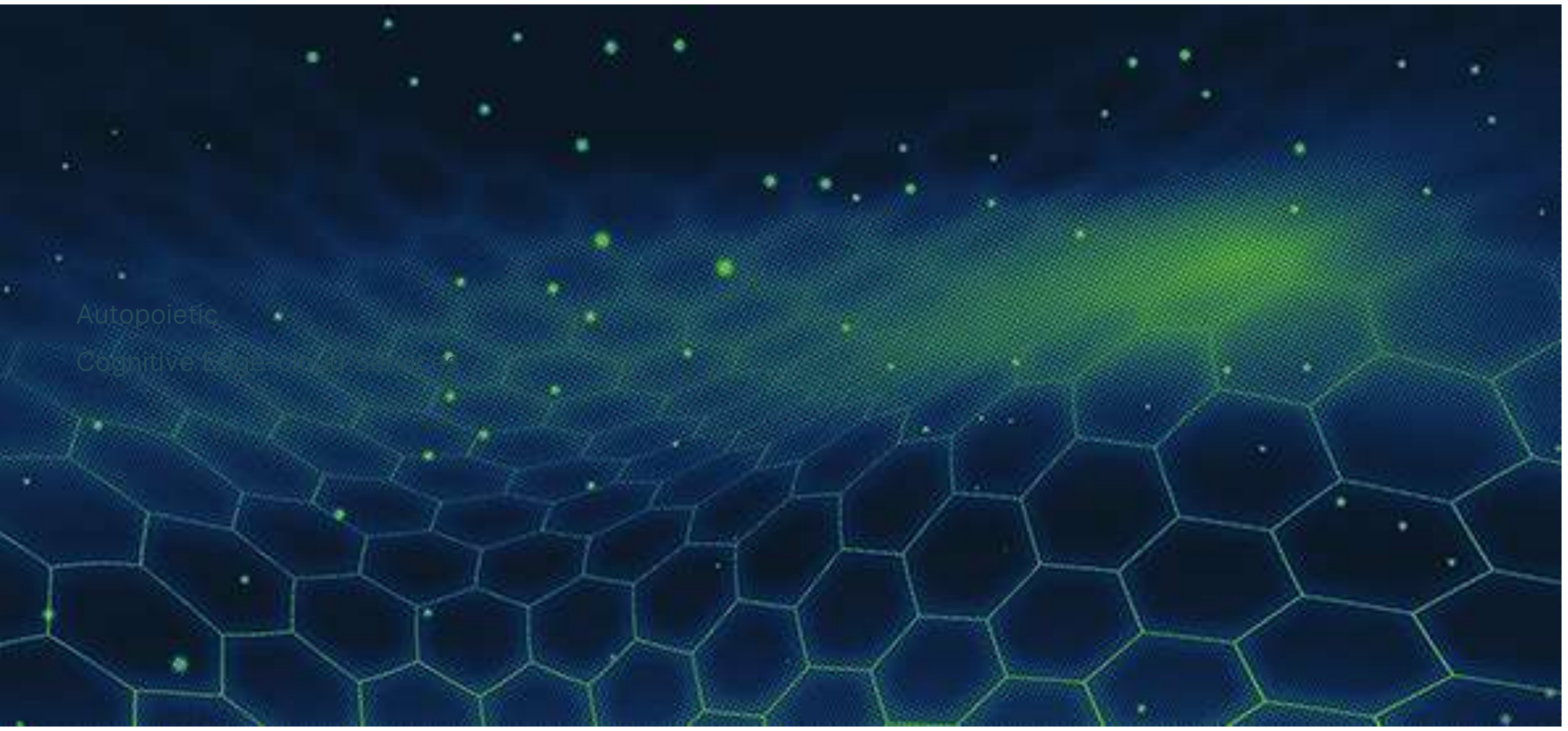
ACES Autopoietic Cognitive Edge-cloud Services

Autopoietic Cognitive Edge-cloud Services

Funded by the European Union



Deliverable D6.6
IPR Management, Business Models, and Business Plan, Exploitation

Grant Agreement Number: 101093126

Autopoietic Cognitive Edge-cloud Services	
Project full title	Autopoietic Cognitive Edge-cloud Services
Call identifier	HORIZON-CL4-2022-DATA-01
Type of action	RIA
Start date	01/ 01/ 2023
End date	31/12/2025
Grant agreement no	101093126

D6.2 – 1st Communication, Networking Plan, Dissemination Strategy Report	
Author(s)	Fred Buining, Luca Alessandro Remotti, Francesco, Mureddu, Giovanni Rimassa
Editor	Fred Buining
Participating partners	DataPower, Martel Innovate, HIRO
Version	Final
Status	Submitted

Deliverable date	M9
Dissemination level	PU - Public
Official date	30 September 2023
Actual date	30 September 2023

Disclaimer

The ACES consortium consists of the following partners:

No	PARTNER ORGANISATION NAME	ABBREVIATION	COUNTRY
1	INSTITUTO DE ENGENHARIA DE SISTEMAS E COMPUTADORES, INVESTIGACAO E DESENVOLVIMENTO EM LISBOA	INESC ID	PT
2	HIRO MICRODATACENTERS B.V	HIRO	NL
3	TECHNISCHE UNIVERSITAT DARMSTADT	TUD	DE
4	LAKESIDE LABS GMBH	LAKE	AT
5	UNIVERZA V LJUBLJANI	UL	SI
6	UNIVERSIDAD POLITECNICA DE MADRID	UPM	ES
7	MARTEL GMBH	MAR	CH
8	SCUOLA UNIVERSITARIA PROFESSIONALE DELLA SVIZZERA ITALIANA	IDSIA	CH
9	INDIPENDENT POWER TRANSMISSION OPERATOR SA	IPTO	EL
10	DATAPOWER SRL	DP	IT
11	SIXSQ SA	SIXSQ	CH

Document Revision History

DATE	VERSION	DESCRIPTION	CONTRIBUTIONS
22/05/2022	1.0	Setup and initial Table of Contents	DataPower, HIRO
13/06/2022	1.1	Completed general sections, partner sections	DataPower, all partners
24/07/2023	1.2	Section on IPR management	MARTEL INNOVATE
10/08/2023	1.2	ACES Blueprint	HIRO
31/08/2023	1.3	Final draft delivered	DataPower/HIRO
25/09/2023	1.4	Full review to the document: changes of editorial nature and comments.	Fernando Ramos (INESC-ID)

Authors

AUTHOR	PARTNER
Fred Buining	HIRO
Luca Remotti	DATAPOWER
Francesco Mureddu	DATAPOWER
Giovanni Rimassa	Martel Innovate

Reviewers

NAME	ORGANISATION
Fernando Ramos	INESC-ID

List of terms and abbreviations

ABBREVIATION	DESCRIPTION
Autopoietic system	A system capable of producing and maintaining itself by creating its own parts.
CLOUD	Cloud computing is the on-demand availability of computer system resources, especially data storage (cloud storage) and computing power, without direct active management by the user. Large clouds often have functions distributed over multiple locations, each of which is a data centre. Cloud computing relies on sharing of resources to achieve coherence and typically uses a pay-as-you-go model, which can help in reducing capital expenses but may also lead to unexpected operating expenses for users.
EDGE	Edge computing is a distributed computing paradigm that brings computation and data storage closer to the sources of data. This is expected to improve response times, security and privacy, and save bandwidth. Edge computing is an architecture rather than a specific technology, and a topology- and location-sensitive form of distributed computing.
KPI	Key Performance Indicators
SEO	Search Engine Optimisation
GA	Grant Agreement
CA	Consortium Agreement
IP	Intellectual Property Management
Swarm Intelligence	Swarm intelligence (SI) is the collective behaviour of decentralized, self-organized systems, natural or artificial. The concept is employed in work on artificial intelligence. SI systems consist typically of a population of simple agents or boids interacting locally with one another and with their

	environment.[2] The inspiration often comes from nature, especially biological systems.
AI, ML	Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind. Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Table of contents

1 Introduction

The ACES project responds to the increased demand for 'cloud-type services at the edge' a.k.a. edge–services, that support the leverage from the rapid growing number of connected devices exchanging vast amounts of data. ACES aims to address the different challenges edge services and architectures have at the edge, such as: i) ability to provide east-west transaction resiliency of applications; ii) reliability and stability of automation in edge services management under increasing complexity; iii) handling of the increasing and latency-sensitive flow of data and applications; iv) explainable AI and transparency of the increasing automation in edge-services platform. ACES will implement solutions by infused autopoiesis and cognition on different levels of edge services management using swarm-based intelligence and AI for workload placement, service and resource management, data and policy management.

The ACES solution provides a powerful and intelligent edge infrastructure that is clearly at odds with the centralized nature of cloud computing. The ACES edge intelligence will be a new type of intelligence that improves scalability, autonomy, and federation of services across a hybrid, distributed, disaggregated edge hardware infrastructure. The edge intelligence will be based on emergent intelligence arising from the interactions of various disaggregated hardware elements and the dynamic demand for a wide range of edge services.

1.1 Edge and Cloud

In cloud computing business models as they are commonly known, compute resources and services are centralized in datacentres to which end users have access. Resources are shared across the network. This architecture has proven to be cost effective and efficient in resource sharing, also considering that cloud service providers install their datacentres in what they believe are optimal locations.

Values in the 5G+Edge global ecosystems[1]

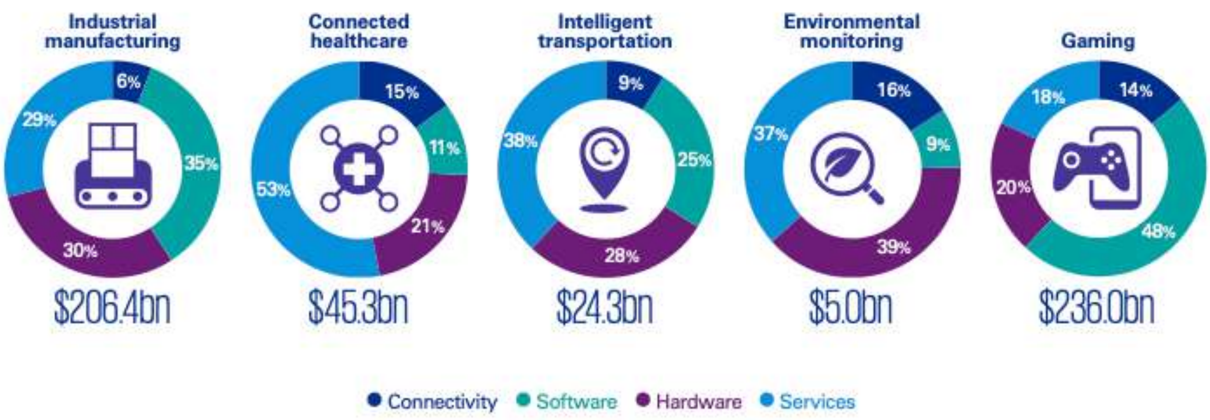


Figure 1: The US$517bn 5G+Edge ecosystem across just five industries. KPMG 2020

[1] KPMG. The 5G edge computing value opportunity. A key pillar in the global economic recovery. 2020.

However, the new approaches to cloud-based applications, the increase of edge-based functions and the expansion of the use of AI/ML-based applications, as well as an increased data protection awareness, are driving a shift of computing and storage power closer to the data sources as well as to the physical processing machines, i.e. at the network's "edge."

KPMG[2] discusses the more than US$500 billion in value across the 5G and edge computing ecosystem via five use cases, enabling telecom industry to emerge from COVID-19 stronger than ever.

Integrating edge computing and mobile computing with IoT (Internet of Things) enables near real-time processes and outputs, warranting less latency and lower cloud server bandwidth requirements, at the same time making available the necessary security layer for sensitive applications and data. Edge computing, compared to cloud computing, offers businesses a more efficient and effective way to process and store data using decentralised enterprise applications and complex systems of sensors.

It needs to be underlined that there isn't always a clear line where cloud ends and edge begins: cloud and edge are related ideas, since processing and storage workloads happen in decentralised spaces, normally outside of what is considered a cloud environment and in proximity of physical devices/the data source. This proximity has a dramatic impact on the performance of latency-sensitive applications, of IoT sensor data processing, and on data-driven real-time decision-making. Services become faster, more reliable and allow organisations and businesses to use and distribute a common pool of resources across a large number of locations that are able to quickly process data without worrying about latency.

However, the edge technologies, such as wireless devices and sensors, lack the necessary compute capacity to process large streams of complex data directly. As a result, smaller, modular data centers are being deployed to provide hyper-local storage and processing capacity at the edge. The opportunity is therefore to supply, build or operate local edge data centers — small micro data centers that are designed to operate near the point of use, supporting applications that are not suited to run in big, remote data centers, even in mid-sized regional colocation data centers.

These Micro Data Centre services provide business operation at a lower cost than traditional data centre services. The modularity and portability of Micro Data Centres and their flexible design allows just-in-time deployment, which may serve as an attractive market growth opportunity. The increasing demand for standardized infrastructure and operability in remote and harsh environments may be capable to drive the growth in the global market.

[2] *Ibidem*.

User Edge, Service Provider Edge and Connectivity: the integrated ACES solution

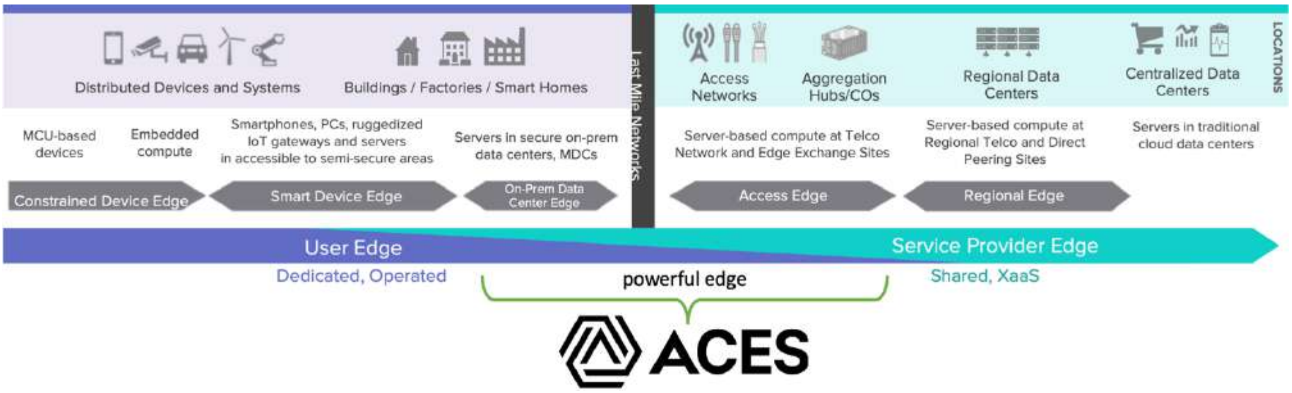


Figure 2: User Edge, Service Provider Edge and Connectivity: the integrated ACES solution. Source HIRO Microdatacenters. 2023.

Unlike most larger data centres, micro data centres will mostly be built, configured and tested in a factory and delivered on a truck. Typical sizes will be 50 kW to 400 kW, and one expects there will be many. Big vendors such as Schneider, Vertiv and Huawei have been rapidly adding to their product lines and redrawing their financial forecasts; startups — companies such as Vapor IO, EdgeMicro, EdgeInfra and MetroEDGE — are pioneering new designs; and established telco specialists, such as Ericsson, along with telco operators, are working on new technologies and partnerships. Also, colocation data center builders and operators, such as EdgeConneX, Equinix and Compass, are assessing where the opportunity lies. Global Market Insights, Inc. estimates that the Micro Data Center market size will grow by more than 23% between 2019–25, growing to $14.5bn by 2025[3].

A new challenge. At present edge-to-cloud architectures have typically a hub and spoke design, in which devices move their data to a distant datacentre (hub) with IaaS and PaaS management systems to manage and process the data with knowledge of workload resource requirements and their availability. The results are then pushed back out to the edge devices. Hyperscale and large datacentres use their economies of scale (incl. overprovisioning) and cognition in workload management, to create stability and performance despite the fluctuations in the demand for, and the availability of, resources.

Edge infrastructures in their current status (FOG, MEC EdgeMicroDataCenters) are more challenged in their stability and performance because of more stringent latency and autonomy requirements, distribution across multiple sites, their local limited size, multi-tenancy and multi-operators, local management, with components being concurrent and asynchronous.

This challenge to edge infrastructures is growing rapidly due to the increasing: - number of connected devices and their data-producing and data-consuming capabilities, - intelligence embedded in edge devices, - atomization of monolithic applications, - scale, speed and complexity of edge device interactivity in a zero-trust environment.

[3] Edge Computing Market - By Component (Hardware, Software, Service), By Application (Industrial IoT, Remote Monitoring, Content Delivery, AR/VR), By Enterprise Size, By Industry Vertical, Forecast 2023 – 2032

1.2 The ACES solution

The Autopoiesis Cognitive Edge-cloud Services (ACES) General Aim is to drive research and innovation activities to find solutions for edge computing as an evolution of cloud computing, designing an edge services cloud with hierarchical intelligence – specifically Autopoiesis and cognitive behaviours – to manage and automate a computing platform, network fabric, storage resources, virtualization, and analytics to increase resilience while managing simultaneous service constraints.

The outcomes of the ACES research will include: i) an autopoiesis cognitive cloud-edge framework; ii) awareness tools, AI/ML agents for workload placement, service and resource management, data and policy management, telemetry and monitoring; iii) agents safeguarding stability in situations of extreme load and complexity; iv) swarm technology-based methodology and implementation for orchestration of resources in the edge; v) edge-wide workload placement and optimization service; vi) an app store for classification, storage, sharing and rating of AI models used in ACES.

The project will demonstrate and validate its functional and technological solutions in three different scenarios requiring highly decentralised computing, and the ability to take autonomic decisions, reducing costs of cloud-edge management, increasing their efficiency, and reducing impact on environment.

Emphasis is placed in developing solutions associated with the specific strengths of the ACES-EDGE framework, namely:

1) end-to-end transaction resilience to fulfil the varying end-user requirements, according to the Quality of Service (QoS), Service Level Objectives (SLO) and Service Level Agreement (SLA); (Challenge 1)
2) autonomy and self-regulating mechanisms that go beyond recursive self-reference and provide systems stability, locally and edge-wide; (Challenge 2)
3) 'horizontal' (east-west) flow of data and applications between sites, and the interactions between data producers and data consumers, in terms of bandwidth, energy efficiency, security, privacy, etc. (Challenge 3)
4) UIX on the platform to deliver transparent information to stakeholders (i.e. operators, application developers, and end-users) on platform operations and platform performance: the relationships between internal and external data-points, machine reasoning and decision making, and - where human intervention should be considered (Challenge 4)
5) deliver low-overhead (costs, latency, energy, labour) security in stream and in real time in a zero-trust environment (Challenge 5).

The autopoietic and cognition functions are of importance at different levels of the workload placement framework, including for service and resource management, data and policy management, to manage their own stability with knowledge of its distributed component deployment and their state of health, along with the knowledge of best practices to deal with fluctuations.

'Cognition' is defined as the ability to acquire and process information, apply knowledge, and autonomously change the inner circumstances to provide better services in response to the fluctuations in the environment. Because cognition creates intelligent self-adaptation up to the same level of complexity as its analysed environment, ACES needs an additional control approach to protect the system from damage or destruction in situations where its cognitive complexity handling capabilities are exceeded.

'Autopoiesis' means self-producing and self-maintaining and the ability to produce more of their own organization principles than the ones produced by analysing the environment (cognition).

ACES ‘Cognition’ refers to a framework that provides AI/ML enhancements based on three concepts. 1) **Awareness**, which refers to (i) resource awareness - enabling the intelligent placement and optimisation of application and data workflows based on primitives, such as resource utilisation, usage patterns and energy consumption, and (ii) context awareness - the discovery and management of resources and data sources to form a coalition for pursuing the joint goals of the system.

2) **Autonomy** refers to self-* properties (with * representing monitoring, managing, organising, adapting, healing, etc.) while being resource and context aware.

3) **Actionability**, refers to the capability to learn from past actions of the Cognitive Cloud and predict suitable courses of action for the future, in real-time and in an automated manner. The most important outcome in ACES is a series of frameworks, toolchains, and runtime mechanisms to harness advanced edge-services, cognitive traits in executing and managing these services, and ease the creation of new cognitive-capable services with greater edge autonomy, reactive and proactive to their environment (e.g., redistribution of workloads, data applications) and/or to functional needs, keeping optimal performance at all times (i.e., fault-free and efficient operation).

This outcome will be the result of combining five interrelated objectives:

- Develop an Autopoiesis cognitive edge-services cloud (Objective 1)
- Research Autopoiesis cognitive edge-services that orchestrate resource and workload management with respect to a wide range of edge-relevant requirements, such as latency, energy efficiency and security (Objective 2)
- Research autopoiesis cognitive edge-services that optimize the data management, storage, replication, movement, and acceleration (Objective 3)
- Develop services (including visualisation and recommendation framework) to improve the experience of operators, software developers, and end-users (Objective 4)
- Test and demonstrate the effectiveness and generality of ACES by evaluating three real-life use cases that exhibit cognitive edge-services with various levels of awareness, autonomy, actionability within the services, the edge-services stack and the hardware (Objective 5).

2 Exploitation strategy overview

2.1 Introduction

Considering the business and technological opportunities of moving cloud computing and storage service towards the edge, the place where data is collected and stored and where the outcomes of their processing are used, ACES goal is to research, innovate and test a new technological solution for edge-cloud services, leveraging the capabilities of Artificial Intelligence and Machine Learning to conceive a self-orchestrating, self-adjusting and self-repairing system offering low latency, resilience, stability, and security.

Even if all major players include edge computing in their offering, the market is still in development, not only in respect to the core concepts the ACES project is exploring, but no structured investment in service offering is currently observed. The different building blocks and the technological solutions to enable them are still explored and investment is still assessing the actual dynamic and pace of “movement towards the edge”, which depends on the investment strategies of edge customers.

The exploitation approach of the ACES research and innovation project takes a “spiral” approach, starting with the configuration of ACES in its entirety and of its building blocks and key exploitable results. The purpose of the exploitation is to strengthen Europe’s position in the cloud and edge services and establish the necessary innovation capacity for EU Policy makers to support this process.

The current exploitation plan intends to be the first version of a regularly updated document, which will evolve according to

a. The consolidation of the technological solutions;
b. The development and the maturity of the market and of customer demand;
c. The evolution of the general cloud scenario and of the overall economic, industrial and social environment.

The exploitation will include a scaling up strategy accompanied by a strategic roadmap for the exploitation and long-term sustainability of the project results.

It is based on a set of key pillars:

1. The Assessment of the resources and assets of the ACES project, the knowledge base generated, and the set of solutions summarised in the ACES-EDGE matrix, where the service and its elements are mapped and described. This matrix puts into relationship the different ACES-EDGE configuration with different types of users from different sectors. This matrix is completed via user segmentation and interviews and focus groups to construct typical use cases.
2. The assessment of the market and stakeholder needs, the requirements of target users and markets for the ACES web services, with primary reference to public administrations, mapped in the target user matrix, such user matrix being the main guidance to associate the ACES-EDGE and its configurations with specific user types.
3. The mapping and description of technological solutions, for the overall ACES concept and specifically for the single system components, i.e., the autopoietic enablers, the AI/ML systems, the swarm computing algorithms, etc.;
4. The possible ACES use cases to demonstrate the applications of data collection and data processing at the edge, mapped in the use-case/functions/performance matrix;
5. The market analysis, assessing the trends in the edge-cloud market, the technologies and service developments;
6. The analysis of the competition;

7. Several market scenarios, depicting ACES solutions associated with use cases, a matrix identifying the key market drivers affecting the development of edge-cloud services, starting from the requirements, and then moving on towards the analysis of more general economic aspects affecting the edge service deployment.

The exploitation will dynamically interlink the seven pillars, and the relevant matrixes, logically and concretely connecting the characterisation of the ACES service and of its elements with the different types of users, leading to a match of the core resources and assets with target stakeholders and identification of potential strategies to reach each type of client and the type of product/service. The matrix will relate the ACES functions with the performance characteristics: latency, QoS, energy consumption, autopoietic capability and performance of the AI and ML in orchestrating, managing, and self-correcting the edge. One important part of the matrix concerns the relationship between the distinctive functions and the relevant performance parameters associated to them.

The analysis will produce the ACES business model based on the Exploitation plan, which will be further deepened and detailed in the revisions over the 36 months of the research and innovation project and beyond. This aspect will be carefully assessed and planned, focusing on the vision of the use of the project outcomes after the end of the project funding, together with the business approach that will be implemented to maximize the impact of the ACES project and guarantee its sustainability in the medium-long run.

The future exploitation and sustainability model of ACES is rooted in the initial design of the project that facilitates post-project market uptake of results, also considering the major environmental, climatic, and societal challenges.

2.2 The structure of the ACES exploitation plan

2.2.1 Introduction

In this section we provide an overview of the spiral approach to designing and implementing the exploitation activities of ACES. Its key pillars were described in the previous section, namely the assessment of ACES assets; the scouting and characterisation of the prospective ACES market; the mapping of technological solutions; ACES use cases and their validation; market and competition analysis; and the definition of the different market scenarios.

In the next sections we describe:

a) The ACES results mapping and analysis;

b) The key exploitable results;

c) How the consortium members claim ownership of the key exploitable results;

d) Market trends and gaps, analysing the maturity of the edge-cloud concept as it is nowadays, as well as the general estimates of edge market development;

e) The ACES business model and business plan, in relation to the current development status;

f) An initial overview of IPR and licensing issues, preliminarily to the comprehensive negotiation in view of the concrete exploitation by the ACES consortium partners, covering the management of results and the dissemination strategies as well as the access rights;

g) The joint exploitation by the ACES consortium, and the individual exploitation plans.

We emphasise again that this exploitation plan is an ongoing development and that it will be updated in follow-up deliverables. As a first step, we plan to undertake specific validation and scenario-building actions to improve the competitive positioning of ACES as a whole, and of its components.

2.2.2 ACES results mapping and analysis

2.2.2.1 ACES value proposition

A large portion of applications, particularly enterprise apps, leverage a heterogeneous network of public clouds, private clouds, on-premises data centres, and edge. Edge as a Service (EaaS) providers are "companies that offer a platform realizing distributed cloud architectures and integrating the edge of the network in the computing ecosystem."[4]

Currently, Cloud providers offer EaaS as a hub-and-spoke model, meaning leveraging on the cloud hosting space and from there offer the service to easily scale and quickly provision new resources (virtual machines, databases) to hundreds, or even thousands, of endpoints across the globe. Since their business models are based on economies of scale in hyperscale data centres, their services will always be driving the centre of gravity to the hyperscale data centre; the location where data is gathered from edge devices, analytics performed, decisions made, AI trained, applications and AI deployed to edge devices. This business model fuels a coarse-grained infrastructure of large data centres around the world, that require enlarged support infrastructures of data transportation, energy supply, water supply and waste heat removal, that often exceed the local supply capabilities.

The ACES consortium believes that a more fine-grained edge mesh infrastructure based on powerful microdata centres with local autonomy over data and processes supported by federated east-west collaboration throughout the mesh has advantages over the cloud hub and spoke model and eliminates some of the disadvantages of the cloud. To operate this as a viable business of an 'Edge as a Service', a wide range of innovations is required, from hardware, software, operating models, etc. While some of these aspects are being researched in other (current and future) projects, the ACES consortium focusses on developing *the software heart* of such an infrastructure while anticipating the disaggregation of hardware that is expected the coming decade: *the orchestration software that virtualizes the hardware and regards them as pools of resources operating jointly under autopoietic principles*.

Autopoietic[5]. Autopoiesis originates from biology and is derived from the Greek words "auto" (for self) and "poiesis" (for creation or production), and it refers to a system's ability to self-organize, self-maintain, construct its own reality. Unfortunately, technology evolution and innovation is becoming increasingly self-referential and self-programming, thereby isolating and immunizing itself from other sectors of society. Technology can be conceptualised as an autopoietic system, recognizing the recursive nature of technology and its co-evolution with the social and physical environment.

ACES adopts the view that, with help of various forms of AI, technology can become a dynamic self-referencing system that continually evolves and adapts in response to human and societal dynamics. By recognizing the reciprocal relationship between technology and society, we can strive to create technologies that are more responsive to human needs and values, and that contribute to a more equitable and sustainable society. ACES autopoietic approach to edge computing technology has important implications for the design and implementation of AI technologies as laid out in the ACES blueprint for orchestration, which is in first instance described in Section 2.2.3 as well as in the forthcoming deliverable of WP2.

[4] https://www.edgeir.com

[5] *Technology as an autopoietic system*. Available from: https://www.researchgate.net/publication/372076687_Technology_as_an_autopoietic_system [accessed Jul 30 2023].

The value proposition of ACES can be defined by the following principles:

Unity

Ability of individual hardware parts and software services, working together in a unified way, to achieve shared goals in ways that were not explicitly programmed.

Regeneration

Ability to dynamically learn from its own behaviour in the context of its environment, and adapt to changes in the environment to become more resilient and better able to handle future events.

Balanced complexity

Ability to manage various interconnected parts, whose interactions are difficult to understand (hence complex), in a way that it maximizes its benefits while minimizing its costs.

Regulation

Ability to comply with environmental conditions regarding security, privacy, safety, reliability, interoperability, waste, and efficiency.

Living system

Ability to self-organize and maintain itself through interaction with the environment through information exchanges.

Embodiment

Ability to grow a body of knowledge between the environmental conditions and its responses to those conditions through awareness, inference, and action.

The above-mentioned principles allow edge infrastructure to be highly flexible and adaptive to the local conditions, autonomously operating to its best and most efficient abilities, and to use the locations of the servers to kick-in when conditions exceed local capabilities.

2.2.2.2 Functional Specifications

In ACES the powerful edge is created from a mesh of Edge-Microdata Centres (EMDC's) that are completely composable, customizable, heterogeneous, and therefore different in capacity and capabilities. The high-level functional specifications for ACES and the operations of a powerful Edge are:

On platform level

- Edge autonomy: offloading applications and data from edge devices (improving the autonomy at the edge), reducing latency, improving energy efficiency and security, reducing costs.
- Distant Cloud connectivity: placing applications (incl. AI) and data closer to the edge (improving distant cloud efficiency);
- Networking and interoperability: supporting the edge-to-edge-to-cloud continuum by creating the required networking connections and interoperability services.

On EMDC level

- local autonomy; meeting the demands of the local edge clients.
- federation: meeting the demands upon groups of EMDC's caused by fluctuating local demand and mobile demand traveling along several EMDC's in a region.
- Networking and interoperability: facilitating east-west traffic locally without the need of a distant cloud.

This translates into an *autopoietic framework* that needs to address the following fine-grained functional specifications:

Awareness; aware of its local and regional:

- hardware configuration;

- hardware capabilities and availability;

- software services (supply);

- local and regional requests for services (demand);

Inference, self-intelligent in:

- matching supply and demand;

- efficient execution of its services locally and in federations with locations in the region;

- efficient in organisation and execution of self-intelligence;

- effective in interpreting ad hoc, mid-term, long-term a) fluctuations in demand and b) changes in the local and regional conditions (network, energy, etc.).

Action, capable of actuating by:

- self-activation based on local and regional conditions;
- change of metrics collection, parameters, agents used, and market mechanism setting for demand and supply that perpetuates the efficient matching of demand and supply;
- execute placement and movement of data, containers, applications, etc.

2.2.2.3 SWOT Analysis

ACES-EDGE uses the SWOT Analysis (strengths, weaknesses, opportunities, and threats) as a framework to evaluate the market positioning of the ACES Edge service. Through our SWOT analysis, which is a developing undertaking and will take advantage of the ongoing research on the market development on the ongoing ACES autopoietic solution.

Our SWOT analysis will progress from the inside to the outside of the ACES solution, starting from the technology and innovation and testing the external requirements through the development of specific use cases, first of all those in the energy sector planned and then linking the other application areas focused on in similar Edge projects clustered in the EUCloudEdgeIoT EU initiative studying the cloud-edge continuum.

Our SWOT analysis is progressively designed to reach a realistic and fact-based assessment of the ACES EaaS, considering internal and external factors, guiding the system's marketing process by bridging the TRL levels, which separate it from the full market deployment.

The full ACES SWOT analysis will involve the demand side to match the technical and operational solutions of the Edge platform, ensuring that user requirements are properly taken into account.

Strengths

Strengths describe where ACES-EDGE will excels at and how it will perform against traditional cloud services and competing EaaS platforms.

- *Flexibility* and scalability
- *Increased performance* and reduced latency
- *Outsourcing:* Some enterprises expect to own their own edge micro data centres, according to Uptime Intelligence research, but many others will likely prefer to outsource this altogether, despite potentially higher operational costs and a slight loss of control. This supports the uptake of EaaS, of which ACES is an example.

Weaknesses

Weaknesses are those factors, which will impede the ACES-EDGE performance at its optimum degree.

- *Maturity:* EaaS as a distributed mesh is lacking the maturity that hub and spoke EaaS offerings from large cloud providers have. These issues include licensing, location, and ownership of sites; support and maintenance; security and resiliency concerns; and management sentiment.
- *Uncertainty about demand:* many edge applications — whether supported by 5G or some other networking technology (such as Wi-Fi 6) — may not require a local micro data centre. For example, high-bandwidth applications may be best served from a content distribution network at the regional edge, while many sensors and IoT devices produce very little data and so can be served by small gateway devices. In fact, among 5G's unique properties is the ability to support data-heavy, low-latency services at scale.
- *Uncertainty about scale of demand:* Suppliers and telcos alike, then, are unsure about the number, type, and size of data centres at the local edge.
- *Need for end-user investment in their local systems,* which will not be required by centralised cloud systems.

Opportunities

Opportunities analyse and assess external factors that have the potential to give ACES EDGE a competitive advantage.

- More control over data and their location, certainty about government access.
- Better compliance of data protection regulations and the EU ones.
- More flexible and effective handling of security of local microservers.

Threats

Threats are those factors that potentially could harm or even hamper the development and market deployment of the ACES EDGE. These can be internal or context elements.

- *The slow roll out of 5G:* 5G will ultimately support a vast array of new always-on, low latency and immersive applications that will require unimaginable amounts of data and compute power — too much to send back realistically or economically to the internet's hyperscale core. Much of this will require low-latency communications and rapid processing of a few milliseconds or less — which, the speed of light dictates, must be within a few kilometres.

In the next steps of the project exploitation activities the ACES Team, with the involvement of a set of selected stakeholders and experts, will deepen and broaden the SWOT analysis as the project and the Edge computing solution and platform become more mature and operational.

2.2.2.4 PESTLE Analysis

The PESTLE strategic analysis focuses on the assessment of political, economic, social and cultural, technological, environmental and legal factors, which concern and affect a business or a system, its performance, development and sustainability.

In the context of ACES, the PESTLE analysis determines the macro factors of context and their dynamics to be considered when designing and operating the system. It identifies framework elements, which are likely to affect market prospects, competitive positioning, operations potential and direction, and therefore are vital for the definition of relevant measures and decisions.

It thus considers six classes of factors:

1. Political, considering the regulatory and government policy framework which is likely to influence the setup and operations of the ACES Edge service, including technical regulations, data protection regulations, environmental regulations, industrial policies and service and product regulations, the general level and availability of human resources;
2. Economic (macro), including the general economic growth perspectives, the growth of the IT sector and one of the target sectors of the ACES Edge service, the financial scenario: availability of resources, interest rates, investment and venture capital dynamics, other general economic and financial variables such as exchange rates and constraints to capital flows, and business demographics;
3. Social and cultural, considering technology awareness, innovation propensity, risk propensity and health consciousness, population growth rate, cultural aspects, life, career and leisure attitudes that can affect availability of professional workers and recruitment;
4. Technology: prevailing trends in R&D and innovation, the rate of technological change, the quality and diffusion of the innovation infrastructure, the level of automation and the customer target base for the ACES Edge services, the ensuing barriers to entry in the sector and the strength of the competitive position of incumbent cloud and edge service providers;
5. Legal issues: technology law, consumer law, employment law, health and safety and all regulations which affect development, operations and performance of the service provision, and the way the ACES Edge business enters the market;
6. Environmental include ecological and environmental elements such as energy consumption, climatic impact, and considering environmentally sustainable energy solutions.

A first PESTLE analysis

At this point in time the PESTLE analysis is at an early stage of implementation and the exploitation concept must be framed carefully to define the steps that lead the ACES technological Edge solutions to the market.

The exploitation strategy shall most likely define the likely configuration of the ACES Edge service at the end of the EU-funded Research and Innovation Action and then define the next steps for a possible market entry. We do not expect to be able to make a direct market entry with a service, but ACES will target a full configuration of the Edge computing system. From there ACES will develop the path towards the market, not only in respect to service configuration, but even more important, developing use cases to support a vertical – sector-specific – cost-benefit analysis to propose “sales cases” to prospect customers.

In this first ACES exploitation deliverable, the focus is on the critical areas. In the next versions of this Report the core building blocks will be defined and also the steps leading to market exploitation.

We present now a first assessment of the PESTLE factor classes for the ACES Edge system, consistent with the current state of development of the ACES solution:

Topic Area	Discussion
Political	Climate goals: with outsourcing EaaS to a distant cloud, the realisation of climate goals through innovation in the www infrastructure lies in the hands of a few large cloud providers. They currently self-regulate their efforts in reaching those goals. That process is slow, for example lowering the PUE through innovations has come to a standstill. Because of their economies of scale, any innovation in their infrastructure is weighed very carefully against all other business objectives. For companies, and cloud end-users, implementing climate favourable innovations in

	their ICT infrastructure is not affecting substantially their bottom-line business. Therefore, incentives to drive cloud computing back to the edge/on premise, supported by regulation (PUE targets) will accelerate the adoption of climate favourable technological solutions and favour the companies developing these technologies. effects. The EU Framework Programme stimulating edge computing (Meta edge operating system, IoT, cognitive cloud), the AI act, the Data act, Gaia-X, dataspaces, SIMPL, are all supportive activities on EU level. What is missing is the taxation of less favourable technologies currently being used by large cloud providers and offered by OEM's. This pathway is a delicate pathway between on the one hand building technologies that climate conscious and democratic societies will embrace and on the other hand competing with technologies developed outside these societies and risking isolationism from the Computing Continuum. International cooperation becomes difficult in the new geopolitical climate The success of the ACES project may influence policy makers into establishing such policies.
Economic	The Computing Continuum is monopolised by a few non-European hyperscaler providers. Lacking industrial adoption due to high concerns regarding data sharing and costs. The OPEX of decentralized cloud-edge solutions is significant, due to their inherent management complexity. ACES aims to reduce OPEX costs by introducing AI-powered automation and to support the establishment of an open ecosystem based on fair and transparent resource sharing
Social	Companies and end-users are reluctant to adopt novel cloud and edge solutions due to reduced trust. ACES aims to increase trust of end-users and ecosystem participants by employing transparency and fairness by design. There is confusion and lack of trust in cloud computing and in AI and ML for system management. For Edge computing, skill gaps are an even bigger Computing Continuum barrier then for Cloud and IoT.
Technological	Software/hardware heterogeneity hampers adoption across industries and countries. Security is not strong enough for heterogeneous device networks in the Computing Continuum. Current state-of-the-art solutions (e.g. Kubernetes) do not employ ML and offer limited capabilities regarding resource sharing across different organizations, which is further complicated due to security issues arising from data transfer across organizational domains. ACES addresses these issues by introducing ML capabilities in the core Kubernetes services and by ensuring secure resource sharing throughout the cloud-edge continuum even when data is transferred between different organizations.
Legal	A single European digital market is hindered by different legal frameworks of EU member states, thus increasing the complexity of comparison between services provided by different cloud providers. The ACES solution will partially address the complexity issues arising from data movement, by minimizing data transfer across the continuum and it will enable

	resources benchmarking, thus facilitating comparison of services provided by different providers. New regulation on AI stifles market entry of Cognitive Computing Continuum applications. Lack of standardisation and certification disrupts the formation of a European supply chain.
Environmental	Information and Communication Technologies have a significant environmental impact, namely on CO_2 emissions of data centres. ACES will pursue the optimized use of spare resources, which in turn will result in reduced CO_2 footprint. The Computing Continuum applications cannot achieve the needed energy efficiency at the moment.

2.2.3 Key Exploitable Results

This section provides a short description of the Key Exploitable Results of the ACES-EDGE system. It also includes an overview of the exploitation options of all its elements, as well as of the integrated system, together with the map of potential users.

Key Exploitable results

First, an open source and modular architecture, supporting virtualisation and multiple architectural patterns for ad-hoc edge- clouds in one site, and across multiple sites; an event-driven, data-centric architecture with high levels of automation and autonomy, but also supporting human operator control. The architecture will have no single point of failure, achieve high throughputs, and offer horizontal scalability (by offering edge resources at runtime without disrupting ongoing operations in terms of communication, processing, storage, and monitoring).

These exploitable results will be:

- A range of coherent edge-services creating an edge cloud, including:
    - intelligent automated orchestration of edge-services (computation, acceleration, storage, networking) in support of the platform (operator);
    - visualisation and recommendation framework to improve the experience of operators, software developers, and end-users.
    - functioning asset monitoring and management (operator, end-user)
    - single pane of view to understand the efficiency of the application deployment on ACES (software developer).
- A cognitive framework driving awareness, autonomy and actionability tools;
- A distributed Knowledge Base;

Second, a cognitive autopoietic framework with: *Awareness*: collecting valuable data across the infrastructure and abstracting relevant knowledge stored in a distributed knowledge base (DKB) and accessible for cognitive agents; *Inference:* creating high level understanding of the context and the operations within; *Actionability:* cognitive agents that are based on knowledge stored in the DKB act and execute a range of microservices.

These exploitable results will be:

-General Collective Intelligence (GCI): a schema including several propositions that determine how subagency functionalities (cognitive agents) can generate collective fitness through cooperation that seeks to create Autopoiesis efficacy by satisfying systems needs of welfare and fitness sustainably.

-A network of processes decomposed into a pragmatic set of contextually related functional task processes, each undertaken by a coalition of cognitive agents (subagencies) in pursuit of collective goals serving the collaboration between sites and overarching local lower-level individual and sometimes conflicting goals.

-UIX's with an informative single pane of view for ACES operators, application developers and end-users with relevant knowledge regarding their role and objectives, and tools to act in the platform.

Third, an intelligent energy grid framework with applications to:

- Operate a functioning local flexibility market (end-user); a
- functioning local energy management system (end-user); and a
- single pane of view to understand the efficiency of the application deployment on ACES (software developer).

In the next subsections we detail the ACES framework blueprint that materialises these results.

2.2.3.1 Edge-services cloud with autopoiesis and cognition

The ACES capabilities are manifested through its stack of edge services. Many companies develop a full stack of services based on open source, partly or full proprietary customizations for Cloud, Fog, Edge services. Their stacks stem from the wide range of available open source services and tools combined with proprietary developments. The range of services and tools is so overwhelming to the extent that the Linux foundation has started in 2021 a working group to work on consolidation and a common stack layout. ACES builds on a stack defined by HIRO-MicroDataCenters and their experience from among others the Braine project[6].

The high-level architecture is displayed in Figure 3: The ACES high-level architecture and consists of the following services:

ACES frontend: providing a single pane of glass to end users and operators.

Service manager: group of software components used to effectively manage the storage, definition, catalogue, and instantiation of workloads (also known as containers). It also manages the QoS for workloads through management of the underlying infrastructure, with supporting components to manage and control EMDC hardware and network infrastructure, including:

- SLA/SLO;
- QoS (Quality of Service);
- Service catalogue;
- Data dependency;
- Privacy & Security.

[6] The BRAINE approach is based on creating a seamless Edge by building clusters of nodes, forming an Edge MicroDataCenter (EMDC), along with smart network interface cards (NIC) capable of pairing with Edge devices to enable high speed connectivity and native AI. The impact of BRAINE encompasses advances in the European video distribution ecosystem, by developing new edge computing architectures and subsystems, improving data processing at the network edge, by supporting scale-up and scale-out approaches. https://www.braine-project.eu

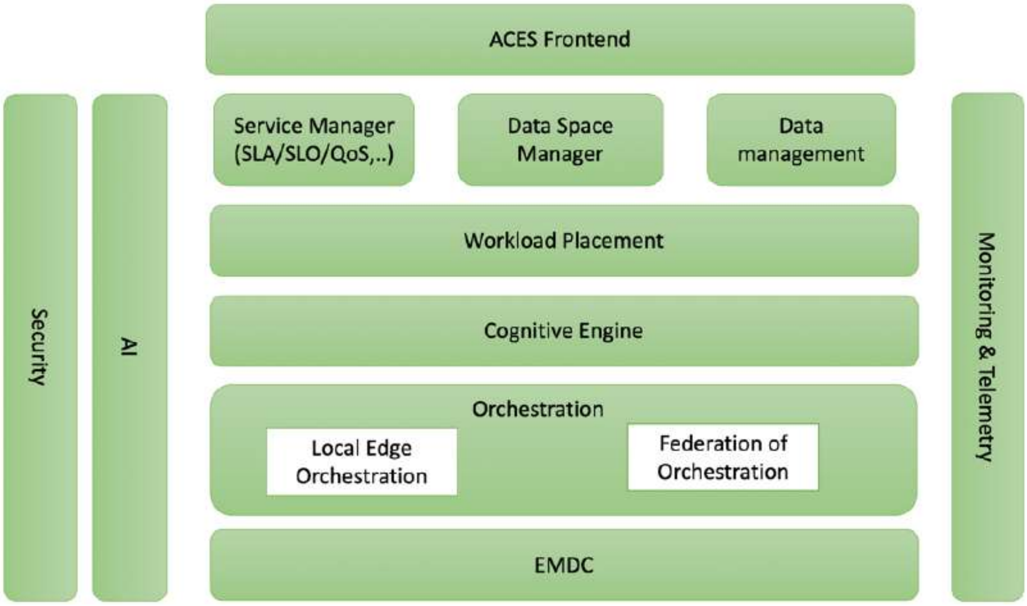


Figure 3: The ACES high-level architecture

Date space manager: providing a secure and privacy-preserving infrastructure to pool, access, share, process and use, and monetize data (datasets, data-products, AI models). It has a clear and practical structure for access to and use of data, training of AI, all in a fair, transparent, proportionate and/non-discriminatory manner with clear and trustworthy data governance mechanisms. It enforces the European rules and values, in particular personal data protection, consumer protection legislation and competition law. The data space manager will assign and acknowledge a wide variety of roles relevant to operate in data spaces, through role-based access management. The Data space will have a store feature that facilitates the marketplace aspect of the data space, a tool to make data available against compensation, including remuneration, or for free.

Data manager: range of software components to underpin the storage and usage of workload and EMDC data on the platform, monitoring and enforcing the management of data through data policies and lifecycle management. It includes:

- DKB (Distributed knowledge base);
- Policy management;
- Data lifecycle management.

Workload placement: is the positioning of workloads (demand) on resources (supply) within the EMDC platform. The decisions in matching demand and supply are made dynamically and ad hoc through a decentralised process of agents that represent demand and supply and obey meta-heuristic algorithms (incl. swarm technologies) that provide a good enough solution through for example some form of stochastic optimization. Meta-heuristics are lightweight programmable algorithms that solve complex problems within a short execution time and provide a good enough solution.

The many independent agents in a meta-heuristic algorithm makes it suitable for parallelization. However, the use of meta-heuristics has a few disadvantages which need to be balanced with the advantages:

- Communication overhead: in parallel implementations, the communication and data transfer between many agents can be intensive and create an overhead that penalizes the execution

of the algorithm and consumes more time. The communication steps require all participants to stop their individual movements and start the communication coherently.
- Lack of explainable behaviour of, and control over, the meta-heuristic: swarm behaviour is emergent group behaviour based on simple rules and agents. Extrapolation from the swarm agents and rules do not provide a predictable behaviour.

To manage the above mentioned challenges the meta-heuristic is designed in a way that it creates minimal communication overhead. By adding non-swarm AI/ML optimization algorithms and iterative methods (in the cognitive engine) we create feedback and a certain level of control on the swarm behaviour. This will also enable some level of explainability of the swarm behaviour and a mechanism to understand and influence the swarm behaviour in a logical manner. These functions will be implemented in the cognitive engine. Note: This does not mean we will be able to create full predictable behaviour and control of the swarm emergent behaviour since that would contradict the concept of emergence.

Cognitive engine: the elements of the Cognitive framework include the distributed AI, ML training, configuration, prediction, validation, and monitoring models for the complete platform and for the autopoietic framework. These AI/ML models will be integrated in the individual hardware modules across the EMDC as CXL becomes more available on hardware component level and implemented across the EMDC platform. [7]

Orchestration: divided into two levels. The first level is tuned to the local custom configuration of the EMDC. It anticipates the future full capabilities of CXL disaggregated infrastructure in which storage, memories and CXU cores can be pooled independently and ad hoc for certain workloads (local orchestration). The second level is the federated orchestration in which local autonomous EMDC's create federations with neighbouring EMDC's to cope with local and regional workloads.

AI: AI functions/services will be developed for the ACES platform to make automatic intelligent decisions for self-optimization and auto healing, while creating federations with neighbouring EMDCs. To create this neighbouring view, the EMDCs will support different forms of federated learning across their regional EMDCs.

Monitoring and Telemetry: The telemetry and monitoring system section represents the components used to collect information to be analysed and processed for improving other software and algorithm functions, and to visually represent telemetry data to an administrator.

- Metrics information & filter
- Metrics analyser
- Metadata database
- Data collector
- Dashboard

Security: will have a range of services to identify anomalies in, and mitigate the poisoning of the data, used by the AI/ ML and swarm technology, and secure the AI algorithms/models. The security approaches will utilize AI for security enforcement, network and system configuration.

[7] Compute Express Link (CXL) 3.0 Debuts, Wins CPU Interconnect Wars. https://www.tomshardware.com/news/cxl-30-debuts-one-cpu-interconnect-to-rule-them-all

2.2.3.2 Framework for autopoiesis and cognition

The complete framework for autopoiesis, cognition and interaction is described in this Section. Please refer to Figure 4: The ACES-EDGE

State of the platform: the state of the platform is collected via metrics (1 and 6 in figure) from the incoming **demand** for resources from the workloads and the **supply** availability of resources (occupied and available).

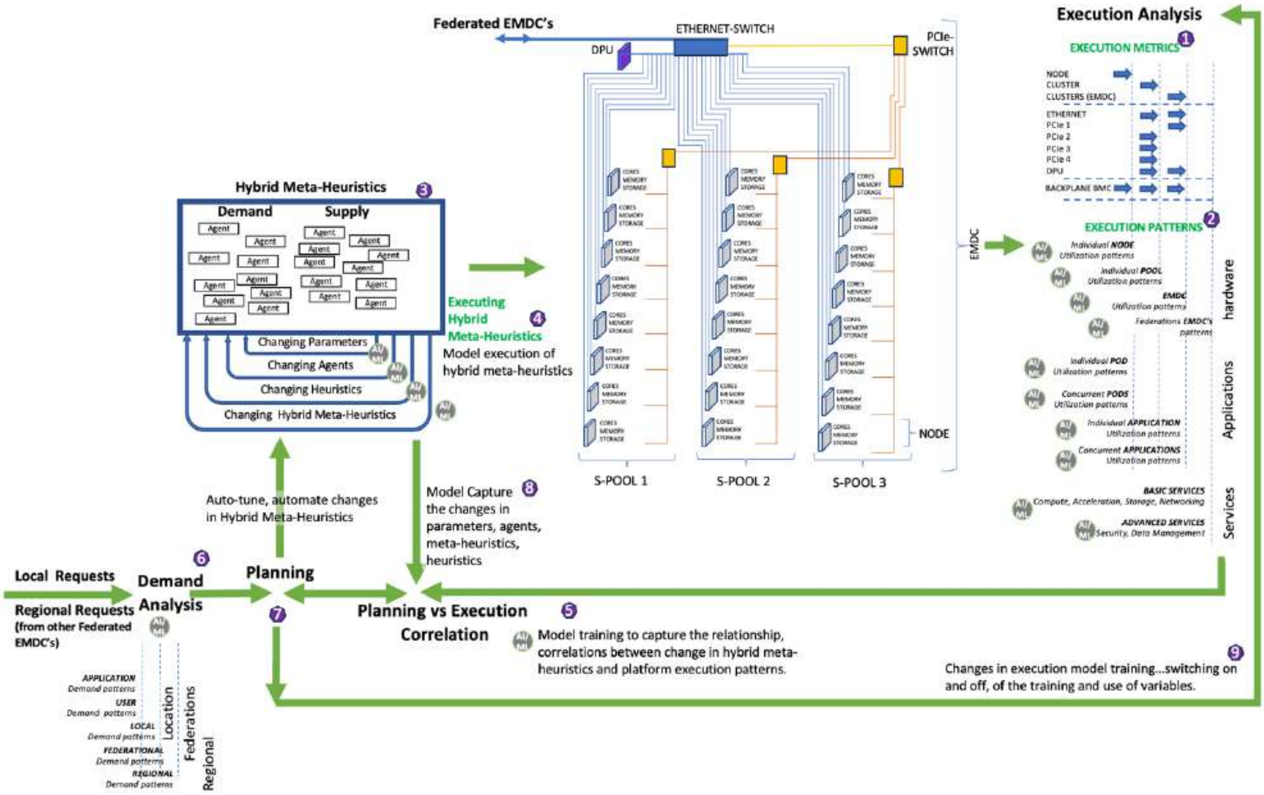


Figure 4: The ACES-EDGE autopoietic framework.

Pattern learning, prediction, and recognition (2 in the Figure): The metrics are analysed and turned into execution patterns (for platform *awareness*):

- Hardware patterns: patterns that describe the various patterns in hardware utilisation on: - node level (CPU, memory, storage), on pool level (e.g., pools of servers, memories, storage, accelerators), bandwidth on EMDC level and on federation level.
- Applications patterns: patterns that describe the various real utilisation levels on the hardware (which differs from the resources the application demanded at entrance) on; - pod level, - concurrent pod level, application level, concurrent application level.
- Service patterns: patterns that describe the utilisation of various services; on basic service level processing, storage, acceleration, networking) and on advanced services (security, data management, etc.).

Hybrid Meta-Heuristics algorithm for executing the matching of demand and supply (3,4, and 8 in the figure): A hybrid algorithm that mixes swarm technologies with other AI, ML technologies to execute the dynamics within the 'marketplace' where demand for resources meets supply of resources (3). Executed by agents (4) that drive the actual ad-hoc deployment and scheduling of micro-services (selection(s) of pods) from applications. The swarm behaviour of demand and supply emerges from the interaction of agents that is defined by their parameters. The various degrees of agents, parameters and algorithms are captured in one or more models (8) of the hybrid meta-heuristics and models.

Capturing correlations (5 in the figure): between the level in which SLA, SLO, QoS requirements of operator (of the platform) and end-users (applications owners) is met, with the settings in the hybrid meta-heuristics. AI/ML tools are used to understand the relationship between the settings of the hybrid meta-heuristic, their changes, and their contributions to meeting demand.

Planning of the changes (7 in the figure): in the hybrid meta-heuristics to meet current and predicted demand patterns.

Finetuning of metrics and patterns collection (9 in the figure): Depending on the behaviour of the platform and the incoming demand the collection of metrics and patterns is scaled up or down in terms of granularity (more or less, variations of the timescale, etc.).

2.2.3.3 Autopoietic blueprint

We now turn to the autopoietic blueprint. Please refer to Figure 5: The ACES-EDGE Autopoietic Blueprint.

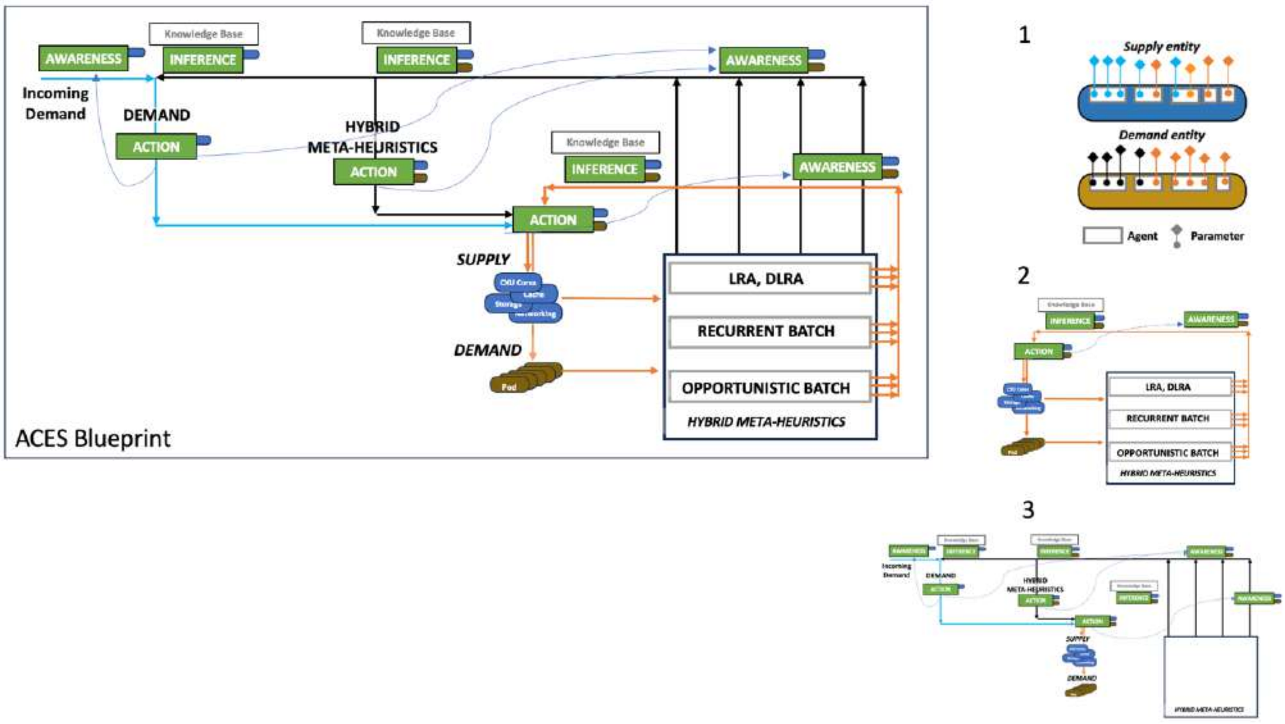


Figure 5: The ACES-EDGE Autopoietic Blueprint

In the ACES autopoietic framework the demand and supply are broken down into entities, representing demand and supply (see item 1 in the figure). The demand comes from the user, application, and pod, and the supply includes CXU cores, nodes, caches, storage, and networking. Their state and their individual behaviour are characterised with the help of agents and their parameter(s). The number, type of agents and their parameter settings are defined based on the:

- Intrinsic behaviours of the entity: For demand, those are the intrinsic behaviours for each of three types of workloads: Long Running Applications (LRA) & DLRA, recurrent batch, and opportunistic batch. For supply, those are the intrinsic behaviours of CXU cores, memory, storage, and networking.

- Desired behaviour of the entity within the hybrid meta-heuristic of a workload type (LRA & DLRA, recurrent batch, opportunistic batch).
- Desired behaviour of the entity within the platform as a whole.

The actual behaviour of entities, groups of entities is continuously monitored (**awareness**) and analysed (**inference**) on two levels:

- the hybrid meta-heuristic where the entity is deployed in (item 2 in the figure).
    - the platform as a whole (item 3), namely across the intrinsic categories that the entity is part of, and across the different hybrid meta-heuristics that are active on the platform.

The agents of demand and supply entities are auto-configured (**action**) to optimize the utilization of the platform, match Service Level Agreements (SLA), Quality of Service (QoS), meet KPI's.

Each type of the three identified workloads (LRA & DLRA, recurrent batch, opportunistic batch), will have one (or more) hybrid meta-heuristic(s) to choose from. Each hybrid meta-heuristic, when active, operates on a local EMDC and can stretch to neighbouring EMDC's (federation) to the extent that neighbouring EMDC's have created an ad-hoc or enduring federation.

Each hybrid meta-heuristic will match demand (users, applications, pods) and supply (CXU cores, memory usage, storage, networking) through cycles of *awareness* (collected metrics), *inference* (created intelligence) and *actions* (changes in behaviour of demand and supply), which are dynamically customized.

The application store will be used as a repository to store the hybrid meta-heuristics, the agents, their parameters, and meta information that describes and characterizes their performance and the conditions within which these performances have been achieved.

2.2.3.4 Uptake of the Blueprint in the WP's research

All technical work packages will revolve around the three functions in the blueprint – *awareness, inference, and action* – and the dynamic relationships between the three functions to facilitate autopoietic responsiveness of the powerful edge. See Figure 6: The three technical Workpackages of ACES and their relationships.

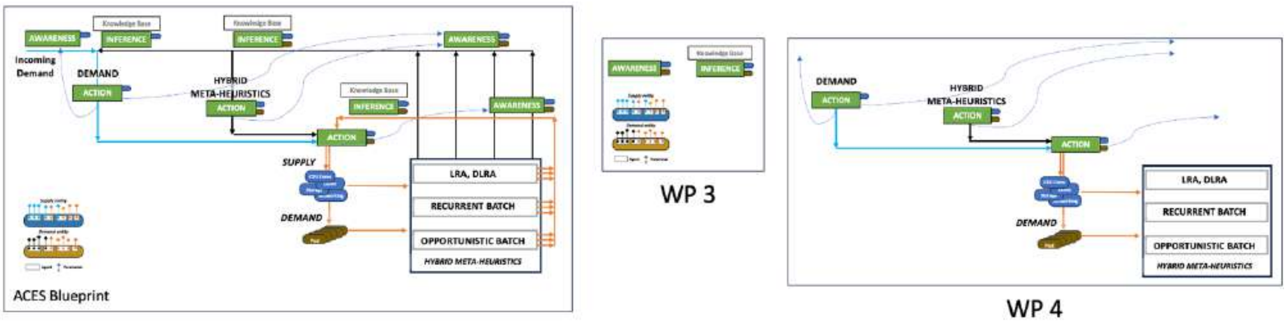


Figure 6: The three technical Workpackages of ACES and their relationships

Challenges in the optimization of demand and supply

Within these cycles, each step of from Awareness over Inference and Actions has its own challenges that will be researched in ACES *(to be completed with input from consortium and as guidance in the WP)*. See Figure 7: The logic components of ACES-EDGE.

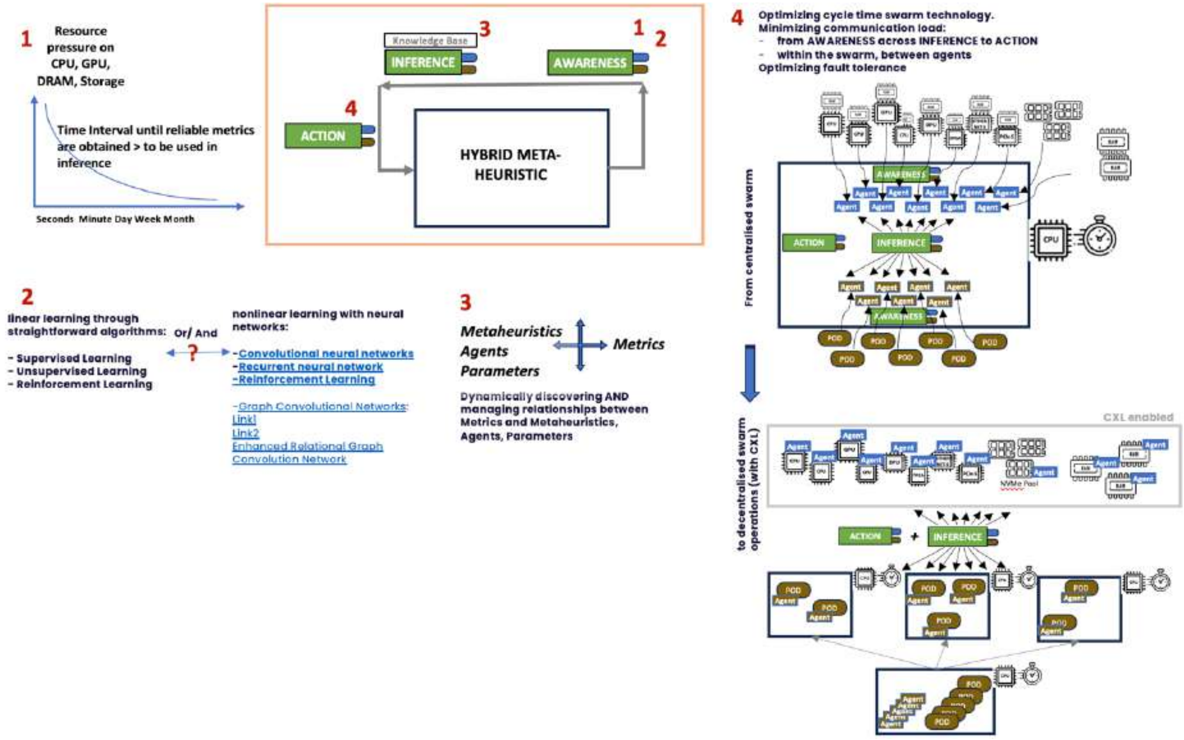


Figure 7: The logic components of ACES-EDGE

In *Awareness*, the challenges (see Figure 7: The logic components of ACES-EDGE, items 1 and 2) are:

- Time intervals: the sample rate used in training and used for monitoring. These two time intervals need to be considered, the time interval used in the training (and retraining) of the AI/ ML, and the time interval to monitor the metric (e.g., to detect an anomaly). Short measuring intervals and high sample rates will add additional resource pressure to the infrastructure.
- The appropriate AI/ML technology to recognize patterns in demand and supply. Research shows that linear learning is well suited to capture patterns of single entities, for example CPU utilization. Limiting conditions for using these technologies could be the absence of training data sets and the cycle time to train and retrain compared to the desired responsiveness in the platform. Non-linear learning technologies can help analyse more complex intra- and inter-relationships of demand and supply. Limiting conditions here can be the relatively large knowledge gap about the possible inter- and intra-relationships and the resources required to create complete and useful maps of the relationships.

In *Inference*, the challenges (item 3) are:

- Discovery tools: that can be used to discover correlations between various metrics from demand and supply collected in the awareness functions.
- Sharing insights with application users and edge devices, application developers, platform operators.
- Inferences based on these correlations and that dynamically control the:
    - Awareness functions: dynamically switch on-and-off the type and number of metrics being monitored.
    - Changes parameters of agents.
    - Changes agents within a hybrid meta-heuristic.
    - Changes a hybrid meta-heuristic.

In *Action*, the challenges (item 4) are:

- Managing the swarm communication load on the platform. Agents have a stop and listen cycle, a heartbeat, to align their communication. Also, as the number of agents grow in a demand-supply space, the communication grows exponentially. This can become a bottleneck in the responsiveness of the swarm behaviour and create a critical network communication overhead.
- Distributing the swarm intelligence, namely the individual agents. ACES ambition is the first step towards managing intelligently a fully disaggregated infrastructure. Fully disaggregated means that agents on the supply side represent individual CXU's, Memories, Storage, Networking devices and run on processing capabilities inside that individual piece of hardware. Currently this technology is being developed and rolled out into the market, called CXL. ACES intelligence precedes yet anticipates the widely availability of this technology and addresses the widely acknowledged challenge of orchestration complexity in such a disaggregated infrastructure.

Another important challenge is the Demand-Supply matching. The problem is that idealized experimental environments often deviate from the workloads of real production environments, especially under the online-offline co-located hybrid containerized deployment mode. Solving these issues could be done by a load interference estimation (highly complex) or a data-driven and machine learning based method for performance prediction. These learning technologies need to be able to handle the different type of workloads and their variations in contents and microservice topology:

- Publish/subscribe system (e.g., a sensors system, a messaging service, etc.) have workloads composed of publications, subscriptions, and unsubscriptions.
- Batch application workload is composed of processing orders.
- Complex event processing system has workloads that contain different events to be processed (e.g., text messages, weather reports, monetary transactions).

Finally, workload prediction might not produce direct information about the possibility of an SLA violation of the predicted workload, which is essential for the decision-making process of triggering a scaling action and the number of resources this action will manage. Also, this prediction might not reflect the number of resources used by the system, which does not provide a snapshot of its performance based on its SLA parameters (high-level metrics). An autoscaling technology will benefit from having additional information related to the SLA parameters (i.e., throughput and response time) to decide about said scaling action. Efficient resource usage and meeting performance goals for the microservices is determined by autoscaling, scheduling and placement of the microservice Pods on the nodes in the cluster. The scalability describes a system's elasticity, ability to adjust to change in demand in terms of time, volume, and complexity. The general aims of vertical and horizontal autoscaling are twofold: a) efficient resource usage and allocation while b) maintaining a good performance for the microservices, which is another hard challenge.

2.3 Mapping the Key exploitable outcomes on the partnership

All ACES beneficiaries shall claim the Key Exploitable Results consistent with their strategies and business plans. In this section, each partner elaborates on the linkages between their technological and industrial interests and the key exploitable results of ACES.

This mapping is necessary to understand the self-assigned roles of each consortium member in the exploitation of ACES in its components and as a whole service concept. Sources of data and information include:

- The ACES-EDGE blueprint
- The functional specifications of the service
- The ACES technical team
- Developer's interviews
- Partner data, if available

2.3.1.1 HIRO

HIRO's mission is to lead and enable the innovation and technological development of the powerful edge infrastructure, in the edge-to-edge-to cloud continuum. The goal is to exploit the potential coming from the growing amount of Big Data coming from Smart Cities, Factories, Autonomous driving and IOT by deploying and appropriate hardware and software infrastructure. The primary focus of Industry leaders is in building Cloud capacity at scale, IOT platforms, machine learning, AI and Cognitive Intelligence, while Edge computing gets relatively less attention.

HIRO has recognized this untapped potential and has created business partnerships with hardware manufacturers, software platform developers to be able to deliver scalable edge infrastructure (bare metal and appliance). These EdgeMicroDataCenters (EMDC's) can deliver at the local edge cloud-type capacity and services that is easy scalable, minimizes latency, safeguards privacy of data locally, guarantees very high in an energy efficient and compact formfactor, making available to customers. By Europe, 'owning' first-hop from edge device to the www, their powerful edge infrastructure is capable of safeguarding the European values (GDPR, Digital Sovereignty) through hardware and software innovation.

The engagement of HIRO in the ACES-EDGE RIA is particularly important, as it researches and develops an innovative orchestration platform that uses distributed emergent intelligence across the infrastructure while using a combination of open common AI and ML and swarm based intelligence. This approach breaks away from the current common hub and spoke, north-south, cloud-edge orchestration that will not be able to handle the dynamics of demand from edge devices at on an ad hoc basis. Aces aims to provide end-to-end transaction resilience, reliability, and stability of automation in edge-cloud management and secure flows of sensitive data and applications.

HIRO is interested in the entire system, as an enabler of value-added Edge-as-a-Service technologies, which can help solving the needs of different users of diverse industries.

However, HIRO targets specific technological elements of ACES, which once developed and deployed will be determinant for the efficient and effective operations of the ACES EaaS, namely all components of the autopoietic configuration and management functions. These will be embedded in the HIRO technological systems for Edge services.

2.3.1.2 INESC-ID

INESC-ID is a non-profit research institution, privately owned by IST and INESC. INESC-ID is also an Associated Laboratory since 2004. It is also viewed and stands as one of the most dynamic research institutes in Portugal in the broad areas of computer science and electrical engineering. Its research and development activities focus on artificial intelligence, information and decision support systems, graphics and interaction, distributed systems, communication networks, high performance computing, nano-electronics and energy systems. INESC-ID promotes cooperation between academia and industry by addressing research on daily life issues, such as healthcare, space, mobility, agri-food, industry 4.0, and smart grids.

The technologies developed by INESC-ID will target non-functional aspects of ACES to ensure its proper behaviour: performance, efficiency, security, and robustness.

For performance, we will leverage several forms of network function (NF) acceleration to achieve low latencies and high throughputs at the edge, namely by running NFs in programmable networking hardware (such as Intel Tofino). As it is a challenge to build accelerated code, given both the variety of accelerator targets and the intricacies of acceleration techniques (e.g., how to parallelize code without breaking its semantics?), we will explore program synthesis to meet this challenge.

For efficiency, we will investigate self-adaptation ML mechanism for ML retraining considering the computational cost of retraining on the edge nodes, particularly.

For robustness we will develop techniques to ensure uninterrupted availability of workload state and data, even in the face of node failures. We will also develop robust and secure ML-based mechanisms for network control, and primitives for AI-based self-monitoring to detect malicious attacks and self-driving network control on various cognitive layers.

Also paramount in ACES security is maintaining the privacy and integrity of sensitive data and workloads. We will explore trust execution environments to support the execution of privacy-preserving workloads and enforce anonymous authentication.

2.3.1.3 IPTO

The mission of IPTO is the operation, control, maintenance and development of the Hellenic Electricity Transmission System, to ensure the country's supply with electricity in an adequate, safe, efficient and reliable manner, as well as the operation of the electricity market for transactions outside the Day Ahead Scheduling, pursuant to the principles of transparency, equality and free competition.

The Autopoiesis cognitive edge-services cloud (ACES) could support the reliability of the Electricity system through distributed market algorithms that take into account the local characteristics of the various regions. Additionally cognitive frameworks with predictive capabilities (such as the prediction of the energy demand) could improve the reliability and efficiency of the power grid by adjusting accordingly the energy supply side. Autopoietic behaviour of the system could improve the safety of the electric equipment through cognitive machine learning algorithms such as anomaly detection. Best practices and insights can be shared also with other Transmission Operators across the globe.

2.3.1.4 SIXSQ

SixSq provides two solutions for devices and applications management to the market, Nuvla.io and NuvlaEdge.

- NuvlaEdge is a plug and play agent software, which enables customers to turn any Linux hardware into a smart edge device.
- The Nuvla.io platform provides customers with a simple dashboard from where they can monitor and manage the edge devices mentioned above. They can also deploy and update their own- and third-party applications, not only at the edge but also in the cloud.

In a nutshell, Nuvla allows users to simply manage cloud and edge infrastructures.

The technologies developed within the ACES project will benefit both the NuvlaEdge software, and the Nuvla.io platform.

First, the cognitive and autopoietic frameworks and features envisioned for ACES will be implemented in the NuvlaEdge software, to bring a new level of decision-making and fluidity at the edge. Adding this intelligence layer at the edge will also enable workload placement optimization, which will result in enhanced performance and capabilities of the edge devices managed through Nuvla.

This update will also improve the Nuvla.io platform, with the ability to offer new functionalities to customers, and improve the performance and stability of the devices.

Customers will benefit from the implementation of this cognitive layer, as it will facilitate and optimize the management of the infrastructure, for both the devices and applications.

2.3.1.5 LAKESIDE LABS

Lakeside Labs is a non-university research centre dedicated to self-organizing networked systems. The mission is to apply related methodologies like swarm intelligence to the coordination of complex technical systems like swarm robotics, smart production, mobility- and transportation systems, IoT and Cloud-Edge computing.

The technologies developed by Lakeside Labs will allow for autopoietic behaviour of EMDCs, groups of EMDCs or the cloud-edge continuum. This will be achieved by considering requests (e.g., workloads, pods) and resources (computation nodes, network nodes, cache, storage) as swarm agents that arrange for distribution of workloads, data and applications and orchestration of resources in a decentralized, self-organizing manner. The desired behaviour of the system is considered as emergent properties of the subtle interplay between agents without requiring a central scheduling entity.

The coordination and orchestration mechanisms based on swarm algorithms are the exploitable outcome for Lakeside Labs in this project. They can be used to even spread swarm technologies throughout other layers of the edge-cloud continuum (e.g., coordination of IoT devices at the edge) or be applied to other domains such as cyber physical systems of systems (CPSoS), optimization of production, logistics, transportation and smart energy systems.

2.3.1.6 Universidad Politécnica de Madrid

UPM is a public university that offers high-quality education and research in engineering, architecture, and science. UPM participation in ACES project takes part through IPTC (Information Processing and Telecommunications Center), a research center that focuses on information processing and telecommunications. IPTC collaborates with academia and industry to develop innovative solutions for various domains and challenges.

In the case of the Key Exploitable Results (KERs) outlined in the ACES proposal, UPM will further develop novel solutions for decentralized, edge distributed systems optimization. The new algorithms and models created during the project will be published in internationally leading journals and conferences, further strengthening the research output of the university. Moreover, university courses and final thesis will be infused with the ACES research and industrial context to further improve its relevance, novelty, and employability for the students that take part in our taught programmes.

2.3.1.7 MARTEL INNOVATE

The set of Key Exploitable Results (KERs) has been already outlined in the ACES project proposal, but it needs to be specialised for Martel's context. Moreover, the project results themselves undergo a discovery process while the engaged partners understand them with increasing depth, adapt their plans, and gain insight on the technological and business consequences. Therefore, this mapping to Martel's positioning and interest is only a first approximation and will be refined in the later iteration of the Exploitation Plan document throughout the project life cycle. Considering first Martel as a software vendor, with an open-source offering in the areas of IoT, Smart Cities, cloud-native software infrastructure and telecommunication services, as well as engaged in several other research and innovation initiatives (other Horizon Europe projects, but also memberships in FIWARE, BDVA, and other associations), the immediately interesting KERs for Martel (as listed in the original ACES proposal, and therefore still undergoing an evolutionary adaptation process) are:

- *ACES Cloud Stack*. The horizontal platform as a whole, being predominantly based on open-source and standards, will blend particularly well with Martel's existing offering for IoT and

Smart Cities, providing a convenient approach to deploy the Orchestra Cities platform[8] and similar Martel cloud-native offerings on a cloud-edge-IoT computing continuum environment. While not yet completely defined, the ACES technology stack already shows some points of contact with what Martel routinely uses (e.g., the Kubernetes orchestrator and associated software ecosystem). Moreover, the coupling with the EMDC specialised and configurable hardware node opens the door to business partnerships opportunities, where turnkey HW/SW solutions could be offered to the market by, e.g., Martel and HIRO.

- *ACES Cognitive Framework and Autopoiesis Framework*. The ACES project seeks to obtain edge intelligence and emergent *self-** properties by adopting multiple AI approaches. In particular, combining coordination-centric (through Swarm Computing) and data-centric AI (through techniques such as distributed deep learning) offers a promising spectrum of possible specialised software components and solutions. Martel's plan, while contributing mostly to the data acquisition and processing aspects, and less to the coordination and actuation parts, is to evaluate the benefits and applicability of these ACES software frameworks as value-added additions to Martel's products.
- *Workload placement and optimisation*. Given the very compatible choices made by the ACES consortium (Kubernetes ecosystem, open-source infrastructure, etc.), Martel expects that results in this area (pending suitable IPR agreements where necessary) will offer a rather direct path to exploitation, by being provided as a combination of scientific publications and (most likely open-source) reference implementations, which Martel and other interested partners could directly take to production-level quality within their own offering or, alternatively, propose to suitable technical organisations for wider adoption and possibly standardisation.

In addition to that, Martel has outlined a further mapping that is suitable to its other core activities of project consulting and communication and dissemination services. This other, non-technical facet of Martel corporate identity addressed the same KERs, but yields indirect exploitation strategies such as finding and fostering synergies and cooperation across multiple projects where Martel has stakes or influence (either as a non-technical partner or as engaged in one of the relevant coordination action projects), creating engagement around ACES approaches and results through discussion panels and speeches when organising events on behalf of the European Commission or other customers, partners, and stakeholders.

2.3.1.8 TU DARMSTADT

TUD is a leading European university of technology performing world-class research in cybersecurity. The results developed within ACES will support TUD in positioning itself as a major research institute in system security in Europe, thus helping to attract world-class new talent and broaden its research collaborations with other renowned research institutions as well as industrial partners. TUD aims to publish project research results in first-tier scientific conferences and high-impact journals in security and AI research. Further, pertinent aspects of project outcomes will also be incorporated into TUD's teaching curriculum, seminars, summer schools, and thesis topics.

2.3.1.9 UNIVERSITY OF LJUBLIANA

The University of Ljubljana's Faculty of Computer and Information Science is a leading institution for computer and information science education and research in Slovenia. The faculty's primary function is educating undergraduate and postgraduate students to develop skills in different areas of expertise. Another essential purpose of the faculty is generating new knowledge through research in

[8] Orchestra Cities is designed as a multi-function platform allowing a broad range of IoT-enabled applications and services, including waste and water management, transportation, air quality monitoring, traffic and parking, and more, to be effectively defined, offered and monitored. https://www.orchestracities.com

order to solve contemporary problems. As such, UL will advance the state-of-the-art by developing novel techniques for edge-wide workload placement and edge-wide workload optimization in the scope of ACES. In particular, it will focus on pushing the boundaries of placing and optimizing the workload across a large-scale infrastructure, considering both edge-wide as well as across cloud and edge cases.

2.3.1.10 IDSIA-SUPSI

IDSIA is an autonomous research institute that maintains a close collaboration with SUPSI University, dedicated to exploring diverse facets of artificial intelligence and machine learning. The institute undertakes both theoretical and applied research endeavours, with the goal of furthering the comprehension of AI algorithms and their practical implementations. Through fruitful partnerships with esteemed universities, research institutions, and industry leaders across the globe, IDSIA fosters a dynamic and cooperative research environment.

One of the research sectors at IDSIA focuses on intelligent multi-agent systems. The IDSIA research centre has embarked on the ACES project, aiming to participate in the developing of an autopoietic cloud framework. This innovative framework opens up as exploitable outcomes new avenues for research and experimentation, particularly in the integration of decentralized optimization techniques with machine learning and artificial intelligence tools. Consequently, IDSIA can expand its knowledge and expertise in this domain, further enriching its research endeavours.

More specifically, IDSIA aims to address distributed workload allocation problems by formulating them in a manner that can be efficiently solved through fully decentralized gradient tracking schemes. The agents within the framework will possess intelligence, enabling them to dynamically update their requirements by leveraging forecasting techniques facilitated by Recurrent Neural Networks and/or Q-learning algorithms. Furthermore, IDSIA seeks to explore the potential of federated learning, which involves sharing knowledge across various layers of the cloud architecture. These research themes hold significant importance for IDSIA, greatly enhancing the institute's expertise and knowledge base.

2.3.1.11 DATAPOWER

DataPower uses data modelling and enrichment and integration to support business processes and evidence-based decision-making processes and risk-balanced assessments, helping the definition of strategies, competitive positioning and the optimisation of processes and value chains.

DataPower’s expertise concerns process design, process analysis, process performance optimisation, Digital Transformation, Supply chains and value chains, Dynamic simulation methods. We integrate digital technology into all areas of a business, changing the way organisations, institutions, administrations, fundamentally changing how they operate and deliver value to customers and stakeholders. DataPower is also interested on the cultural change that requires organizations to continually challenge the status quo, experiment, and get comfortable with failure.

Per se, DataPower has no direct interest in any of the specific building blocks of ACES-EDGE, except for the entire edge service as such and specifically the AI/ML enabling the autopoietic functions. In this sense DataPower will take up the whole ACES-EDGE concept and use it within the ongoing and forthcoming digital transformation initiatives.

2.4 Market trends and gaps mapping

2.4.1 The maturity of the edge-cloud concept

Several analyses[9] discuss the technology and market life cycle of edge computing. The technological developments drive the need for computing and storage power towards the edge, that is closer to where the data is recorded, created and/or needed. There are estimates of the demand that expect a ten-fold increase in the next decade.

The computing ecosystem is evolving, as is the servitisation progresses that involve Software as a Service, Application as a Service, Platform as a Service, Infrastructure as a Service, with the cloud being a core element.

Major ICT providers, i.e., technology developers; system integrators; solution providers; IT providers; hyperscalers; telecom operators, are moving towards edge computing, at least in their communication. All these players are currently assessing the value chain of the edge to determine where to place their focus and offering on the ecosystem; on hardware; on orchestration solutions; on system integration; or system operation. It is currently not clear where the value is placed and how to set up the whole service sale and resale.

One of the key questions concerns the timing of the action within the edge-cloud life cycle. In this context, investment risk assessment is a particularly sensitive aspect. The current state of the life cycle confirms[2] that a) there is the need for compute and storage infrastructure outside datacentres and close to where data is captured and processing required; b) there is potential market value to edge computing; c) that the current challenge is about positioning services and offerings, rather than directly deciding about investments.

At the present state of development of ACES solution development and of the definition of the market strategy, even though most of the players mentioned above are presenting their edge offering in one way or the other, the real challenge is constructing demand scenarios to calibrate the investment to set up the offering. At the same time the question to answer will be how to balance the on-premises compute and storage capacity development in respect to moving the compute capacity and storage power towards the edge.[10]

That depends significantly on demand, on bandwidth, on the characteristics of applications, on the required and delivered SLAs, the necessary latency, and the mandatory data sovereignty. Very much depends on the ability of edge computing providers to motivate their clients to adopt processes and applications that process data at the edge, constructing fitting cases with attractive cost-benefit ratios for potential adopters. The latter not only concerns hardware and software (application) CAPEX and OPEX, but also the distinctive characteristics of the edge in terms of latency, resilience, cybersecurity, Quality of Service, system stability, bandwidth requirements, performance documentation, and compared economic and functional benefits in respect to currently operating solutions.

In other words, the challenge lies in optimising technology-pushed innovation towards the edge and at the same time creating the relevant demand, knowing there some relevant changes on the part of edge users may be necessary, with the relevant costs.

We summarise this initial discussion with the following three points:

[9] ArthurDLittle: "Edge Computing: Hype or Ripe" (https://www.adlittle.com/en/insights/report/edge-computing-hype-or-ripe)
[10] Gartner Group "Notes From the Edge – Is Edge Computing All Hype?" https://blogs.gartner.com/bob-gill/2021/10/05/notes-from-the-edge-is-edge-computing-all-hype/

1. The edge market value
   - Creating demand for workloads shifted towards the edge, leveraging security, data protection and latency concerns.
   - Relocating datacentres on the edge, optimising CAPEX and OPEX.
2. Edge as an enabler
   - AI and ML applications, Augmented Reality, Robotics and Cyberphysical systems
   - Computer vision, video analytics
   - Data aggregation at the source of creation
   - Device offloading
   - Leveraging 5G apps.
3. The deployment of Edge computing is based on the verification of
   - Size of the market and requirements of the market (functional specs; service levels; latency; processing power)
   - Market growth: do current remote and geographical processing by organisations actually require fast migration towards the edge?
   - The actual reduction of latency, in the cases where a "call home" for diverse reasons (such as authentication, encryption keys, or even some logic, etc) may reduce the latency advantage;
   - Costs of edge computing: how does delocalisation and move towards the edge compare to the economies of scale of centralised data centres?

At present times and considering current developments, edge computing and edge storage appear appealing and promising. The challenge is to verify how the supply will shape its offering and how demand is ready and willing to adopt different system models, with their relevant costs.

2.4.2 Market trends

This section will highlight the main market trends and segments relevant to EDGE-Cloud services. It will explore the main developments in demand as well as the specific characteristics required by users.

Edge computing is an innovative, more performing, and more flexible way to configure the cloud and to comply with levels of service, security and data protection requirements. The transition from cloud to edge is still in its inception and both concepts "are not mutually exclusive paradigms. Instead, they interact and complement each other, providing a more comprehensive and efficient computing ecosystem."[11]

[11] Edge Computing and the Future of Cloud Computing: A Survey of Industry Perspectives and Predictions. A.Shaji George, A.S.Hovan George,T.Baskar. 2023, PUIRJ | DOI:10.5281/zenodo.8020101

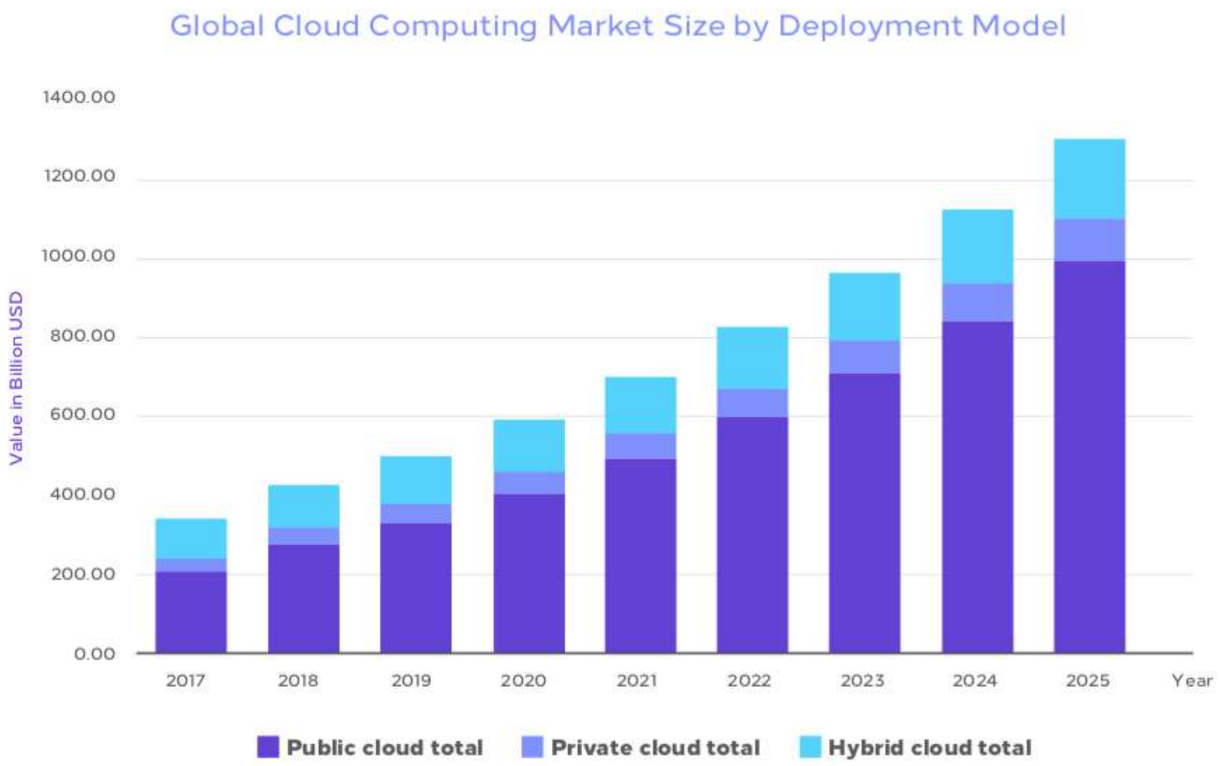


Figure 8: Global Cloud Computing Market Size by Deployment Model

According to the IDC analysis, presented in the chart above Figure 8: Global Cloud Computing Market Size by Deployment Model, the edge and cloud computing markets are expected to grow fast. In 2017, the global market for cloud computing was valued at less than USD 400 billion, and it is estimated to expand to USD 1,3 trillion in 2025. Obviously, the factors which affect these developments are many and these estimates may need to be adjusted, yet the outlook is very promising. The market scenarios for Edge computing are still uncertain, mostly due to the still-developing technological solutions and the need for corresponding user investment in new implementations and investment in migration from present centralised cloud services.

The edge computing market already includes a vast number of players, some (AWS, Microsoft, Google, Alibaba, IBM, Oracle) with unique and distinguishing features, and are expected to dominate the developing global market in the future. To understand the edge market, it is necessary to interpret the series of underlying causes and the accompanying patterns and trends: for example, the supply of edge services is not the only driver, but also the propensity of users to migrate their applications, computing, and storage to the Edge. Performance indicators are also relevant to lay out a proper cost-benefit analysis to underpin cloud-based service users to take advantage of concrete benefits.

Considering consulting industry reports, market research, and news sources that cover the edge cloud market, we can see that the trend is dynamic and in the next few years we should expect a continuous growth.

2.4.3 Recent market research

Edge computing enables rapid and real-time data analysis by processing and storing data closer to the source. It consists of a mesh network of micro data centres that process and store data locally, while also sending it to cloud storage. The Edge approach leads to significant advantages in terms of performance, resource allocation, optimisation, computing speed, data control, and protection compliance.

Figure 9: Edge Computing fundamentals. Source: McKinsey & Company[12]

The global market for edge computing is expected to grow exponentially following an increased adoption of edge devices. Edge devices include a range of technologies such as IoT devices (e.g., mobile point-of-sale kiosks, smart cameras), industrial PCs, medical sensors, gateways, and computing infrastructure. In 2022, the global market for edge computing was valued at $11.99 billion. It is estimated to expand from $15.96 billion in 2023 to reach $139.58 billion by the year 2030, with a CAGR of 36.3%.[13]

2.4.3.1 Specificities of Cloud Computing

Cloud computing is a method of processing, storing, and managing data on a network of remote computers hosted on the internet, rather than on local devices or servers. It involves a shared storage area that allows all network devices to access data simultaneously.

Cloud computing is an IT service delivery approach where third-party service providers use the internet to provide computing resources and software tools. In this model, customers only pay for the time they spend using the computer, as well as the storage and bandwidth they utilize.

The global cloud computing market is driven by factors such as the increasing digital transformation in businesses, widespread adoption of internet and mobile devices, and the growing consumption of

[12] McKinsey Technology Trends Outlook 2022. Cloud and edge computing. August 2022
[13] Fortune Business Insights Market Research Report. March 2023.

Big Data. The adoption of technologies like IoT, 5G, edge computing, real-time analytics enabled by Machine Learning (ML) and Artificial Intelligence (AI) is expected to further increase the value of cloud computing in businesses.

The global cloud computing market was valued at USD 569.31 billion in 2022. It is projected to reach USD 677.95 billion in 2023 and by 2030, the market is expected to reach USD 2,432.87 billion, growing at a CAGR of 20.0%

The chart below by McKinsey and company (Figure 10: Cloud computing key facts. Source: McKinsey & Company), provides a snapshot of the current scenario of cloud computing, mostly centralised, and it depicts the weight of cloud computing across industries. The challenge for Edge computing will be to drive the migration of those application areas, which will benefit from a decentralised operation at the edge according to sound use case scenarios.

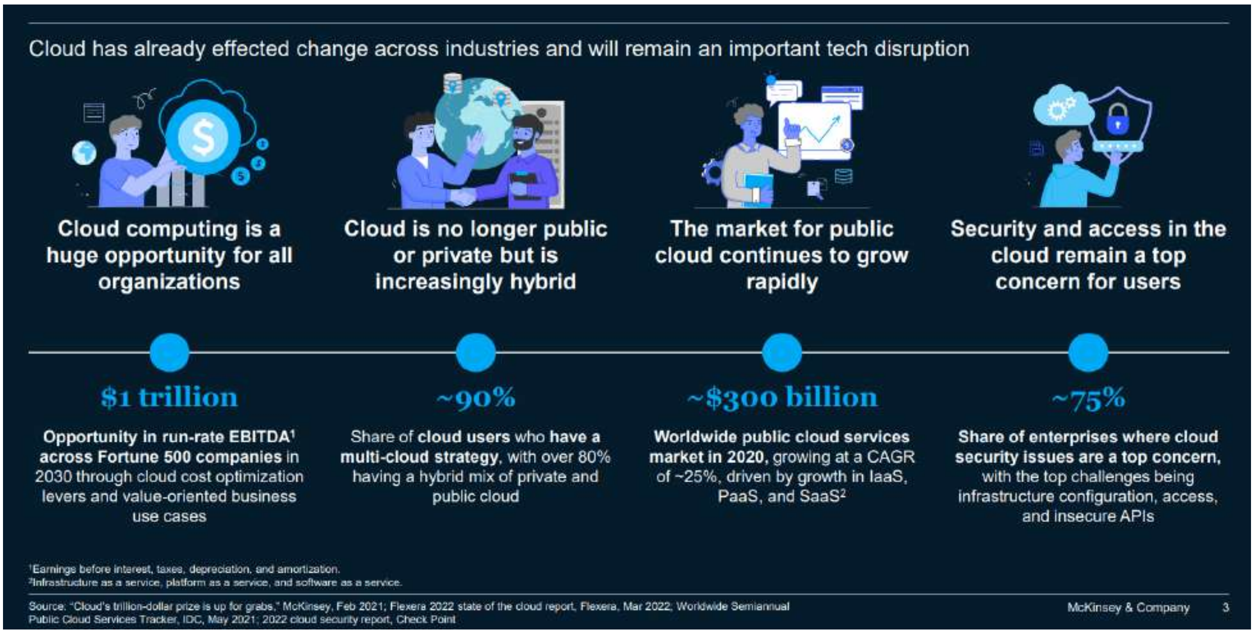


Figure 10: Cloud computing key facts. Source: McKinsey & Company[14]

2.4.3.2 Specificities of Edge Computing in Europe

The European edge computing market was valued at USD 815.1 million in 2020 and is projected to grow at CAGR of 26.5% till 2030 to reach a market size of USD 8.554 billion by 2030. The European cloud computing market was valued at USD 60 billion in 2022 and is projected to grow at a CAGR of 25% to reach USD 650 in market size by 2032. In 2021, 41% of EU enterprises utilized cloud computing, primarily for hosting e-mail systems and storing electronic files. Compared to 2020, the adoption of cloud computing increased by 5 percentage points. Among those enterprises, 73% used advanced cloud services such as security software applications, hosting databases, and computing platforms for application development, testing, or deployment.[15]

---

[14] *Ibidem*.
[15] GMD Research. Europe Edge Computing Market 2020-2030 by Component, Technology, Deployment, Application, Industry Vertical, Organization Size, and Country: Trend Forecast and Growth Opportunity. 2021

Cloud computing services cater to various business ICT needs. Among enterprises using cloud computing, 79% rely on cloud solutions for their email systems. Approximately 66% of enterprises use the cloud for file storage. Office software, such as word processors and spreadsheets, is utilized by 61% of enterprises, while 58% employ security software applications. Additionally, 46% use the cloud to host their databases.

The information and communication sector has the highest proportion of cloud computing users (76%), followed by other sectors ranging from 32% to 48%. Additionally, the retail trade sector experienced the highest increase in cloud computing usage compared to 2020 (+7 percentage points).

Cloud computing adoption varies significantly across countries. Sweden (75%), Finland (75%), the Netherlands (65%), and Denmark (65%) have a high percentage of cloud users, while Greece (22%), Romania (14%), and Bulgaria (13%) have lower adoption rates. Large enterprises show the highest adoption rate at 72%, followed by medium-sized enterprises at 53% (compared to 46% in 2020), and small enterprises at 38%.

2.4.3.3 Some estimates on Edge computing market trends

The global edge computing market size was valued at $11.99 billion in 2022 & is projected to grow from $15.96 billion in 2023 to $139.58 billion by 2030. The U.S. dominated the market in terms of the presence of key players. In Europe edge computing market was valued at $815.1 million in 2020 and is expected to grow by 26.5% annually over 2020-2030.

According to International Data Corporation's (IDC) Worldwide Edge Spending Guide "enterprise and service provider spending on edge computing will reach $40 billion in 2022 in Europe and significantly increase over the forecast period, reaching nearly $64 billion through 2025, with a five-year compound annual growth rate (CAGR) of 16.4%."[16]

Available studies indicate that Edge-powered SaaS (Software as a Service) on 5G networks are becoming increasingly popular. There is a trend of Companies worldwide who are integrating their services with the cloud, leading to a surge in data and processing traffic due to the reliance on IoT and AI devices. Global companies are increasing their investments in 5G as more than 380 telecom operators seek to enhance speed and agility.

The combination of edge computing and 5G infrastructure offers benefits such as reduced complexity, cost savings, and strengthened cybersecurity defences. In June 2021, IBM Corporation collaborated with various edge partners to drive digital transformation for enterprise clients. The partnership involved more than 30 partners, including Capgemini, Atos, Foghorn, HP Inc., and Lenovo. This collaboration aims to enable businesses to leverage artificial intelligence, hybrid cloud, and 5G technologies to incorporate edge solutions into client operations.

It is evident that Edge computing is a promising technological system development but faces a number of challenges related not only to the maturity of technological solutions, related to setup, orchestration and AI/ML-rooted management and optimisation, which are on the supply-side. Similar issues arise on the demand side, which are related to user investment and digital transformation decisions.

We want to consider that:

- Building a comprehensive and resilient edge computing solution can be demanding and resource intensive due to need to invest in edge nodes, devices, and data centres.

[16] 2022, the International Data Corporation (IDC)

- Ensuring the security of the entire computing network also entails significant costs for providers, which may limit market expansion.
- users may need to invest in process and system innovation.

No final word can be said on these aspects at this point in time. A more detailed analysis will be made when the ACES project will undertake a cost-benefit analysis based on a number of system-user scenarios, which will demonstrate the advantage of an Edge solution in respect to the centralised cloud solution, also considering data protection and security, as well as performance issues.

2.4.4 Factors affecting the development of Edge computing

There are several factors driving the uptake of Edge computing and storage systems. These factors need to be assessed against the functional needs, the level of service required as well as the need to invest in the digital transformation of systems currently in use, considering the advantage of the Edge system in respect to a centralised cloud system.

2.4.4.1 Driving factors

Recent studies have identified several driving factors of Edge computing, which are synthesised hereafter:

- Market growth is driven by the increasing exploitation of Big Data, Artificial Intelligence (AI), Machine Learning (ML), and other technologies. These technologies enable users to monitor, analyse, and visualize unprocessed data, transforming the market landscape.
- The widespread adoption of AI and ML in businesses has increased data usage and mitigated data storage issues. AI and ML technologies contribute to increased business productivity, improved decision-making processes, and reduced operating expenses.
- Cloud solution providers focus on research and development, product enhancements, and new product launches that integrate Big Data, AI, and ML. Therefore, the increasing adoption of Big Data, AI, and ML is expected to drive further market growth.
- For example, Microsoft Corporation launched Azure Health Data Service, a Platform as a Service (PaaS) platform designed to support transactional and analytical workloads.
- Cloud providers are set to increase investments in edge computing for faster AI deployment. While AI models are typically trained in centralized data centres, utilizing edge computing allows cloud providers to deploy models closer to end users, enabling rapid queries and responses.
- Businesses with distributed user bases will utilize edge networking for faster access, while the post-pandemic rebuilding process may drive increased cross-border connectivity needs. Edge computing providers enable rapid expansion with reduced risk and simplified management.
- Edge computing presents significant value and cost-saving potential, serving as a valuable complement to cloud and CDN services for optimizing operational efficiencies. As organizations increasingly adopt edge computing, they gain the ability to evaluate its overall worth and strategically leverage its capabilities to reduce costs and improve performance, ensuring a well-rounded approach to meeting their computing needs.
- Cloud gaming is revolutionizing the gaming industry by shifting game processing to the edge, freeing users from device constraints. Game content is streamed from edge servers, enabling power savings and access to a wider range of content. However, cost challenges exist in delivering edge-based game content compared to traditional methods. Despite this, investing in edge computing is crucial to meet customer demands for simplicity and efficiency.
- Cross-border businesses with distributed user bases will increasingly adopt edge computing to host latency-intensive applications and services at the network edge. Edge compute servers come in two options: bare metal servers and virtual machines (VMs). Bare metal

servers offer dedicated resources and enhanced control, while VMs provide virtualized environments and resource sharing.

2.4.4.2 Hampering Factors

Concerns surrounding data security and privacy, data loss, data breaches, unforeseen emergencies, application vulnerabilities, and cyber-attacks pose challenges to the advancement of cloud computing.

While cloud services offer efficiency and cost-cutting benefits to businesses, storing data in the cloud exposes it to the risk of cyber-attacks. Cyber-attacks such as cloud malware injection, meltdown, account or service hijacking, and man-in-the-cloud assaults can exploit vulnerabilities in company data, leading to business closures, financial losses, and impeding market expansion.

Also, the need to meet regulatory and compliance requirements is continuously increasing as companies adapt to changing business needs.

Finally, edge computing increases the attack surface and potential data vulnerability, with 66% of respondents concerned about ransomware attacks and 74% emphasizing the impact of compromises at the edge, according to an AT&T study[17]. Furthermore, 64% express concerns about attacks on cloud workloads. As companies shift to distributed edge frameworks, security plans must be re-evaluated and enhanced. Innovative approaches are required to strengthen endpoint security and mitigate risks associated with edge computing.

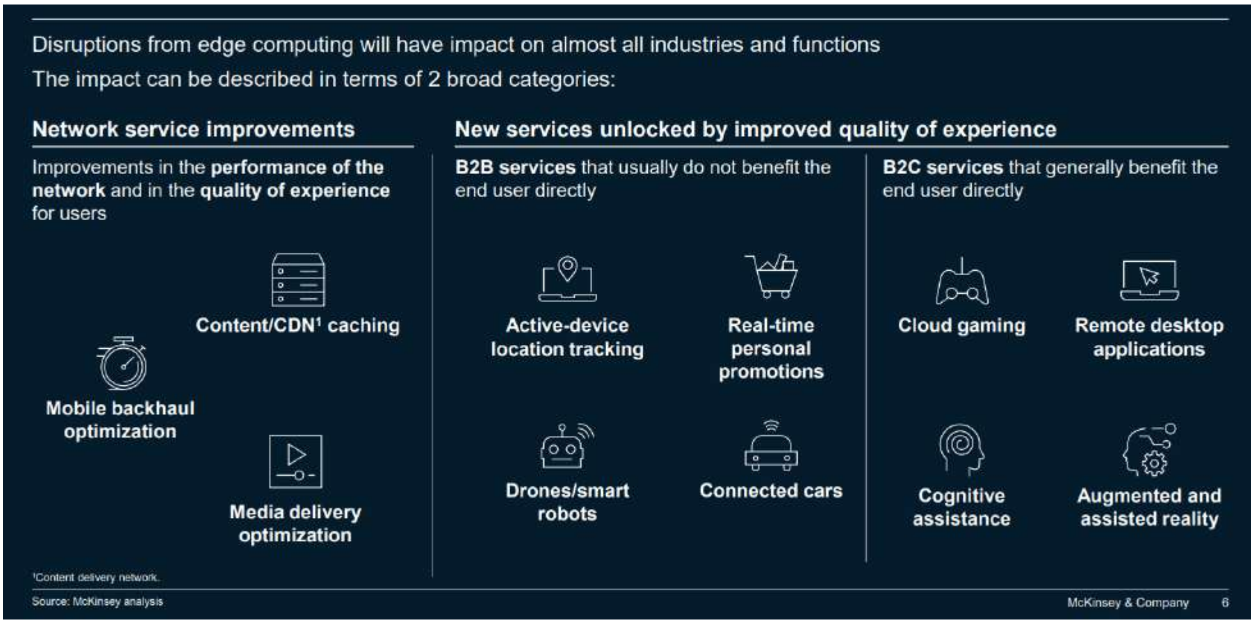


Figure 11: Impacts of disruptions in Edge Computing on industries. McKinsey and Company.[18]

[17] 2023 AT&T Cybersecurity Insights Report: Edge Ecosystem
[18] McKinsey Technology Trends Outlook 2022. Cloud and edge computing. August 2022

2.4.5 KPIs for Edge computing

The specific characteristics of Edge computing can be described and quantified by a number of Key Performance Indicators reflecting the core elements. These KPIs can be used to construct the use cases and their cost-benefit scenarios. The performance of edge computing is fundamentally related to four types of KPIs:

- Site availability,
- Network performance,
- Storage capacity
- Compute resources availability and performance.

According to research[19], there are several indicators that can help us to define the cost-benefit scenarios and thus understand the market:

- Latency: Measures the time it takes for data to travel from the edge devices to the cloud and back. Lower latency indicates faster processing and better performance.
- Bandwidth Utilization: Evaluates the efficiency of network bandwidth usage between edge devices and the cloud. Higher utilization suggests optimized network resources.
- Service Availability: Measures the percentage of time that edge computing services are accessible and operational. Higher availability indicates reliable and uninterrupted service.
- Edge Device Utilization: Assesses the usage and efficiency of edge devices deployed in the network. Higher utilization implies optimal utilization of resources.
- Data Transfer Rate: Measures the speed at which data is transferred between edge devices and the cloud. Higher transfer rates indicate faster data processing.
- Cost Optimization: Evaluates the effectiveness of cost management strategies in edge computing deployments. This KPI assesses if the costs associated with infrastructure, maintenance, and operations are optimized.
- Scalability: Measures the ability of edge computing systems to handle increasing workloads and accommodate growth without compromising performance. Higher scalability ensures flexibility and adaptability.
- Energy Efficiency: Evaluates the energy consumption and efficiency of edge computing infrastructure. Higher energy efficiency indicates reduced environmental impact and cost savings.
- Quality of Service (QoS): Assesses the overall performance and user satisfaction with edge computing services. QoS metrics may include response time, error rate, throughput, and other user experience indicators.
- Edge Device Reliability: Measures the stability and uptime of edge devices in the network. This KPI assesses the frequency and duration of device failures or disruptions.
- Data Security: Evaluates the effectiveness of security measures implemented to protect data at the edge. KPIs can include the number of security incidents, data breaches, or unauthorized access attempts.
- Edge-to-Cloud Data Synchronization: Measures the accuracy and timeliness of data synchronization between edge devices and the cloud. This KPI ensures that the data remains consistent across the entire system.
- Edge-to-Cloud Data Transfer Efficiency: Evaluates the efficiency and optimization of data transfer between edge devices and the cloud. Higher efficiency indicates reduced latency and improved overall performance.

[19] https://www.techtarget.com/searchdatacenter/tip/4-essential-KPIs-for-edge-computing-management

- Workload Distribution: Measures the balance and distribution of computing workloads between the edge and the cloud. This KPI ensures optimal resource allocation and prevents overloading of either edge devices or cloud servers.
- Edge Device Management: Evaluates the effectiveness of device management processes, including provisioning, configuration, monitoring, and maintenance of edge devices. This KPI assesses the efficiency of managing a large number of distributed devices.
- Time to Market: Measures the time it takes to deploy new edge computing services or applications. This KPI indicates the efficiency of development, testing, and deployment processes.
- Edge Resource Utilization: Evaluates the efficient utilization of computing resources available at the edge, such as CPU, memory, and storage. This KPI ensures that resources are optimally allocated and not underutilized.
- Service Level Agreements (SLAs): Measures the adherence to predefined SLAs that define the performance, availability, and reliability of edge computing services. This KPI assesses whether the agreed-upon service levels are being met.
- Customer Satisfaction: Evaluates the overall satisfaction of customers or end-users with the edge computing services. This can be measured through surveys, feedback, or user experience metrics.
- Innovation and Agility: Measures the ability of the cloud edge computing infrastructure to support rapid innovation, development, and deployment of new services or applications. This KPI reflects the agility and adaptability of the system.
- Storage capacity: Edge computing relies on local storage resources for storing and processing large volumes of data. Storage capacity is measured in gigabytes or terabytes, depending on workload and data expectations. Storage utilization metrics help determine when additional storage capacity is needed and inform data retention and backup decisions.

Selecting the appropriate KPIs depends on the specific use case, goals, and priorities of the edge computing deployment. It's essential to align the chosen KPIs with the desired outcomes and continuously monitor and adjust them as needed to achieve optimal performance. The KPIs will be tested in different scenarios, use cases and settings.

2.4.6 Service Advantages of a subset of providers

The most important providers have specific added values to offer.

Provider	Added Values
AWS	Leader in breadth of services: AWS is the CSP with the greatest number of services worldwide, offering more than 300 services. Prolific ecosystem: AWS offers a very strong network of partners an ISVs that prioritize developing on the AWS platform first before other CSPs
Microsoft	Solutions-oriented: The Microsoft Azure Cloud environment offers numerous differentiated solutions in different industries such as telecommunications, healthcare, manufacturing, retail & financial services Hybrid Cloud & Multi-cloud: Microsoft places significant focus on offering hybrid & multi-cloud capabilities to continue to support enterprises who want to keep hybrid environments

Google	Sales execution: Google Cloud is consistently focused on perfecting their sales strategies with the recent increase in market share specifically from enterprises
Alibaba	Engineering leadership in Asia: Alibaba cloud offers leading engineering skills in China and the overall Asia pacific region, allowing them to create differentiated products Data & Analytics capabilities: Alibaba Cloud offers partners access to numerous digital trans- formation opportunities given its data & analytics expertise
IBM	Focus on regulated industries: IBM has outlined its focus on regulated industries such as financial services which has allowed them to create innovative solutions based on their computing technology Container Management: IBM's OpenShift container management software is the dominant software in the market and supports IBM in extending its service offering via other software solutions
Oracle	Global standard pricing: Oracle offers a unique pricing model that is standard in all regions across the world Innovative multi-cloud: Oracle offers innovative multi-cloud architecture that leverages their database & data warehousing technologies

2.5 The ACES business model and exploitation plan

The definition and design of the ACES business model and exploitation plan still requires a significant elaboration effort, also considering that the overall ACES concept is in the making and that at least two, if not three TRLs are still to be bridged to come near the market.

It also needs to be considered that setting up an Edge computing system requires a close interaction and cooperation with client organisations and therefore needs to build on specific use cases and case studies, as the ones which the ACES project will pilot.

The ACES business model and exploitation plan will be developed in the next iterations of the present document (D6.6). First, consolidating the consortium members' position and strategy. Second, starting to reach out to Edge-cloud system experts and stakeholders to interactively define the elements of the advancing market scenario and the determinants of the development of the ACES solution.

The conception of this chapter, which will likely require several iterations and will grow across more than one version of this deliverable, will involve:

- The ACES technical team
- The functional specifications of the service
- Partner's team interviews
- User's interviews
- Partner data, if available.

The goal is to determine the potential market scenario, bridging the Technology Readiness Levels ahead of the market stage. The approach will determine:

- The ACES business model
- The Competitive edge.
- Sharing of views and strategies of the individual partners
- The roles, the exploitation objects, the market and technological targets.
- Proposing a starting model and, where possible, an initial target scenario
- Completing the SWOT analysis and possibly using the business model canvas as an evolutionary representation tool.

The ACES team will use different approaches and tools to fulfil this task:

- Surveys and interviewes
- Focus groups
- Webinars.

2.6 IPR management and licensing

This complex issue requires the precise identification of the different knowledge elements which make the ACES Edge service. This section is based on the ACES KERs and the mapping of ownership. Sections 2.6.1, 2.6.2 and 2.6.3 outline the provisions included in the Consortium Agreement signed by all partners.

2.6.1 Management of Results

Ownership

Results are owned by the partner that generates them. Joint ownership is governed by Grant Agreement Article 16.4 and its Annex 5, Section Ownership of results, with the following additions. In the absence of a joint ownership agreement, or pending its conclusion, or unless otherwise agreed: (i) each of the joint owners shall be entitled to use their jointly owned Results for non-commercial research and teaching activities on a royalty-free basis, and without requiring the prior consent of the other joint owner(s); (ii) each of the joint owners shall be entitled to otherwise Exploit the jointly owned Results and to grant non-exclusive licences to third parties (without any right to sub-license), if the other joint owners are given: (a) at least 45 calendar days advance notice; and (b) fair and reasonable compensation. The joint owners shall agree on all protection measures and the division of related costs in advance.

Transfer of Results

Each consortium partner may transfer ownership of its own Results, including its share in jointly owned Results, following the procedures of the Grant Agreement Article 16.4 and its Annex 5, Section Transfer and licensing of results, sub-section “Transfer of ownership”. The transferring Party shall, however, at the time of the transfer, inform the other Parties of such transfer and shall ensure that the rights of the other Parties under the Consortium Agreement and the Grant Agreement will not be affected by such transfer. Any addition after the signature of the Consortium Agreement requires a decision of the General Assembly.

2.6.2 Dissemination

The confidentiality obligations set out in Section 10 of the Consortium Agreement apply to all dissemination activities described in the Section 8.4 of the Consortium Agreement.

During the Project and for a period of 1 year after the end of the Project, the dissemination of own Results by one or several partners including but not restricted to publications and presentations, shall be governed by the procedure of Article 17.4 of the Grant Agreement and its Annex 5, Section Dissemination, subject to the following provisions. Prior notice of any planned publication shall be given to the other partners at least 20 calendar days before the publication. Any objection to the planned publication shall be made in accordance with the Grant Agreement by written notice to the coordinator and to the Party or Parties proposing the dissemination within 10 calendar days after receipt of the notice. If no objection is made within the time limit stated above, the publication is permitted.

Any objection to the planned publication shall be made in accordance with the Grant Agreement in writing to the coordinator and to the Party or Parties proposing the dissemination as follows:

- For dissemination of short publications, the objection(s) should be given as soon as possible and no later than 7 (seven) calendar days after receipt of the notice,
- For long publications, the objection(s) should be made within 15 calendar days after receipt of the notice.

If no objection is made within the time limit stated above, dissemination is permitted.

An objection is justified if:

a) the protection of the objecting Party's Results or Background would be adversely affected,

b) the objecting Party's legitimate interests in relation to the Results or Background would be significantly harmed,

c) the proposed publication includes Confidential Information of the objecting Party.

The objection must include a precise request for necessary modifications.

If an objection has been raised, the involved partners shall discuss how to overcome the justified grounds for the objection on a timely basis (for example by amendment to the planned publication and/or by protecting information before publication) and the objecting partner shall not unreasonably continue the opposition if appropriate measures are taken following the discussion. The objecting partner can request a publication delay of not more than 90 calendar days from the time it raises such an objection. After 90 calendar days the publication is permitted, provided that the objections of the objecting Party have been addressed.

A partner shall not include in any dissemination activity another partner's Results or Background without obtaining the owning partner's prior written approval, unless they are already published. The partners undertake to cooperate to allow the timely submission, examination, publication and defence of any dissertation or thesis for a degree that includes their Results or Background subject to the confidentiality and publication provisions agreed in the Consortium Agreement. The partners' prior written approval is required in order to use the name of the partners or any of their logos or trademarks.

2.6.3 Access Rights

In the Consortium Agreement, the partners have identified and agreed on the Background for the Project and have also, where relevant, informed each other that Access to specific Background is subject to legal restrictions or limits. Anything not identified in the Consortium Agreement shall not be the object of Access Right obligations regarding Background.

Any partner may add additional Background to the Consortium Agreement during the Project provided they give written notice to the other partners. However, approval of the General Assembly is needed should a Party wish to modify or withdraw its Background. Each partner shall implement its tasks in accordance with the Consortium Plan and shall bear sole responsibility for ensuring that its acts within the Project do not knowingly infringe third party property rights.

Any Access Rights granted exclude any rights to sublicense unless expressly stated otherwise in writing by all the partners concerned. Access Rights shall be free of any administrative transfer costs. Access Rights are granted on a non-exclusive basis. Results and Background shall be used only for the purposes for which Access Rights to it have been granted. All requests for Access Rights shall be made in writing. The granting of Access Rights may be made conditional on the acceptance of specific conditions aimed at ensuring that these rights will be used only for the intended purpose and that appropriate confidentiality obligations are in place. The requesting partner must show that the Access Rights are needed.

Access Rights for implementation

Access Rights to Results and Background needed for the performance of the own work of a partner under the Project shall be granted on a royalty-free basis.

Access Rights for Exploitation

Access Rights to Results if needed for Exploitation of a partner's own Results shall be granted on Fair and Reasonable conditions. Access rights to Results for internal research and for teaching activities shall be granted on a royalty-free basis. A request for Access Rights may be made up to twelve months after the end of the Project or, in the case of Section 9.6.2.1.2 of the Consortium Agreement, after the termination of the requesting partner's participation in the Project. For the avoidance of doubt any grant of Access Rights not covered by the Grant Agreement or the Consortium Agreement shall be at the absolute discretion of the owning partner and subject to such terms and conditions as may be agreed between the owning and receiving partners.

As regards Results developed before the accession of a new partner, the new partner will be granted Access Rights on the conditions applying for Access Rights to Background. Access Rights granted to a Defaulting partner and such partner's right to request Access Rights shall cease immediately upon receipt by the Defaulting partner of the formal notice of the decision of the General Assembly to terminate its participation in the consortium. A non-defaulting partner leaving voluntarily and with the other partners' consent shall have Access Rights to the Results developed until the date of the termination of its participation. Any partner leaving the Project shall continue to grant Access Rights pursuant to the Grant Agreement and the Consortium Agreement as if it had remained a partner for the whole duration of the Project.

The general provisions for Access Rights provided for in Section 9 of the Consortium Agreement are applicable also to Software. Partners' Access Rights to Software do not include any right to receive source code or object code ported to a certain hardware platform or any right to receive respective Software documentation in any particular form or detail, but only as available from the Party granting the Access Rights.

3 Individual exploitation plans

This section presents the initial exploitation ideas of the ACES consortium members. Along the refinement of the business model, market scenario and the exploitation actions, this section will be developed and improved.

3.1 HIRO

HIRO's exploitation strategy is focused on the development of Edge platform components, which provide its customers value added in the configuration, flexibility, resilience, orchestration, performance management (latency, efficiency, effectiveness, energy impact).

HIRO aims to develop a full-fledged powerful edge infrastructure going very much beyond the hardware and network components and embedding high value-added functions, which allow:

- seamless IOT network integration and linkage with lawful interception systems.
- mobile Edge Computing enabling the provision of the edge-cloud and IT services within the close proximity of a mobile base station to obtain extreme low end-to-end latencies and compute intensive applications in resource constraint devices and machines with the collaborative computing involving the cloud servers.
- operation of a hybrid infrastructure, a mix of on-premises, cloud, and outsourced data centre services to meet the needs of risk mitigation of Enterprises and Institutions
- real time data stream processing, building on low latency and high throughput analytical processing as an effective alternative to off-loading processing to high-capacity edge data centres.
- setup and operation of a hyperconverged infrastructure incl. storage enabling the working & learning environments address mission-critical challenges and ensures better performance and higher SLA's from key applications, meeting the growing storage demand for enterprises, optimizing rackspace with scalable storage solutions.
- The provision of hybrid cloud models, guaranteeing control, security, and availability on premise and of premise scalability.

HIRO MicroDataCenters has defined an exploitation strategy, which will, once the ACES solutions will be developed and tested, embed them in its technology offering.

The ACES RIA will be the route not only towards innovative technological developments, but also to the comprehensive analysis of the market scenarios to define the appropriate solutions configuration.

This version of the ACES exploitation strategy shall be considered just the first step towards the full-fledged marketing of this innovative Edge-cloud concept.

3.2 INESC-ID

INESC-ID main technical contributions to ACES include advancing state-of-the-art on network intelligence, security, privacy, availability and performance by developing robust AI-based network solutions, edge-based acceleration techniques, and synthesis-based mechanisms for the automatic generation of network functions.

Our main goals involve publishing the project results in top-tier scientific conferences and high-impact journals in networking, systems, security, and AI venues. Some project outcomes may also be incorporated into INESC-ID/IST University of Lisbon teaching curriculum, seminars, summer schools and thesis topics. The results developed within the project will support INESC-ID as a major research institute in networked systems and security in Europe, helping attract new talent and broaden research collaborations with other research institutions and industrial partners in these areas.

3.3 IPTO

IPTO, under the ACES Project, will deliver three use cases focused on: Marketplace & distribution, Distributed Process Management and IoT based Asset Monitoring and Management. These use cases will investigate matters of improved grid stability through congestion management and voltage control, improved asset health and predictive maintenance operations which will be exploited by enriching the company's technical expertise and knowhow and hopefully set guidelines for future investments.

The modules supporting the use case will utilize the distributed architecture of ACES for improved performance, autopoiesis and scalability adding to the overall technical knowhow within IPTO in relevant matters.

Furthermore, through ACES, IPTO will procure 3 or 4 high-end Data Centers, that will be utilized for the implementation of the Use Cases. This infrastructure, along with the ACES platform with its components and outcomes (Autopoiesis cognitive edge-services cloud, Cognitive framework, Autopoiesis framework, Swarm orchestration, Workload placement and optimization, and Application store), and the gained technical knowledge will be exploited in other R&D Projects, as those elements build a very robust infrastructure for such endeavours.

The impact of the use cases will be communicated to partners in the engineering and energy sector through publications, exhibitions, conferences, workshops, seminars, press releases, promotion materials, and web presence as well as within the premises of the ENTSO-E.

3.4 SIXSQ

As described in Section 2.3.1.4, SixSq provides solutions for edge and cloud infrastructures management to the market.

On the one hand, SixSq will bring to the ACES project its knowledge of edge, cloud and edge-to-cloud domains, and will share its expertise in running a commercial edge management platform and marketplace for AI applications.

The ACES project will also benefit from the Nuvla software itself, as an edge-to-cloud platform which offers automatic discovery, categorisation, and data routing of IoT peripherals that are connected to the NuvlaEdge appliance software.

On the other hand, ACES will give SixSq the opportunity to improve the Nuvla.io and NuvlaEdge software to meet the project challenges, especially by exploring the possibility of adding an intelligence layer at the edge. ACES will potentially bring new cognitive features to the Nuvla product, in terms of workload placement and optimization, functionalities which could bring Nuvla to the next level.

In fact, the lessons learnt from ACES will benefit Nuvla as a tool for all SixSq customers, especially service providers in the field of telecommunications. As most of them use universal Customer Premises Equipment (uCPE) infrastructure to deploy applications which virtually generate network components (network monitoring, security software for instance), improving Nuvla with new and updated features related to the performance optimization of the device will be of added value for them.

More broadly, these features will enhance both the performance and the fluidity of the infrastructure. This should reduce latency, enable a better and faster management of the edge devices, and reduce the overall energy consumption of the system. All in all, this will result in cost savings for the customers.

To finish, the Nuvla software is also involved in several Proofs-of-Concept in the fields of public transport, logistics, energy, health data management. Improvements to the software resulting from ACES will directly benefit the remote management of edge devices, and deployment of applications in these contexts.

All the exploitable results and achievements made through ACES will be shared on our website, social media, and events the SixSq team is attending.

3.5 LAKESIDE LABS

Lakeside Labs is a research organization that contributes to R&D of novel technologies in the field of self-organizing networked systems. Lakeside Labs is not bringing services or products to the market on its own but transfers the developed technologies to company partners.

Lakeside Labs will bring its expertise in swarm technologies to the ACES project and will collaborate with the project partners in generating exploitable results in terms of swarm-based methodologies for e.g. workload placement and orchestration that is planned to be part of future EMDCs facilitating autopoietic behaviour.

The exploitable results for Lakeside Labs are therefore novel swarm-intelligence based coordination algorithms used to manage WL distribution, resource management (caching, storage, etc.) and for orchestration of infrastructure beyond EMDC-limits.

One exploitation channel are ACES-follow up projects with the ACES-partners to further increase TRL-levels of the swarm solutions for the current coordination and orchestration problems in ACES or to further expend swarm technologies into coordination of further parts of the infrastructure (e.g. coordination of the IoT-edge layer).

Lakeside Labs will also endeavour to bring the knowledge and experience gained in ACES to other fields of application in the robotics-, energy-, smart production-, mobility/logistics-domains. Lakeside Labs will publish the results in peer reviewed journals and on scientific conferences and will

3.6 UNIVERSIDAD POLITÉCNICA DE MADRID

UPM will apply the ACES KERs to enrich two taught programmes: GITST (Grado de Ingenieria de Sistemas Telecomunicación), and MUIT (Master Universitario de Ingeniería de Telecomunicación). Both of them are undergoing a major programme review, with changes across the whole syllabus. The innovations in distributed systems, edge computing, and applied ML including Autopoiesis obtained in ACES will be integrated into the process to further improve the resulting quality of the teaching. Moreover, every year, at least two BSc /MSc project topics in the lines of ACES will also be offered ti prospective students.

3.7 MARTEL INNOVATE

Through its contribution to the ACES project, Martel aims to further consolidate its experience and position in the cloud-edge-IoT continuum. As a result, Martel's position to offer new innovation management and consulting services in the R&D&I context with a focus on cognitive cloud/edge services, AI/ML-infused cloud services will be strengthened. Martel will significantly expand its network of contacts in the European Cloud-Edge-IoT continuum, which will enable the company to take a more active role in shaping new project proposals and contributing to the next generation Cloud, Edge and IoT computing. Martel will leverage its network to bring together potential partners from academia, industry, and SMEs in strategic EU initiatives such as EUCloudEdgeIoT, in order to enhance the visibility of ACES results and their further use beyond the end of the project. Additionally, Martel will contribute to the implementation of tailored communication and dissemination of the exploitable assets of ACES, which will allow it to effectively disseminate information about the company's findings and best practices to stakeholders and the wider public. In its capacity of a research and technology consulting company, as well as the main developer and vendor of the Orchestra Cities platform, Martel Innovate will contribute to and exploit a selection of Key Exploitable Results from ACES, as detailed in Section 2.3.1.7 of this document. The strong momentum of the ACES project towards far edge, hyper-distributed applications leveraging multi-agent systems and swarm intelligence will provide scouting, new ideas, and concrete validation to evolve the cloud-native architecture and implementation of Martel's open-source platform. Beyond the core product improvement, this will also strengthen Martel's impact on the technical associations and initiative it is part of (e.g., FIWARE, 6G-IA). From a more market-oriented perspective, it is envisaged that ACES will grant Martel a stronger, better market positioning and presence in both the Smart Cities and Industrial IoT domains, boosting the breadth and effectiveness of Artificial Intelligence within Martel offering, with a particular focus on small-data and coordination-centric trustworthy AI.

The individual exploitation strategy outlined here (and to be considered a first iteration that will be evolved throughout the project lifetime) is mostly horizontal, focussing on Cloud-Edge-IoT computing infrastructure and generally reusable AI approaches. However, there is also a verticalized strategy that could apply to Martel: the planned ACES use cases are all within the Energy domain, which is of Martel's interest for a number of reasons. On the one hand, it is contiguous to domains such as Smart Cities where Martel already operates with established partnerships with energy utilities. Then, it is one of the domains of European interest that are targeted by other projects where Martel participates, as well as being of paramount importance for current and future initiatives such as the European Data Spaces and the whole twin digital/green transition for Europe. No further position is done by Martel at this moment, as the project activities related to the use cases haven't fully started yet; it is nevertheless already clear that the Energy domain, the three specific ACES use cases within it, and the competence and commitment of the IPTO partner (and of the other ones involved) are all promisingly relevant for Martel's exploitation.

3.8 TU DARMSTADT

TUD is a leading European university of technology performing world-class research in cybersecurity. Through the ACES project, TUD's contributions are to develop novel scientific results and technical approaches for solving security problems of distributed AI algorithms and applications. In particular, TUD aims to develop a resilient federated learning framework against poisoning and adversarial attacks. The framework will combine several novel approaches to identifying and mitigating poisoned model updates as well as protecting the privacy of users' data. The project research outcomes will be published in top-tier scientific conferences and high-impact journals in security and AI research. Further, pertinent aspects of project outcomes will also be incorporated into TUD's teaching curriculum, seminars, summer schools, and thesis topics.

3.9 UNIVERSITY OF LJUBLIANA

In the scope of ACES, UL will contribute towards advancing state-of-the-art by developing novel techniques for edge-wide workload placement and edge-wide workload optimization that are currently unavailable in container-based scheduling systems.

As the project progresses, the potential for commercially exploiting and disseminating the outcomes of ACES will be identified. UL will submit any suitable results and outcomes to high-ranking scientific journals as well as conferences to reach the wider research community. Relevant outcomes of ACES will also be covered in courses on distributed systems, data-intensive systems, and software engineering, as well as various thesis topics, workshops, and seminars.

3.10 IDSIA-SUPSI

Through its active participation in the ACES project, as previously described in Section 2.3.1.10, IDSIA aims to fortify its expertise and position within the realm of intelligent multi-agent systems. This endeavour will enable IDSIA to transfer state-of-the-art technologies and knowledge to its company partners, fostering advancements in novel research directions.

The ACES project serves as a platform for IDSIA to establish new collaborations, both with consortium partners and through networking activities integrated into the project. IDSIA's dedicated efforts in developing innovative machine learning approaches within the domains of autopoiesis and intelligent multi-agent systems will yield exploitable outcomes such as publications, participation in conferences, panels, and the formation of new partnerships. These collaborations encompass research collaborations with universities and research centres, as well as engagements with interested company partners, further promoting advancements in the field.

3.11 DATAPOWER

DataPower focuses on digital transformation, on process design and optimization and on their redesign based on digital technologies. We integrate digital technology into all areas of a business and of an administration or smart city, producing fundamental changes to how businesses and institutions operate and how they deliver value to customers and stakeholders. DataPower facilitates the development, helping organisations, where necessary, to walk away from long-standing business processes that companies were built upon in favour of relatively new practices that are still being defined.

DataPower sees edge computing as a breakthrough opportunity to boost the performance of information systems, in particular the ones which rely on remote processes and distributed value chains. DataPower is not interested in the technological solution per se but strongly in the application potential and therefore the process innovation and technological breakthroughs which can be achieved in process management in businesses, administrations, smart cities, etc.

The main value of ACES for DataPower therefore is placed in consulting for the digital transformation of processes of their clients and in the research and innovation in new concepts and solutions.

To enable and support this development line, DataPower will make available its qualitative and quantitative toolbox to define marketing and business plans, to lead businesses in the choices and configuration of their systems in the edge-cloud-IoT continuum.

4 Preliminary Conclusions in Month 9 of the project

The ACES-Edge computing RIA pursues the design, implementation, and testing of and edge-cloud system to offer customers who need distributed services a more efficient, effective, sustainable, resilient, performant and secure solution to effectively manage distributed processes.

ACES is an ambitious 36-month research and innovation action, which aims to push out a flexible Edge service exploiting the advantages of autopoiesis, self-configuration, self-organisation, self-management, and self-creation, using various technologies enabled by Artificial Intelligence and Machine Learning.

An increase demand for Edge-cloud services has been recognised, but migrating to adopt edge-cloud based processes requires a completely different business and market model. The ACES team is aware of the fact that, even though the building blocks of the Edge system are identified and characterised, their custom implementation and operation requires research and innovation work before entering the testing and validation stage. Even though we are adopting and iterative and spiral development approach, the critical path needs to be respected and before developing any business model and approach to the market, the foundations of the ACES Edge, its concept and building blocks need to be in place and functionally validated before any market-oriented validation can be initiated. The market-oriented validation, which in part concerns the horizontal edge computing functions but also, and as important, the vertical use cases and their specificities, which will start with the energy sector and will be extended according to the work plan.

In month 8 of the RIA timeline the system concepts and components have been identified and characterised. The current research and innovation work is focused on the completion of the blueprint and on the refinement of the implementation of the single components of the ACES Edge architecture. More work will be necessary to define the ACES business model and market approach, however.

The current document is due to be revised and updated and the next revision is scheduled for month 12 of the timeline. The description and design of exploitation activities will go hand in hand with the technological innovation and system development, maintaining a close connection between the technological specification and development of the exploitation activities.

The next revision of this document will therefore include:

- An improved version of the functional specifications of the ACES Edge service;
- The more detailed characterisation of the system components, the Key Exploitable Results;
- The claims of consortium members of the specified KERs;
- The specification of the three case studies of the energy sector, starting to define a cost-benefit analysis;
- The specification of the approach to the ACES Edge cost-benefit model compared to a centralised cloud service;
- A refinement of the Cloud and Edge market trends and gaps and an initial review of use case requirements with vertical stakeholders.

Special care will be placed in identifying ongoing research and innovation activities within EU funded programmes and outside, which deal with similar technological issues and market challenges. A structured assessment will be undertaken to support networking, interaction, cooperation, and solutions sharing.