Challenge 6), Transport (D-T Challenges 5 and 7).
- Based on the statistics about cloud-based innovation and software development, presented above, the pool of individuals with skills and expertise in the cloud-style IT methodologies and software may be much smaller in the EU than in the US and Asia. For those skilled people working in academia (e.g., participating in EC-funded research and innovation projects), the lack of adoption of new cloud innovations by EU-based cloud providers implies a more limited flow of skills from academia to industry than is seen in the US.

5.3. Performance and Reliability

The same IDC survey referenced earlier [114] identified "performance issues" as another one of the four leading causes of cloud implementation failure. This concern mixes two kinds of performance issues, those related to the architecture of cloud applications, and those related to how these applications are deployed to, and behave on, one or more cloud infrastructures. Addressing this problem requires optimization of both an organisation's cloud architecture and

[124] https://octoverse.github.com

its choice of infrastructure. Such performance -- and reliability -- concerns are heightened for sectors operating critical infrastructure, such as energy and transport (see D-T Challenge 2). H-CLOUD's analysis of Skills in section 5.2 above drives recommendations to the EC to enhance skills development and training, specifically encouraging application of the principles of Service Reliability Engineering (SRE) and Continuous Integration and Deployment (CI/CD) will better equip EU organisations with the skills they need to engineer their cloud applications to achieve required levels of performance and resiliency.

Optimised applications can then be deployed -- in a deliberate fashion – to infrastructure with the required reliability. Increasingly, such deployment should be instrumented and automated so that performance and reliability problems of running applications are automatically detected and remediated, rather than requiring manual intervention by (scarce) DevOps staff. In addition to proper training for both user organisations and cloud service provider staff, all EU IT providers, including both cloud service and network providers, must be prepared to make clear commitments to their customers regarding reliability, availability, and quality of service in their service level agreements. For critical infrastructure (e.g., in the energy and transportation sectors), the EU should support both sector enterprises and IT suppliers to develop accurate measurements of reliability that can enable robust service planning. Such measurement efforts should include a clear baseline for on-premises IT solutions to support accurate consideration of cloud-based alternatives, without compromising security and reliability.

5.4. Security and Privacy

The IDC survey referenced earlier [114] identified "security" and "trust" as the last two of the four leading causes of cloud implementation failure. As discussed in [16], a key change involved in moving applications and data to the cloud is learning to operate outside the corporate security firewall. Security considerations are no longer the responsibility of a separate security function; instead, applications must be developed from the ground up with intrinsic, zero-trust security[125] built in, and vulnerability testing and remediation should be standard procedure.
Similarly, maintaining the privacy of sensitive data, as required by the GDPR, and as may be required by future legislation, requires that privacy must be engineered into applications at the design phase -- applying the principles of "privacy by design," especially in an exposed cloud environment. As with security more generally, vulnerability testing and remediation should be standard procedure.
These concerns are challenges for every sector considered by the H-CLOUD team (e.g., see D-T Challenge 2) and are amplified for smaller organisations or for organisations that work primarily with sensitive data (e.g., healthcare providers). Larger organisations are finding ways to fulfil the requirements of GDPR, but smaller organisations may be delaying their move to the cloud because of these concerns. H-CLOUD's sectoral analysis shows that cloud adoption is lower and slower among such organisations.
Best practice in security typically involves creation of security operations centres (SOCs) to consolidate a range of security monitoring, evaluation, and response functions. However, establishing and operating these SOCs can be challenging for larger organisations and unfeasible for smaller organisations. Shared service approaches, for example subscribing to "dark web" monitoring services (an example of "threat intelligence") on behalf of multiple organisations, would extend best practice benefits to more organisations.
Organisations prepared for actual deployment of workloads to the cloud have expressed a need for formally defined "labels" for cloud services that describe, in a reliable way, the security and privacy characteristics of those services.

- Early efforts to characterise these aspects of cloud services focussed on adherence to "codes of conduct". For data privacy in particular, codes of conduct have been

[125] NIST. Zero Trust Architecture. 2020

established by both CISPE[126] (an industry association of cloud service providers, focussing on infrastructure-as-a-service) and the EU Code of Conduct[127] (also covering platform-as-a-service and software-as-a-service products). The CISPE Code of Conduct allows service providers to guarantee that customers data 1) are stored and processed in Europe if the customer desires, 2) cannot be reused by the provider, and 3) are stored and processed on infrastructure whose location can be disclosed at the urban area level. However only 10 services have been certified under the EU Code of Conduct (and of these only one reflects full third-party audit verification). 108 services[128] report their adherence to the CISPE Code of Conduct from 33 distinct providers, of which 26 are headquartered within the EU. Out of these 108, 60 have been "certified" (through a process defined by CISPE), although almost 75% of these refer to non-EU based services. Of the over 60 services provided by EU-based providers, only 16 have been "certified", while the rest are based on self-assessment.

- Several national and regional initiatives have proposed "labels" or "codes of conduct". Germany developed the Cloud Computing Compliance Controls Catalogue (C5)[129], defining a baseline security level for cloud computing and enabling organisations to select, control and monitor their cloud service providers. France developed the SecNumCloud[130] certification to qualify trusted cloud service providers. Under the EU Cybersecurity Act, ENISA has also proposed a cybersecurity certification scheme for EU Cloud Services (EUCS)[131]. Efforts are being made to align and harmonize the different approaches.
- The EUSD[132] proposed creation of an "EU Cloud Rulebook" to consolidate all relevant legislation and regulation applying to cloud services. The Gaia-X initiative compiled its own early version of this "rulebook" in its "Policy Rules and Architecture of Standards" (PRAAS) document[133], but scaled back this approach in its Policy Rules Document 21.04[134] (PRD) to distinguish between "policy" and "techno-architectural" requirements. The PRD describes objectives to fulfil in terms of transparency, data protection / GDPR, cybersecurity, contracts, portability of data or data Sharing. Gaia-X is now discussing the development and implementation of "labels" that can be applied to services as a result of certified compliance with a set of requirements (Policy Rules, Architecture of Standards, Technical Requirement, Data Spaces Requirements) developed by the three main committees of Gaia-X (Policy Rules Committee, Technical Committee, Data Spaces Business Committee). These requirements, where possible, will rely on existing third-party schemes (e.g. codes of conduct, ISO certifications) and third-party certifications to ensure a proper level of conformance. Automated checks of conformance will be used where possible. The Gaia-X requirements expect to demonstrate an industry consensus between users and providers of cloud services, compared to the EU Cloud Rulebook which would consolidate regulatory requirements from across Europe.
- Meanwhile Cigref in France has acted to formalise the approach originally taken by Gaia-X, developing a "Reference Document on Trusted Cloud"[135] detailing a range of characteristics desired by Cigref's member IT organisations, covering contracts, security, GDPR, reversibility, interoperability, and sovereign immunity.

All these efforts seek to simplify an organisation's search for cloud-based services that are as secure and private as are needed. However, they ignore the fact that security and privacy

[126] CISPE
[127] EU Cloud Code of Conduct. 2020
[128] CISPE. Code of conduct public register. 2021
[129] BSI. Cloud Computing Compliance Criteria Catalogue. 2020
[130] ANSSI. SECNUMCLOUD. 2020
[131] ENISA. Consultation on the draft of the candidate Certification Scheme on Cloud Services (EUCS). 2020
[132] EC. A European strategy for data. 2020
[133] Gaia-X, Policy Rules and Architecture of Standards, June 2020
[134] Gaia-X, Policy Rules Document 21.04. 2021
[135] Cigref. Trusted Cloud. 2021

responsibilities start with the organisation itself, and that those responsibilities can at most be "shared" with IT partners and suppliers that include cloud service providers as well as a range of software providers and systems integrators. These responsibilities are extensive[136], are often under-resourced, and always represent a significant risk factor for cloud adoption.

5.5. "Build vs. Buy", Procurement and Financial Management

The adoption of cloud-style methods in information technology triggers changes in several traditional IT management approaches.

- Application architectures are being "deconstructed" from monolithic, tightly integrated structures into looser collections of interacting services and microservices. Business functions (e.g., credit card authorizations) that previously might have been hidden deep inside a complex software application might now be performed by an external specialised service provider or as part of an on-demand "function-as-a-service" capability from a cloud provider. Instead of giving a suite of business requirements to a development team, those requirements might be split up among internal developers, external service providers and software-as-a-service vendors for implementation, integration, testing, and deployment. What was once a pure software design and development activity now requires a combination of, and the coordination of, technology landscaping, vendor qualification, financial management and purchasing/contracting activities.
- Many smaller organisations, both SMEs as well as public administration and healthcare organisations, have common IT requirements that are strategic to their missions. Common, shared solutions could be developed collaboratively if group governance and organisational challenges could be addressed.
- The traditional split between responsibility for and management of capital and operating expenses, particularly in public sector organisations (public administration and healthcare) must now be managed in new combinations, with increasing levels of "shadow" IT spending beyond the control of formal IT departments. For the public sector in most member states most IT spending is linked to hardware purchases (and personnel). Even if the procurer would prefer to use the cloud, the budget cannot be used for operational costs.

As a major purchaser of IT services, with ambitions to increase the extent of cloud-style service delivery to citizens and stakeholders, the EC and member states should pave the way for cloud-based procurement and adoption. These actions should also be available to create benefits for SMEs in specific sectors and generally.

5.6. Addressing the Challenge of Legacy Applications: Refactoring, Reengineering, Re-assessing Requirements

For many organisations, including SMEs, legacy applications remain critical to their operations yet represent major impediments to digital transformation. Legacy re-engineering has many challenges[137]. It is easy to overlook this challenge but addressing it would advance digital transformation for many organisations.

[136] For example, Gary Duan in "The AWS Shared Responsibility Model for Kubernetes" explores the many aspects of securing a Kubernetes-based cloud stack on AWS.
[137] CIO. 11 Dark secrets of application modernization. 2021.

6. CHALLENGES TO CLOUD DEPLOYMENT: STRENGTHENING THE EU IT SUPPLY MARKET

Once an organisation adopts cloud-style IT methodologies, it can consider different forms of cloud deployment, including hybrid and "multi-cloud" combinations of:

- cloud providers whose ultimate parent company comes from outside the EU,
- cloud providers with global headquarters in the EU,
- distributed edge compute infrastructure, and
- cloud-style infrastructure on-premise, in co-located facilities or through managed private cloud.

Some of these choices raise specific challenges for EU organisations.

6.1. Non-EU Cloud Providers of Infrastructure- and Platform-as-a-Service (IaaS and PaaS)

Section 2.3 illustrates how the three US-headquartered hyperscalers (Amazon Web Services (AWS), Microsoft Azure and Google Cloud) have top-ranked market shares in IaaS and Paas, not only in the United States, but also in most other countries and regions around the world including the EU. Each of these providers offers an extensive range of products and services that is well-known to the development community, interoperable with other products from the same provider and with many open-source options, priced competitively and designed to make cloud adoption easy and flexible for client organisations.

- The three hyperscale platforms have traditionally not been interoperable across platforms, creating *de facto* vendor lock-in[138] that is felt by customers, not only in the EU, but also worldwide. As one US customer notes: "The actual differences between public clouds are minute, but getting there is not. If I spend the time to figure out how to get up on, say, Azure, I'm going to stay there, because it's too time-consuming to learn multiple proprietary cloud platforms."[139] (Note that new managed Kubernetes capabilities from these hyperscalers do allow greater interoperability for those customers able to work in this mode. Dependencies on platform specific services, such as security services can limit this flexibility.)
- The use of non-EU-headquartered cloud providers highlights the problem of digital sovereignty, where EU organisations relying on non-EU-headquartered providers face the risk that a foreign government could take actions that would limit their ability to fulfil their obligations, thrive and succeed on the world stage. This risk has been made real by events such as the US ban on supply of certain chip technologies to Huawei and ZTE, by the US Patriot and CLOUD Acts jeopardising security of sensitive data, and by possible scenarios such as actions by the foreign governments that serve their strategic or commercial interests at the expense of EU stakeholders. As described above, Cigref[140] proposes to identify cloud services that can be certified to be either "Safe" or "Trusted", where cloud providers are expected, among other things, to disclose details regarding any dependencies on non-EU parties and the extent to which they are subject to foreign regulations, laws, or state control.

[138] Microsoft has introduced specific bundling and licencing business terms that can add costs for organisations looking to host Microsoft software on cloud services other than Azure. (The Economist. How Satya Nadella Turned Microsoft around. 2020)
[139] Nutanix. Third Annual Enterprise Cloud Index. 2020
[140] Cigref. Trusted Cloud. 2021

6.2. EU-headquartered cloud providers, IaaS and PaaS

The EU IaaS and PaaS market segments are dominated by US-headquartered hyperscalers, while EU-headquartered cloud providers control a much more limited share of the EU market. Several strategies have proved successful, including strong orientation toward both the customer and the developer, using new business models (e.g., charging for smaller time slices or on a transaction basis), and assembling a sovereign value chain in their offerings including building unique partner ecosystems to meet customer needs. Nevertheless EU-headquartered cloud providers face several inherent challenges in this market:

- Most EU-headquartered cloud providers' infrastructure is significantly more limited than that operated by any one of the hyperscalers and is physically located in only one or a few EU countries, limiting scalability, particularly for EU organisations operating in multiple countries. Such geographic limits also make it harder for EU providers to offer services that depend on geographic diversity (such as disaster recovery and backup, and "follow the sun" service models).
- EU customers can construct most of the same technology stacks on either US-based hyperscale facilities or EU-headquartered cloud providers' infrastructure. The difference is that these options are available directly from the US-based providers, while for EU-headquartered providers, customers need to do more of the work themselves, requiring them to employ or contract staff with the right skills.
- Most EU organisations face a skills gap as they make their journey to cloud-based IT, and larger organisations often use the services of systems integrators (such as Accenture, Capgemini, Deloitte, etc.) which are more familiar with the technology stacks of the US hyperscalers than with those of EU providers.
- Smaller EU organisations face more extreme skills gaps and may tend to favour the more complete, more easily adopted offerings of US hyperscalers, despite the risks of vendor lock-in and potential loss of data sovereignty.
- The EU market's awareness of US hyperscalers and their latest offerings is much higher than of EU-based providers and their offerings. This is partly an inevitable consequence of the global relevance of US hyperscalers based on their market share. It is also a function of the importance of the EU market to the US hyperscalers[141] .
- Like EU organisations themselves, EU cloud providers suffer from skills shortages and generally do not have as many skilled personnel to offer the range of human-based services that might contribute to increased market share, such as outreach and awareness, training, client consulting, implementation advice and guidance.
- EU-headquartered cloud providers have not benefited from major procurements from public administrations and national organisations from the EU, nor are there any regulatory considerations that might make it more appropriate for sensitive projects (e.g., involving personal data or critical infrastructure) to be hosted on EU-headquartered facilities.

The current small market share of EU-headquartered IaaS and PaaS providers represents a challenge for these companies. As noted, the EU market for IaaS and PaaS is the second largest market for the US hyperscalers after the US itself, so these hyperscalers will work hard to keep this business.

Our recommendations in this target both the EC and EU IT suppliers -- in each case, both sides should work together to achieve the desired result.

[141] The EU market of US hyperscalers in IaaS and PaaS (worth an estimated US$9 billion in 2020) is their second largest market after the US itself (where the top 3 sold over $ US$23 billion in services in 2020). (See market analysis in Section 2.3)

6.3. Software-as-a-Service, Application Layer

The IaaS and PaaS segments are not the biggest or the most rapidly growing segments of the cloud market. Throughout the world, including in the EU, software-as-a-service (SaaS), or the “application layer”, is a comparable or larger opportunity than IaaS and PaaS combined (see section 2.3). This segment is also the most diverse, with the top three players, in almost every market worldwide including the EU, controlling a comparatively small proportion of the SaaS market.

After an organisation’s own IT resources, SaaS and application providers, along with system integrators, sit closest to end customers and are in an excellent position to help them benefit from digital transformation and digitalization. SaaS players are actively working to improve their ability to create value for their customers. For example, the EU-based SAP, one of the world’s largest SaaS/application providers, has recently partnered with Google to leverage Google’s big data analytics and AI skills to increase the value SAP can provide to its customers.[142]

The SaaS segment comprises traditional software applications being converted to cloud-based delivery by software vendors, as well as growing segments of big data analytics, artificial intelligence applications and frameworks, and data-driven services in general.

- As legacy applications (including “shrink-wrapped” and other software installed on premise -- or even on specific computers) are migrated to the cloud by their independent software vendors (ISVs), those vendors may choose to deploy their applications on foreign-controlled infrastructure, for many of the same reasons described above. This raises concerns about the security, sovereignty and trustworthiness of the complete “stack” of software and hardware used to deliver SaaS functionality. ISVs attempt to manage these concerns in various ways. For example, Salesforce indicates that it has 37 certifications -- however none of them on their own may be sufficient to give customers full confidence in using this SaaS application[143], and customers continue to bear ultimate responsibility for these issues.
- Even for experienced cloud-style IT adopters, new data analytics and AI applications can motivate new uses of public cloud, triggering renewed concerns about security, sovereignty, and trustworthiness. These may arise even for organisations that operate their own on-premises facilities or well-established hybrid cloud structures, but still use foreign-supplied software components whose dependencies and integrity are not transparently disclosed.

H-CLOUD has developed several recommendations intended to support the creation of new IT solutions that address the needs of EU organisations, particularly SMEs, and accelerate their adoption of cloud-style IT solutions. Many of these recommendations should also be embraced by EU SaaS vendors as pathways to new markets and expanded revenues.

We make additional recommendations targeting SaaS players, particularly in their offerings to SMEs, but as above, they are not divided between the EC and the EU cloud industry -- in each case, both sides should work together to achieve the desired result.

6.4. Private Cloud and On-Premises Deployment

Section 5 makes it clear that the use of the public cloud is not in itself necessary to enable the digitalization that is so important for the EU economy. While H-CLOUD was unable to identify detailed market data on the colocation, data centre, and private/managed cloud markets in the EU, actions should be taken to ensure that these infrastructure options enable the same

[142] SAP. Google Cloud and SAP Partner to Accelerate Business Transformations in the Cloud, 2017.
[143] Salesforce. Compliance engineered for the Cloud. 2022

support for digitalization as public cloud, while also providing the same attributes of security, sovereignty and sustainability desired from the public cloud and the cloud-to-edge computing continuum.

Many EU organisations have made strategic decisions to continue to operate significant portions of their IT capability on premise (in their own data centres or in co-location facilities), in private or managed cloud infrastructure, or in "virtual private clouds" (VPC) where customers "rent out" dedicated (not multi-tenant) infrastructure from an IaaS provider. Some of the challenges addressed above, such as concerns over security, privacy, and data protection, motivate these decisions. For example, questions of security and privacy have been identified by German organisations as impediments to moving to the public cloud, and this may in part explain the lower-than-expected enterprise cloud adoption statistics for France and Germany mentioned in Section 2.3.

6.5. Hybrid and Multi-Cloud, Federation

Fewer and fewer companies operate in a "pure public cloud" or "pure on premise" IT environment. While in some cases the use of hybrid and/or multi-cloud architectures is a deliberate strategic choice, many companies find themselves in the multiple cloud world without planning it, mostly in the context of SaaS offerings, often by virtue of independent decisions about the use of the cloud across the organisation. Without careful management and supervision, this can create interoperability challenges as organisations increase the extent of digital integration, increase cybersecurity attack surfaces, and increase costs.

Several hybrid- and multi-cloud management tools have been developed to help with this problem. Most of them have been developed by specific vendors (including Google Anthos, Azure Arc, AWS Outpost, as well as many others), raising the question of whether the management tools themselves might reduce the ability of EU organisations to make sovereign choices about their own digital footprints.

The challenge of integrating infrastructure and services from multiple providers will be exacerbated as edge computing grows in importance. Edge computing customers should not be restricted to the use of infrastructure owned by a single service provider. Robust integration of edge infrastructure from multiple providers across Europe will be essential for broad deployment and adoption. This challenge has been identified by the European Alliance for Cloud, Data and Edge[144], Horizon Europe[145] as well as H-CLOUD's "The Potential of Cloud Federation"[146] (see S-F Challenge 2: "Defining, Evolving, Selecting, Agreeing on and Managing the Architecture, Technical Standards and Tools for Federated Clouds and Distributed Data Access and Exchange"). Although the underlying technologies are rapidly evolving, and the scope of integration must extend to the more heterogeneous edge computing environment, it is critical to find agreement on the architectural framework and the standards to be adopted within that framework. Establishing a platform architecture enables technical discussions to be modularized and compartmentalised and facilitates agreement on standards that enable specific services to interoperate.

Note that "multi-cloud" and "federated cloud" are two views of the same technical activity: the integration of infrastructure and services from multiple providers. In multi-cloud deployment, the customer organisation is responsible for this integration and for reaching the service levels it needs to meet its objectives. In federated deployment, participating providers agree on technical and operational standards and procedures that allow the different services to be integrated, with defined service levels, for the benefit of customers.

---

[144] AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. 2021
[145] EC. HORIZON-CL4-2021-DATA-01-05: Future European platforms for the Edge: Meta Operating Systems (RIA). 2021
[146] H-CLOUD. Supply Analysis: Cloud Federation (S-F). 2021

In a commercial context, where the customer pays for services and infrastructure through contracts with providers, federation may not offer significant operational benefits. However, federation is a common approach to multi-supplier integration in the public research community since it respects the separate funding structures and governance of the participating research infrastructures. Federation could be an attractive mechanism for increasing the scale and capability of resources in a variety of other publicly funded activities, such as healthcare, education, and public administration more broadly.

Several initiatives can be seen in the global market on federation models. There are grids of interconnected hardware devices that use blockchain technology to manage the federation. Most of the initiatives federate part of the cloud stack, such as storage (Filecoin, Maidsafe, SIA, Storj) or compute (Golem, SONM), while ThreeFold Tech offers the ability to federate across the full cloud stack.

6.6. Edge Computing

Section 4.2 explored how edge computing is a critical ingredient for the digitalization of organisations as they increase their level of “data awareness”. Section 2.3 explained how edge computing could become an important IT deployment option for EU organisations, potentially even more important than today’s Infrastructure-as-a-Service and Platform-as-a-Service offerings. This section explores the technical and business challenges associated with this trend.

Edge computing offers several technical benefits, particularly lower latency and increased responsiveness that are required for many use cases (autonomous driving, robotic surgery, etc.), which could have a positive impact on sustainability and energy-efficiency. Increased use of edge computing avoids the need to move growing data volumes to core data centres, which costs money, takes time, and triggers increased CO_2 emissions.

Potentially edge computing could disrupt the dominance of hyperscale data centres. A similar trend was observed in the energy industry, where solar panels and wind turbines released energy production from the exclusive control of large power generators. A geographically dispersed grid of lower-cost small cloud capacity nodes could democratise the cloud industry. Advanced automation -- which is needed anyway given that manual management of a multitude of cloud devices is not realistic, both from a physical access perspective as in terms of workload -- could reduce the operating costs of cloud hosting facilities.

Realising the potential of edge computing will require several technology challenges to be solved[147]:

- What capabilities (compute, storage, bandwidth, as well functions implemented by software) will be needed to enable distributed infrastructure to work together to meet the needs of customers? How can these capabilities be provided using low-power, efficient hardware combined with lightweight software systems?
- What coordination mechanisms can be used to enable processing of distributed data, while still preserving privacy and security?
- What federation mechanisms can be used to seamlessly integrate these devices and services both technically and operationally, recognizing that edge computing applications will build on a potentially large number of edge and IoT devices and connecting networks, distributed over a large territory, owned and operated by many providers.

[147] ATOS. A 2021 perspective on edge computing. 2021.

- How can the hyperconnectivity available to edge services be managed to maintain the latency and bandwidth benefits of edge computing?
- How can the environmental impact of widely distributed IT systems be controlled and optimised, particularly with the dense network of wireless communications links needed between edge computing nodes?

There are also several business challenges to be addressed[148]:

- Edge computing will transform the relationships among the cloud, fixed telecommunications providers, and mobile service providers.
- While the cloud suggests a business model for edge computing, it is a simpler model to understand than the distributed model that will be needed by "edge service providers." ESPs will have to integrate "network slices", orchestration services, and shifting ensembles of distributed devices into stable edge computing and storage offerings with defined levels of service and pricing.
- Obstacles to SMEs entering the edge service market will be higher than for the cloud market, requiring widely distributed assets, and correspondingly distributed personnel to install and maintain those assets.
- Geographic variation in service availability and coverage density will challenge edge service providers and edge customers alike, making deployment difficult for providers, and adoption difficult for customers.

[148] 5GPPP. Edge Computing for 5G Networks. 2021

7. CROSS-CUTTING FOUNDATIONAL ELEMENTS

Several aspects of cloud computing are foundational, touching every component of the cloud-to-edge continuum.

7.1. Ensuring that the IT Capabilities Required for Digital Transformation are Environmentally Sustainable

The EC has clearly communicated how important it will be for Europe's data centres to be climate neutral by 2030[149]. However, this is only a part of the challenge. This White Paper identifies what is needed to encourage Europe's digital transformation, so we must explore what is needed to ensure that the IT capabilities required to enable that transformation are themselves environmentally optimised and sustainable.

For this discussion, the environmental impact of IT capabilities can be explored along five major dimensions of sustainability: software (including operating systems and virtualization), hardware, data centre, network, and energy sources, although there are significant dependencies between these dimensions.

7.1.1. Software, Operating Systems and Virtualization

Software is where the productive work of digital transformation occurs, but its contribution to the environmental impact of IT has received little attention, except in the research community, specifically in connection with high performance computing (HPC) and artificial intelligence.

In those domains, algorithm design, complexity and performance are examined in detail, since at the scale of HPC calculations, small changes in performance can allow large increases in the scope of problem that can be analysed in a single HPC system (e.g., the number and size of genomes, the resolution and geographic coverage of climate models). While individual HPC systems have consistently grown over the last 30 years[150], the demand for such systems has grown at least as fast, and HPC resources are "supply constrained" across the globe.

Similarly, the artificial intelligence research community has increasingly considered both the cost and the environmental impact of training and optimising deep learning models[151]. OpenAI has estimated that the computational effort required to train state-of-the-art deep learning models has increased at least 3x per year -- far more than the increased computational capacity that might be enabled by Moore's Law[152]. Although expressed in terms of the "hardware" required to train these models, this trend results from the growing complexity of AI model training, combined with comparatively stable performance of the software used to train those models.

As HPC, AI and big data analytics grow in importance for organisations and demand sectors outside of the research domain, and as their needs for these capabilities increase, the efficiency of the software used to support these functions will become increasingly important to IT's impact on the environment. As these activities move out of the publicly funded research domain, resource growth will be tempered by cost and budget considerations, but the need for efficient software tools will continue.

For less advanced, but more common software requirements -- such as online banking or customer relationship management, the question of efficiency usually translates into trade-offs among the speed of execution, the response time offered to a user and the cost of running the

[149] EC. EU Digital Strategy, February 2020.
[150] Top500, Performance Development. 2022.
[151] E. Strubell, A. Ganesh, A. McCallum, Energy and Policy Considerations for Deep Learning in NLP. 2019
[152] OpenAI, Measuring the Algorithmic Efficiency of Neural Networks. 2020

application. This can be improved through good algorithm design, as well as the use of best practises in overall architectural design. Poorly designed algorithms do not scale well and cost their organisations (including software-as-a-service providers) in poor performance and/or excessive hardware needs, again triggering economic costs that will motivate improved efficiency.

At this level, the efficiency of the underlying operating system, and more generally the efficiency of the “stack” of software technologies required by the applications should be considered. Moving IT applications to the cloud represents an opportunity (albeit not fully exploited) to redesign those applications into smaller interoperating software modules that can be deployed across a variety of hardware systems (through “virtualization” or “containerisation”), creating two benefits: the ability to increase or decrease the resources (“elasticity”, “scaling”) for specific software modules as requirements increase and decrease, and the ability to deploy those modules to maximise the utilisation of physical hardware. Both benefits allow hardware resources, and related energy requirements, to be optimised to current demands, reducing the use of energy to power unused hardware.[153]

This benefit of virtualization is typically associated with moving workloads to the cloud, but it is available even in on-premises facilities. For example, Microsoft describes a 98% reduction in energy requirements when Microsoft Exchange (email) workloads are moved to the cloud, however, almost all of these benefits result from the differences between an on-premise server version of Microsoft Exchange and the virtualized version that Microsoft itself uses to provide its “email-as-a-service” version of Exchange, rather than from inherent efficiency of the cloud itself.[154] Similar improvements in efficiency are possible on premise as a result of software redesign.

Accessing the energy-efficiency benefits of containerization is more challenging in on-premises and edge computing environments, since there is a smaller pool of software workloads in a single enterprise or at a single edge node than might be found in a larger cloud data centre. The hybrid cloud approach is critical for maintaining high levels of hardware utilisation, where demand in excess of local capacity (either in an on-premises data centre or in a specific edge node) can be satisfied by other resources (e.g., the public cloud in a hybrid cloud arrangement, or nearby edge nodes for edge computing).

Optimising hardware utilisation must increasingly be balanced against optimising environmental impact. As will be discussed below, environmental impact is increasingly a function of the environmental character of local power sources, and workloads will increasingly need to be distributed to where environmentally friendly power is available. This “bring the compute to the green power” approach adds an additional layer of complexity to the resource optimisation process.

Optimization of software efficiency would be better enabled with consistent reporting and management of hardware deployment tracking and statistics, such as is available from cloud service providers, as well as increased awareness of applications of their own demand drivers. Ideally, an application should self-manage by being self-aware and detect bottlenecks, redeploy to avoid the bottlenecks, and even "recompile" some of its own modules to exploit future hardware accelerators such as GPUs, FPGAs or TPUs as they are available. This kind of data, combined with data on costs and environmental impact, can enable automated application management to optimise cloud deployments -- if similar data were available for all kinds of deployment, this would allow client organisations to optimise the mix of on-premises, private cloud, public cloud, and eventually edge computing resources.

[153] Koomey, J. and Taylor, J., Zombie/Comatose Servers Redux, April 2017.
[154] Microsoft, The carbon benefits of cloud computing. 2020

7.1.2. Hardware

The previous section highlights how software and hardware efficiency and environmental impact are closely linked. Once software has been optimized and virtualized, the energy efficiency of the hardware itself should be considered. For many years, Moore's Law[155] has enabled the regular refresh of compute hardware, with regular increases in capacity at roughly constant prices, while Dennard Scaling[156] has enabled improvements in power efficiency at almost the same rate, keeping the power required by that increased capacity under control.

However, both technology trends are waning, and it is unclear if historical trends can continue. For example, chip manufacturers are working to increase the density of the integrated circuits that make up today's compute, storage, and networking hardware. Some of the most advanced chips are fabricated at a scale of 7 nanometers (nm) for each transistor, representing the width of 35 silicon atoms or 1,200 atoms altogether. Moving to 2nm transistor spacing would mean that transistors would have to be implemented with only 100 atoms of silicon.

If these trends wane, continued increases, and even increasing rates of growth, in IT compute demand will translate into increased requirements for compute hardware, which will start to cost more to purchase and which can be expected to see slower increases in energy efficiency. Increasing amounts of compute capacity will therefore translate into increased energy requirements.[157]

A slowing of transistor density improvements (Moore's Law) also translates into changing economics for equipment replacement. Historically a 5-year replacement cycle was perceived by IT professionals as economically efficient, however longer useful lives might make economic sense. This in turn would allow the embedded environmental impact of IT equipment to be better "amortised" over a longer useful life, although this aspect of IT's environmental impact has not been fully analysed.

Hardware providers -- including both manufacturers and cloud-based service providers -- should track and report energy and environmental efficiency of their resources to assist customer organisations in making informed decisions about the right hardware to acquire. For example, cloud service providers could list energy and environmental impacts associated with the different types of hardware they offer, along with the different price options for each hardware type. Actual energy consumption and CO_2 generation could be reported as part of periodic billing. Hardware manufacturers could provide validated factors for energy and environmental impact, which can be incorporated into calculators for resource management and orchestration tools. More generally cloud providers should optimise their internal operations to not only minimise costs but also environmental impacts.

Most cloud hyperscalers have transitioned to hardware manufactured to standards developed for the Open Compute Project (OCP)[158]. This approach saves money by eliminating many product features desired by enterprise customers, which hyperscalers can provide for themselves. OCP designs are also increasingly optimized for more energy efficient data centres. IONOS and OVHcloud have taken a similar "production engineering" approach to the design of their own servers and packaging[159]. Adoption of such approaches by more EU-based organisations could translate into capital and operating cost savings, but those organisations would also need to build required capabilities in managing their hardware infrastructure.

[155] G.E. Moore. Cramming more components onto integrated circuits, 1965
[156] Robert H. Dennard , Fritz H. Gaensslen , Hwa-nien Yu , V. Leo Rideout , Ernest Bassous , Andre , R. Leblanc, Design of ion-implanted MOSFETs with very small physical dimensions, 1974
[157] Dietrich, M, Chen, A.. Projecting Optimum Replacement Cycles for Scientific Computing Equipment (forthcoming at http://people.cs.uchicago.edu/~aachien/lssg/research/zccloud/lifetime/)
[158] Open Compute Project
[159] OVH production engineering approach.

Manufacturers should provide greater detail and granularity regarding embedded environmental impacts of IT. Current data is somewhat generalised and not appropriate for rigorous "Scope 3" GHG reporting or for careful planning and optimization. Typically, generic server configurations are used that do not address potentially large impacts of using different components. For example, a "high end" server configuration was analysed and found to have very high embedded CO_2 levels as a result of the use of solid-state drives (SSDs), instead of traditional "spinning disk" hard drives[160].

7.1.3. Data Centres and IT Hosting Infrastructure

The efficiency of Europe's data centres is an important factor in the overall environmental impact of European IT and digital transformation. Data centres house IT infrastructure and are purpose built to allow high density delivery of power to compute and storage equipment, to efficiently remove the heat generated by that equipment, and to ensure the physical security of hosted equipment and systems. Other IT hosting approaches are possible particularly for remote equipment and equipment at the network edge. Since remote and edge computing equipment requires less space than servers in a data centre, dedicated facilities may not be required, power is often supplied from "household" grade wiring, and heat can be ejected into the surrounding environment. Nevertheless, the physical environment and physical security of these facilities should be properly planned and managed. Supplying reliable power to remote locations is a particular challenge for edge computing deployments, where even comparatively power-efficient racks might consume 6-12 kW of power, which cannot be cooled with fresh air alone.

As compute densities increase (even with a waning Moore's Law), the amount of power that needs to be supplied to each compute server increases, and the amount of heat generated by that server also increases, necessitating increased investment in heat removal technologies. The total power efficiency of a data centre is often expressed as its power utilisation effectiveness (PUE). This metric is not perfect, and other metrics have been proposed.

Air cooling has been the traditional method for heat removal, requiring industrial air conditioning units and external heat exchangers and cooling towers, which are not energy or environmentally efficient, contributing to high PUEs for many data centres. Increasing power densities are driving the use of new technologies that use liquid cooling, from rear-door heat exchangers to direct-liquid-to-chip and dielectric immersion cooling, all of which offer energy and environmental efficiency benefits, but which require purchase and installation of new data centre equipment. The EC might consider incentives for installation of these power-efficient technologies if the payback on such investments by data centre operators is not sufficient to encourage more rapid adoption.

As power densities increase, data centres must also arrange for increased power supply from local energy providers – or locate new data centres where environmentally friendly (renewable) power is available and optimise the distribution of workloads to minimise environmental impact. Increasing power supply in high demand areas can be a challenge in many locations, where there are constraints on available power, relevant power distribution capacity, and even tolerance for expansion from local communities. For demonstrably efficient data centres, incentives might be provided to those energy providers to support such supply upgrades. Costs per unit of energy (kilowatt hour) are expected to continue to rise over time, particularly as the energy ecosystem invests in new generation (including renewable sources) and new transmission and distribution (to replace ageing infrastructure and to cope with increasingly challenging weather conditions). Data centre operators thus have a natural incentive to themselves invest in energy efficiency.

[160] Thinkstep AG. Lifecycle Cycle Assessment of Dell R740, June 2019 .

Electric power is clearly a critical input, but water consumption and possible re-use of exhaust heat are also critical factors. Some metrics have been developed in this regard (e.g., water utilisation efficiency, WUE), but additional metric definitions and application are needed. Incentives should be put in place to encourage improvements (or adoption in the case of heat re-use) in these aspects of data centre operation.

EU data centre operators and trade associations have signalled their commitment to the European Green Deal through the Climate Neutral Data Centre Pact, agreeing to make data centres climate neutral by 2030. Some commitments are very specific[161]:

- By 2025 new data centres operating at full capacity will meet an annual PUE target of 1.3 in cool climates and 1.4 in warm climates. Existing data centres will achieve these targets by January 1, 2030. (Applicable to all data centres larger than 50kW of maximum IT power demand.)
- Data centre electricity demand will be matched by 75% renewable energy or hourly carbon free energy by December 31, 2025 and 100% by December 31, 2030.

Other commitments are more directional:

- Trade Associations will work with the appropriate agencies or organisations toward the creation of a new data centre efficiency metric. Once defined, trade associations will consider setting a 2030 goal based on this metric.
- By 2022, data centre operators will set an annual target for water usage effectiveness (WUE), or another water conservation metric, which will be met by new data centres by 2025, and by existing data centres by 2030.
- Data centres will set a high bar for circular economy practises and will assess for reuse, repair, or recycling 100% of their used server equipment. Data centre operators will increase the quantity of server materials repaired or reused and will create a target percentage for repair and reuse by 2025.
- Data centre operators will explore possibilities to interconnect with district heating systems and other users of heat to determine if opportunities to feed captured heat from new data centres into nearby systems are practical, environmentally sound and cost effective.

7.1.4. Renewable Power Sources

Data centre energy efficiency is a stated goal of the EC and of data centre operators, however reducing the environmental impact of data centres and increasing their sustainability should also be policy objectives and encouraged by government(s). Data centres should be required to provide transparent reporting, using well defined metrics, of environmental efficiency, including greenhouse gas generation per unit of net power supplied to IT infrastructure. Even if low-cost power supply contracts have been arranged, data centre operators should still report full Scope 1, 2 and 3 impacts according to the GHG Protocol. Incentives and supports should be provided to data centres to improve both their energy and environmental efficiency, consistent with best practises and supervised to ensure that efficiency improvements are not achieved by shifting the burden elsewhere (for example through heat pollution).

7.1.5. Network Energy-Efficiency, including 5G

Data centres, and IT hosting infrastructure in general, must be connected with data networks. Most data centres are equipped with high bandwidth network connections, linking them to public networks (such as the Internet) and to private networks to receive and generate the high volumes of data being stored, processed, and generated. Historically, the energy requirements of these networks, which are mostly wired (as opposed to wireless) networks, have been

[161] Climate Neutral Data Centre Pact. Self Regulatory Initiative. 2021.

significantly less than that of the data centres, particularly compute servers, even for long distance data transmission, and this aspect of energy consumption has not been a point of focus.

As more data is generated and received by mobile devices, and increasingly by other devices at the network edge, energy consumption by wireless networks will be a source of growing energy consumption and negative environmental impacts. While the design objectives for 5G wireless communications specified very low energy required per bit, early experience with 5G rollouts is highlighting higher than expected energy required to power the 5G radio networks. Coupled with the higher number of 5G base stations that will be required, and the simultaneous operation of old and new wireless networks, this could translate into significantly increased energy consumption. The location of these base stations may also make it more difficult to arrange for renewable power sources for remote 5G base stations. As the move to edge computing accelerates, it will be important to understand the dimensions of this problem and determine whether additional actions are needed to avoid unintended increases in greenhouse gas emissions associated with increased use of edge computing.

7.2. Cybersecurity

Cybersecurity is out of scope for the H-CLOUD project, but it addresses a range of extremely critical challenges and concerns, with a vast range of initiatives underway and planned. H-CLOUD can only recommend that cybersecurity needs to be an intrinsic component of every initiative related to the cloud-to-edge continuum. It is no longer sufficient -- or effective -- to "add" cybersecurity features at the end of a project's development. Cybersecurity, and related privacy and data protection, must be built in at the design phase.

7.3. Skills

As with environmental sustainability and cybersecurity, the stock of relevant skills and expertise among the EU workforce is a fundamental issue that touches every aspect of this analysis. H-CLOUD's prior report[162] identified this as one of four major challenges "Major Challenge M2. Limited skills and expertise especially in smaller organisations".
Each step of the digital transformation value chain requires a workforce with the required skills and expertise:

- Becoming Data Aware requires workers who can understand how data is collected inside and outside their organisation and can support efforts to find essential data that can drive efficiencies and growth. There will be continued growth in several new job categories, associated with data collection, cleansing and analysis, artificial intelligence, data analytics, data ecosystems, data governance, etc.
- Becoming Data Driven requires organisations to reimagine their own operations and requires workers who can take this transformational journey, who are excited about the opportunities it opens up, and confident that this transformation will create opportunities for them personally.
- Moving to cloud-style IT methodologies requires a range of new digital skills, not only from traditional IT specialists, but also from those workers directly involved in these transitions. Many of these skills challenges are addressed in Sections 5.3.
- As cloud-style IT capabilities are deployed across the EU IT supply sector, "traditional" cloud providers will need more workers with a growing range of skills, and the projected cloud-edge computing continuum will drive a major change in workforce patterns and skill requirements.

[162] H-CLOUD. D3.1 – Strategy Analysis Report and Cloud Computing v1.0. 2021

Pressures on employees and workers along this entire value chain must be addressed with broad training and workforce development programs. The recently released “European Skills Agenda”[163] focuses on 5 key actions:

- A call for collective action, mobilising business, social partners and stakeholders, to commit to working together, in particular within the EU’s industrial eco-systems;
- The definition of a clear strategy to ensure that skills lead to jobs;
- The measures to support people build their skills throughout life in an environment where lifelong learning is the norm;
- The identification of the financial means to foster investment in skills;
- Ambitious objectives for up and reskilling to be achieved within 2025.

While the “European Skills Agenda” does not only focus on digital skills, it builds on the digital gaps and requirements identified in key policies such as the European Green Deal, the Digital Strategy, and the Industrial and SME Strategies, giving to digital skills a central role, resulting for example in the Digital Skills and Jobs Coalition[164]. Furthermore, concrete actions are already devised in linked EU programmes such as Digital Europe Programme[165] and the Digital Education Action Plan 2021-2027[166]. In particular, the Digital Europe Programme will support the design and delivery of specialised programmes and traineeships for future experts in key capacity areas like data and AI, cybersecurity, quantum and HPC; while the Digital Education Action Plan defines actions to foster the development of a high-performing digital education ecosystem and to enhance digital skills and competences for the digital transformation. The European Commission also defined the target KPIs to be reached in term of Digital Skills by 2030 as part of the Digital Compass:

- 80% of all adults have basic digital skills (2020 baseline: 58,3%) as defined in the European Pillar of Social Rights Action Plan[167];
- there are at least 20 million employed ICT specialists in the EU, with convergence between men and women (2019 baseline: 7.8 million).

7.4. Foundational Cloud Research

As the demand for cloud-style IT increases, driven by Europe’s increasing digitalisation and digital transformation, and as the modes of deployment expand from more traditional cloud offerings across the cloud-to-edge continuum to federation and to more environmentally friendly approaches, research will be essential to guide effective and practical development and deployment. Section 2.6 has summarized recent EC-funded cloud-related research, covering self-adaptive cloud infrastructure and services, cloud-edge orchestration, DevOps for cloud services and security and trust for cloud infrastructure, services and data.
Additional research investment is needed to prepare for the large-scale investment in advanced cloud capabilities envisioned for Europe. Two analyses of this topic provide useful guidance:

- Research Roadmap developed by the Future Cloud Cluster[168]
- Research topics included in the “European industrial technology roadmap for the next generation cloud-edge offering”, prepared by the European Alliance for Industrial Cloud, Data and Edge.

[163] EC. European skills agenda for sustainable competitiveness, social fairness and resilience. 2020
[164] https://digital-skills-jobs.europa.eu/en
[165] https://digital-strategy.ec.europa.eu/en/activities/skills-digital-programme
[166] EC. Digital Education Action Plan 2021-2027. 2020.
[167] EC. European Pillar of Social Rights Action Plan. 2020.
[168] https://eucloudclusters.wordpress.com/future-cloud/

7.4.1. "*Research Roadmap*" of Future Cloud Cluster

The Future Cloud Cluster assembled participants working across the spectrum of cloud-related research in Europe, including representatives of over 30 existing EC-funded projects and collaborating institutions in this area. This group developed its Research Roadmap[169] to reflect their vision and perspectives, based on working directly at the forefront of research in the areas of federated clouds, edge computing, computing continuum and multi-cloud. Thirteen major research areas, grouped into four main topics, were identified and are described in Table X.

The main topics align directly with the pillar areas identified in the sections above and underscore a key message of H-CLOUD that the related areas of multi-cloud, federated cloud, the cloud-to-edge continuum and integration with edge computing are both strategically important and will still require research and development to enable effective deployment. Aside from "Area 1" which includes "consideration of data privacy … and applicable legislation" (touching on the Data Awareness pillar, Section 4 above), most research areas address the "Strengthen EU IT Supply" pillar (Section 6 above).

Main topic	Research Area	Description
Cloud Federation	Area 1: Federation of clouds	Facilitate Cloud and Edge interoperability and portability, considering data privacy, security and applicable legislation
Multi-Cloud	Area 2: High Performance Heterogeneous Cloud Services	Management of high-performance computing and high-throughput computing resources
Multi-Cloud	Area 3: Security mechanisms across Clouds	Security controls across providers
Multi-Cloud	Area 5: Knowledge-Based Service Harvesting and Acquisition	Ability to discover and compose services from existing cloud offerings marketplaces
Multi-Cloud	Area 10: Software defined Infrastructures and Novel composition models	Mechanisms for abstraction and virtualisation of resources
Cloud Continuum	Area 4: QoS and SLAs	Enforcement of quality of service across cloud models
Cloud Continuum	Area 6: Dynamic Configuration, Provisioning, and Orchestration of Resources	Orchestration of applications across diverse Edge and Cloud environments
Cloud Continuum	Area 9: Energy efficiency and Sustainability for Edge and Cloud Continuum	Energy optimisation in Edge and Cloud environments
Cloud Continuum	Area 11: Novel Data Storage infrastructures and services	Strategies for data intensive application execution
Cloud Continuum	Area 12: Software application development for the computing continuum	Software engineering, design and programmability of applications in Cloud continuum

[169] Future Cloud Cluster. Research Roadmap. 2020.

Main topic	Research Area	Description
Cloud Continuum	Area 13: Research artefacts for the computing continuum	Researchers access to tools, data and infrastructures
Edge computing	Area 7: Deployment and management of resources: in a decentralised, autonomous way	Autonomic resource management
Edge computing	Area 8: Resilience and Scale	Continuity in service provision for Edge limited operation

Table 3. Key research areas identified by the Future Cloud Cluster (The areas listed by main topic rather than sequentially.)

These 13 research areas were refined to identify over 30 specific challenges, and these challenges are further assessed with respect to impact and timeframe, illustrated in Figure 17. Note that, despite the immediate strategic importance of all these challenges (given H-CLOUD's analysis), some of these challenges will require concerted and sustained effort before stable solutions can be made available.

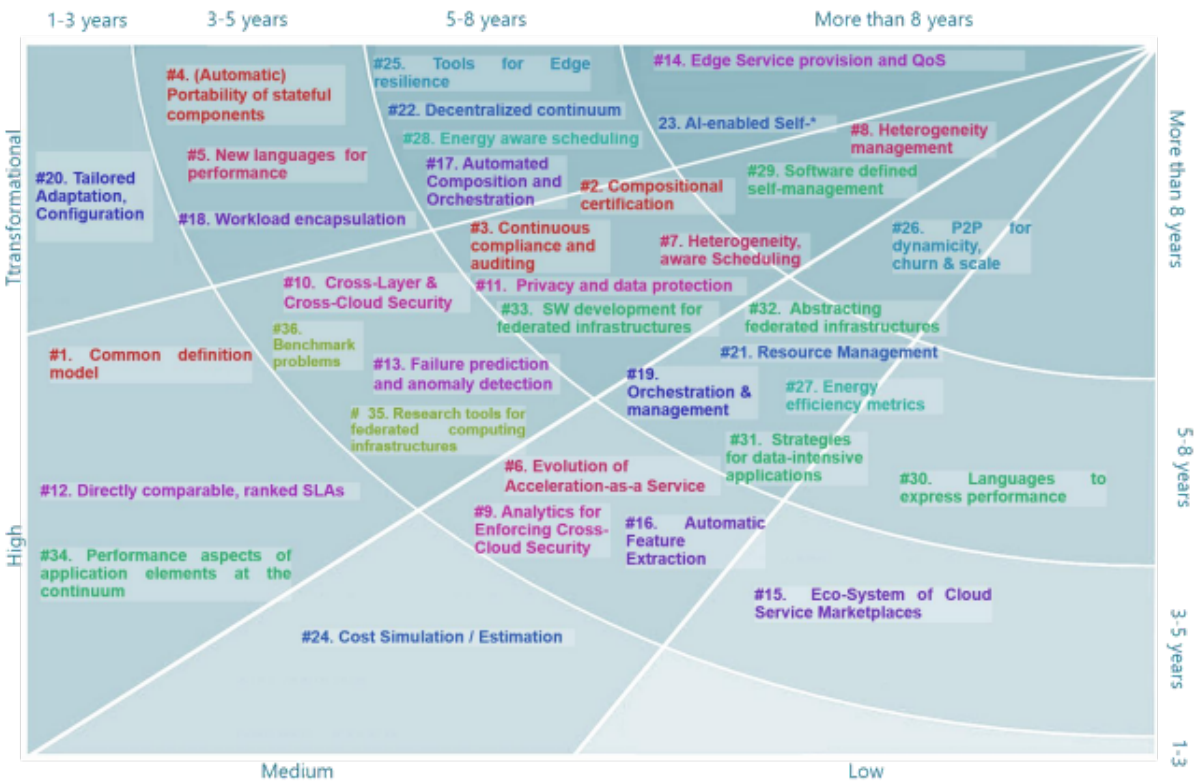


Figure 17. Timeline of key research areas identified by the Future Cloud Cluster.

7.4.2. *"European industrial technology roadmap for the next generation cloud-edge offering"* of European Alliance for Industrial Cloud, Data and Edge

The European Alliance for Industrial Data, Edge and Cloud[170], a consortium of 27 key players in the European IT landscape, developed a detailed analysis of the challenges faced by Europe's cloud sector: documented in "European industrial technology roadmap for the next

[170] https://digital-strategy.ec.europa.eu/en/policies/cloud-alliance

generation cloud-edge offering" (EITR)[171]. As with the Future Cloud Cluster's Research Roadmap, H-CLOUD has incorporated the Alliance's full range of recommendations, including those that include a research component, for consideration by the broader cloud community.

The EITR proposes investment of over 3 billion euros for research topics alone plus the estimated research component of mixed research and deployment topics. Investment figures combine contributions from industry and government (both the EU and Member States), but for R&D topics most of the investment support is requested from government. These are significant figures, but they are consistent with the role that cloud and, as we term it, "cloud-style IT" are expected to play in Europe's digital transformation. The proposed research and research and deployment topics can be mapped to H-CLOUD's strategic framework.

Data Awareness: Three topics are proposed along with resources estimated at 750 M Euro:
- Pan-European Data Sharing Platforms (500 ME)
- End-to-end data pipelines (130 ME)
- Ecosystem middleware and runtimes (120 ME)

Data Driven: Two recommendations are proposed supported by 300 ME are intended to create better platforms to enable data driven businesses.
- Creating an open-source open standard software stack (250 ME)
- Creating managed platforms optimized to support sectoral IT requirements (50 ME)

These both also contribute the next pillar.

Strengthen EU IT Supply: Several significant research topics are proposed by the Alliance in this area, with 650 ME of total investment:
- Creating a cloud infrastructure control plane (200 ME)
- Creating a multi-cloud control plane (150 ME)
- Orchestration capabilities working end-to-end on the cloud-to-edge continuum (150 ME)
- Three efforts enabling a range of edge-based capabilities (totalling 150ME)

Foundations: Five topics are proposed by the Alliance, with roughly1.4 Billion Euros of investment:
- Cybersecurity-related: Data encryption (500 ME) and zero trust security architectures (500 ME).
- Carbon Neutral Digital Transformation: Advanced data centre infrastructure management (DCIM) tools (100 ME), foundational research to support zero carbon cloud (100 ME), and investments in low-carbon infrastructure (200 ME).

This level of investment is significant, strategic and can be fully justified by the impact expected from accelerated digital transformation across the EU.

[171] AA.VV. European industrial technology roadmap for the next generation cloud-edge offering. May 2021.

A. APPENDIX: PUBLIC CLOUD MARKET SEGMENTS IN SELECTED MARKETS

Market (2020)	Public Cloud Revenue $M (Total)	Revenue $M	Market share %	Growth y/y %	Top 3	Revenue $M	Market share %	Growth y/y %	Top 3	Revenue $M	Market share %	Growth y/y %	Top 3
		IaaS				PaaS				SaaS			
EU - Top 3		5,791	68%	36%	AWS, Microsoft, Google	3,342	49%	38%	Microsoft, AWS, Google	7,092	23%	33%	Microsoft, Salesforce, SAP
EU - Rest of the market		2,677	32%	15%	OVHcloud, IBM, Orange, etc	3,515	51%	30%	Salesforce, SAP, Atlassian, etc.	23,480	77%	20%	Google, Visma, ServiceNow, etc.
EU - Total	**45,897**	**8,468**	**18% of total market**			**6,857**	**15% of total market**			**30,571**	**67% of total market**		
USA - Top 3		7,092	23%	33%	AWS, Microsoft, Google	16,186	56%	32%	Microsoft, AWS, Salesforce	21,627	18%	23%	Salesforce, Microsoft, Oracle
USA - Rest of the market		23,480	77%	20%	IBM, Digital Ocean, Oracle, etc	12,928	44%	25%	Google, Oracle, IBM, etc.	98,244	82%	15%	Intuit, SAP, Google, etc.
USA - Total	**178,702**	**30,571**	**17% of total market**			**29,113**	**16% of total market**			**119,870**	**67% of total market**		
Korea - Top 3		589	85%	40%	AWS, KT, Microsoft	69	43%	32%	Microsoft, AWS, Oracle	188	21%	21%	Microsoft, Ahnlab, SAP
Korea - Rest of the market		100	15%	-3%	Naver, IBM, Tencent, etc.	91	57%	28%	Google, Salasforce, HERE Technologies, etc.	711	79%	19%	Salesforce, Douzone Bizon, Bentley Systems, etc.
Korea - Total	**1,749**	**689**	**39% of total market**			**160**	**9% of total market**			**900**	**59% of total market**		
Japan - Top 3		3,293	78%	29%	AWS, Microsoft, Fujitsu	1,048	52%	31%	Microsoft, Salasforce, AWS	1,662	29%	19%	Microsoft, Trend Micro, Salesforce
Japan - Rest of the market		953	22%	20%	Google, IBM, Oracle, etc.	986	48%	23%	Google, IBM, Open Text, etc.	4,063	71%	20%	Google, McAfee, Zoom, etc.
Japan - Total	**12,006**	**4,246**	**35% of total market**			**2,035**	**17% of total market**			**5,725**	**48% of total market**		
UK - Top 3		1,885	76%	36%	AWS, Microsoft, Google	997	44%	34%	Microsoft, Salesforce, AWS	1,937	18%	33%	Microsoft, Salesforce, Workday
UK - Rest of the market		603	24%	17%	UKCloud, IBM, Century Link, etc.	1,252	56%	28%	Google, MongoDB, Celonis, etc,	8,994	82%	17%	Oracle, Google, Sap, etc.
UK - Total	**15,668**	**2,488**	**16% of total market**			**2,248**	**14% of total market**			**10,932**	**70% of total market**		

Table 4: Public Cloud Market Segment Sizes and Shares, EU, US, Korea, Japan, UK