Fraunhofer

Fraunhofer Cluster of Excellence Cognitive Internet Technologies (CCIT)

Expression of Interest (EoI)

»EoI 2025-2027«

April 14, 2023

Prof. Claudia Eckert
Spokesperson for the Steering Committee, Fraunhofer CCIT
Managing Director, Fraunhofer Institute for Applied and Integrated Security AISEC

Michael Fritz
Head of office, Fraunhofer CCIT

1. Motivation

The Fraunhofer Cluster of Excellence Cognitive Internet Technologies (CCIT) combines the competencies of a number of Fraunhofer Institutes in the key technologies of IoT, data spaces and AI/ML into a common solution portfolio. To this end, the first-class research and development work anchored in the institutes is brought together in the Fraunhofer CCIT and linked through joint research and development projects in selected application domains.

The Fraunhofer CCIT's goal for the coming years is to research, test and apply cognitive technologies for data-intensive, networked and highly automated systems. Due to the different needs and maturity levels of digital transformation in different industries such as automotive, health, manufacturing or agriculture, CCIT is developing a spectrum of adaptable and modularly combinable trustworthy technology building blocks to realize the technological concept of the Edge Cloud Continuum (ECC) from sensor to cloud.

Due to the success of its own research and development in recent years, Fraunhofer CCIT can draw on an extensive portfolio of solutions from trusted sensor technology/IoT, sovereign data space technologies and new approaches of hybrid AI/ML. CCIT research has already developed integrated, innovative cognitive platform solutions to enhance existing technologies with e.g. security, AI or sustainability features in an integrated, transparent and controllable way. CCIT has also demonstrated the added value of combining sensor/IoT, dataspace and AI technologies by building dedicated sector-specific MVPs such as Smart Screw, Smart Notch, Smart Tool or Smart Intersection.

2. Current Status

Today's IT landscape is characterized by multiple layers of cognitive processing: cloud, edge, IoT devices. However, these tiers often follow very heterogeneous and different approaches to infrastructure, technology or development methodologies.

The use of cloud computing (as public cloud, private cloud or multi-cloud) has become indispensable for users in most domains (e.g. manufacturing, energy, logistics, healthcare or agriculture) and is described by the keyword »cloudification«.

In addition, edge computing is becoming increasingly important and is finding its way into IT infrastructures (»edgification«). The reasons for this are primarily constraints such as latency requirements for real-time applications, high costs and large carbon footprints for continuous and especially unprocessed cloud loads, or security concerns with cloud platforms and the need to meet security and compliance requirements.

In addition, the use of communication-enabled and powerful intelligent sensors (IoT) continues to advance in all application areas, continuously delivering data to hubs such as edge devices or cloud platforms. With the introduction of new communication technologies (5G/6G, LPWA), »IoT-fication« is gaining even more momentum. Current and future sensors and IoT devices will be able to perform intelligent pre-processing in controllable, local environments, using federated learning to take advantage of distributed intelligence. They will be able to automatically offload computing and storage tasks to edge and cloud devices to dynamically optimize stated requirements such as security, energy consumption, and real-time behavior.

IoT devices are already connected to edge and cloud platforms, thus data transport across these different layers is not an issue. However, the great potential of data-driven infrastructures cannot be realized by the prevailing silo-like layers. What is missing is an AI-driven, intelligent management middleware that provides intelligent, multi-tier services to automatically manage applications and data running on heterogeneous devices. The envisioned policy-driven management middleware must provide adaptive, verifiable, and trusted services to balance and dynamically move the execution of applications and the required data depending on the current

context and the policy rules (e.g., GDPR compliance, carbon footprint) to be met. The question of interoperability and common interfaces is fundamental.

Common to all of the »-fications« mentioned above is the lack of such interfaces, the lack of appropriate interoperability between infrastructure layers, and the lack of a common policy-driven layer that provides intelligent and trusted management for future cognitive applications from the sensor to the cloud. Lack of interoperability leads to high costs, including manual management and high energy costs. Existing silos lead to sub-optimal decision making because the required data is not available where it is needed. In addition, the silos suffer from a lack of agility to respond quickly to new requirements and contexts.

There are existing solutions for cost optimization (e.g., load balancing between cloud providers), climate compatibility (e.g., carbon footprint of the infrastructure used), and cybersecurity compliance. However, these currently have the character of isolated solutions and still require the use of specialized professionals to achieve comprehensive, holistic and systemic optimization. Specialists who are hard to find in a competitive job market. Combined with the increasing demands for compliance with regulatory requirements (e.g. Cyber Resilience Act, AI Act, Data Act, Digital Services Act, Corporate Sustainability Reporting Directive, etc.), the need for automated assistance solutions to meet the multiple requirements with less use of resources is accelerating.

3. Research Challenges

The key challenges to achieving a Cloud Edge IoT Continuum (CEIC) for individual organizations must be divided into several equally important areas. However, it is imperative to work holistically to realize the full value of the CEIC, such as increased efficiency, cost savings, compliance with regulatory requirements such as climate targets, security concepts, or data usage control and sovereignty.

1. »ANY-TO-ANY-Infrastructure«

With any-to-any, we envision connecting the infrastructure layers of sensors/IoT, edge computing and cloud computing to the familiar CEIC. In addition, future expansion to include High Performance Computing (HPC) and Quantum Computing (QC) capabilities will be necessary to keep increasing data volumes and complex interrelationships manageable. For HPC, a short to medium term timeframe (by 2025) should be targeted; for QC, the foundations for interoperability between digital and quantum-based computing need to be laid; a medium to long term timeframe (2025 - 2030) is required.

2. »ZeroTouch Cognitive System of Systems (ZTCSoS)«

Across all infrastructure layers (CEIC or extensions to HPC or QC), there is a need to develop a policy-driven, intelligent, and trustworthy management architecture that provides means to address proven or evolving IT operations paradigms in a consistent manner. These include:

- SecOps: Integrating security considerations into the development and deployment process
- DevOps: Delivering software applications faster and more efficiently
- FinOps: Optimizing the cost of IT services
- GreenOps: Optimizing the environmental impact of IT and infrastructure projects
- MLOps: Developing ML models and integrating them into software applications

The goal of ZTCSoS is to automate the overall control of IT processes in the CEIC while meeting changing optimization requirements, e.g. due to business policies or external constraints.

- *Example A: Cost-optimized load balancing across an enterprise's edge or cloud/multi-cloud environments.*
- *Example B: Training a ML model with Prio 1 of lowest possible C02 load and Prio 2 compliance with a cost cap.*

- *Example C: Numerical simulation of product features under mandatory compliance with data usage control regulations, cost-optimized and as CO2-neutral as possible.*

The ZTCSoS must therefore be understood as a system for multi-dimensional optimization of a company's various IT systems. Due to the increasing complexity of these systems, constantly changing objectives and potential conflicts of objectives, the use of highly automated, autonomous systems is indispensable.

3. »R3ST Conformance« (Resilience, Safety, Security, Sovereignty, Trust)

The basis for the implementation of CEIC and ZTCSoS must be conformity with the requirements of R3ST. This includes the aspects of technical resilience (e.g. cyber resilience), operational resilience (e.g. business continuity) and organizational resilience (e.g. human in the loop) in the area of resilience. In addition, the aspects of safety (in the sense of operational safety), security (in the sense of protection against manipulation and misuse), sovereignty (in the sense of data sovereignty and digital sovereignty) and trust (in the sense of trustworthiness in the reliability, integrity and security of systems) must be considered.

4. »OPEN & EXPLAINABLE«

The implementation of the CEIC using the ZTCSoS and compliance with the R3ST confirmation must be aligned with the principles of openness and traceability. This requires the creation of software as open source solution, based on open standards and using open innovation approaches. In the context of cognitive solutions, these must be comprehensible and interpretable, i.e. created according to the methods and processes of Explainable AI.