Nils Jahnke*, Nina Großegesse, Brian-Frederik Jahnke, David Jilg, Justus Leskow, Rebecca Schmook, Alexander Schultheis, and Jiaqi Wang

Toward a Holistic Evaluation of Edge-Cloud Systems: Insights from a Task Force

Abstract: Edge-Cloud Systems (ECS) are essential for realizing the potential of Internet of Things (IoT) data by integrating software and hardware components along the Edge-Cloud Continuum and combining the advantages of both computing paradigms. Given their diverse configurations and viable alternatives such as cloud-only solutions, ECS must be systematically assessed to demonstrate their tangible benefits to stakeholders. However, existing evaluation and benchmarking approaches across industry and academia remain fragmented, and a holistic evaluation framework is lacking. This article synthesizes the outcomes of a 2024–2025 task force involving 27 researchers and practitioners from Germany and Austria, consolidating existing approaches, tools, and metrics toward a comprehensive framework for ECS evaluation. The framework defines 16 evaluation criteria across three dimensions: technology (e.g., reliability, availability, resilience), organization (e.g., maintenance and administrative effort, interoperability, cost), and environment (e.g., regulatory compliance, data localization, environmental impact). Each criterion is linked to context-specific metrics and measurement methods. We illustrate the framework's practical application through use cases in networking, edge-enabled Artificial Intelligence, and carbon accounting. Finally, drawing on task force insights, we propose a practice-oriented research agenda to advance holistic ECS evaluation.

Keywords: Edge-Cloud Systems, System Evaluation, Edge-Cloud Use Cases, Edge Computing, Cloud Computing

***Corresponding author: Nils Jahnke,** Fraunhofer ISST, Mobility & Smart Cities, 44147 Dortmund, Germany, e-mail: nils.jahnke@isst.fraunhofer.de
Nina Großegesse, Salzburg Research Forschungsgesellschaft m.b.H., Intelligent Connectivity, 5020 Salzburg, Austria, e-mail: nina.grossegesse@salzburgresearch.at
Brian-Frederik Jahnke, Rebecca Schmook, Fraunhofer ISST, IT Service Providers, 44147 Dortmund, Germany, e-mail: brian-frederik.jahnke@isst.fraunhofer.de, rebecca.schmook@isst.fraunhofer.de
David Jilg, German Research Center for Artificial Intelligence (DFKI), Branch Trier University, 54296 Trier, Germany, e-mail: david.jilg@dfki.de
Justus Leskow, Hochschule Osnabrück, Faculty of Management and Technology, 49809 Lingen, Germany, e-mail: justus.leskow@hs-osnabrueck.de
Alexander Schultheis, German Research Center for Artificial Intelligence (DFKI), Branch Trier University, 54296 Trier, Germany, e-mail: alexander.schultheis@dfki.de, Artificial Intelligence and Intelligent Information Systems, Trier University, 54296 Trier, Germany, e-mail: alexander.schultheis@uni-trier.de
Jiaqi Wang, Robert Bosch GmbH, AI in Production, 71272 Renningen, Germany, e-mail: jiaqi.wang5@de.bosch.com

1 Introduction

As *Internet of Things* (IoT) devices are expected to double from 2025 to 2034 [57], an unprecedented amount of data from infrastructures, operations, and human behavior is becoming available. To generate business, ecological and societal benefits from this data, it needs to be collected, aggregated, analyzed and leveraged to improve business processes, decision making, or product development, potentially by using *Artificial Intelligence* (AI) technologies [14, 40].

Edge-Cloud Systems (ECS) are crucial to realizing the potential of IoT data by combining the benefits of different resources along the continuum from edge to cloud [63]. By placing compute and storage capacities near the data sources (at the edge), data preparation and analysis are performed in a decentralized manner, which ensures low latency, high availability, efficiency, and data protection among other critical success factors of data-driven applications [42, 58]. On the other hand, cloud resources are leveraged for tasks such as system orchestration, data aggregation and provisioning, and computing-intensive tasks [23]. Additionally, network technologies account for connectivity between the used resources [26].

When developing an ECS for a specific practical context, a thorough evaluation is needed to define the most suitable configuration of hardware and software resources able to fulfill the needed tasks[58], to demonstrate tangible advantages over possible alternative solutions such as cloud-only approaches[27], and, more generally, to prove the added value of the developed edge-cloud application to its stakeholders, e.g., ini terms of environmental sustainability. However, evaluating ECS comes with its complications. Their inherent heterogeneity and evolving hardware and software components as well as application-specific stakeholder interests preclude a universal evaluation approach [5]. Instead, evaluators must identify assessment criteria and select appropriate metrics and measurement methods tailored to the application context. Additionally, potential trade-offs between different criteria need to be considered. Yet, research and practice lack a comprehensive overview of the possible evaluation criteria of ECS as well as in-depth insights into the evaluation of dedicated ECS components. For example, current publications on the evaluation of ECS mainly consider technical criteria [3, 5, 31] and neglect factors such as data sovereignty or environmental sustainability, which become more and more important for the acceptance and success of IT applications in many domains [38, 47]. However, in practice, such isolated evaluation approaches may neglect relevant stakeholder concerns and undermine the success of applications based on ECS.

This study reports on the collective outcomes of the task force "*Evaluation of Edge-Cloud Systems - Approaches, tools, and metrics*" [1]. The task force brought together insights from 27 collaborators from industry and academia, all affiliated with edge computing research projects in Germany and Austria across domains such as manufacturing, smart living, energy, telecommunications, and software services. Four half-day workshops were conducted between 2024 and 2025. The aim of the task force was to present and discuss individual approaches to ECS evaluation and consolidate them into a holistic framework for ECS evaluation that goes beyond the isolated approaches present in the current literature. It focused on gaining a structured overview about criteria relevant for ECS evaluation, gathering greater insights about the practical application of these criteria in real-world applications, and deriving currently pressing challenges as well as possible paths for future research related to the evaluation of ECS. In detail, we provide a holistic overview of the criteria relevant for ECS assessment, present its practical applicability through selected use cases developed by task force participants and a practice-inspired agenda for future research in the domain.

The remaining sections of this article present our findings as follows: Sect. 2 introduces the background and related work, outlining the concept of ECS and elaborating on existing research in the area of their evaluation. Sect. 3 presents the developed framework of evaluation criteria, structured into technological, organizational, and environmental dimensions. Subsequently, Sect. 4 illustrates the evaluation of ECS through selected use cases, focusing on network performance, AI-based applications, and industrial scenarios emphasizing sustainability. In Sect. 5, we present current challenges and our agenda for future research on ECS evaluation derived from the task force insights. Finally, Sect. 6 concludes the paper.

2 Background and Related Work

This section presents the background required for this work. In Sect. 2.1, the concept of ECS and their role within the *Edge-Cloud Continuum* (ECC) are introduced. Subsequently, Sect. 2.2 outlines preceding work regarding the assessment of ECS in practical contexts and elaborates on the focused evaluation criteria.

2.1 Edge-Cloud Systems

ECS represent a paradigm shift in distributed computing, combining the strengths of edge and cloud architectures to address the growing demands of modern applications. What Edge Computing specifically entails is however still contested in the literature. Willner and Gowtham [62] shed light on the cascading evolution of distributed computing systems, changing from centralized to local data processing, following the trends driven by the "Vs of big data" which are (data) volume, variety and velocity. An edge computing definition provided by LF Edge, a project under the Linux Foundation, highlights the improvements in performance, operating costs and reliability of applications and services [29]. As emphasized in recent literature [2, 9], the surge of interconnected devices and the limitations of traditional cloud-centric models, such as high latency, bandwidth constraints, and privacy concerns, have driven the evolution toward the ECC. This continuum is composed of various resources that process data along the data value chain [54]. Computational resources are strategically relocated closer to data sources, enabling real-time processing, improved responsiveness, and optimized resource utilization. In addition, the ECC architecture enhances data protection, as it allows organizations to process sensitive data locally instead of transmitting it to external centralized cloud platforms [49]. A key feature of ECS is their context-specific deployment: resources and services are distributed across geographically dispersed nodes, tailored to the requirements of each application domain. This approach improves system reliability and fault tolerance by minimizing the impact of failures and security breaches. The literature review reveals that low latency is the most cited benefit, making ECS essential for time-sensitive and large-scale real-time applications, such as industrial automation, IoT analytics, and AI-driven processes.

Architecturally, ECS are characterized by their distributed and decentralized nature, often acting as intermediaries between end devices and centralized cloud services. They support resource-constrained devices by offloading intensive tasks, thereby extending device capabilities and operational lifespans. For such approaches efficient and high-performance data transmission is crucial. Furthermore, strict partitioning and data protection mechanisms are integral, ensuring compliance and security in sensitive environments.

Despite their advantages, ECS face challenges, including hardware limitations at the edge, the need for unified definitions, and the complexity of optimal node placement. Ongoing research focuses on standardizing terminology, refining architectures, and developing context-aware frameworks to fully realize the potential of edge-cloud integration in future information technology landscapes.

2.2 Evaluation of Edge-Cloud Systems

Comprehensive evaluation approaches are essential to address the challenges inherent in ECS, such as ensuring *Quality of Service* (QoS), maintaining secure data processing, achieving system reliability, mitigating network congestion, and managing limited computational resources at the edge. Evaluation criteria are required to support the monitoring, analysis and planning of ECS to maintain consistent performance under dynamic conditions [20]. Furthermore, well-defined evaluation metrics constitute a fundamental basis for the development of efficient and adaptive orchestration mechanisms [9]. Beyond performance optimization, such metrics contribute to a holistic understanding of system behavior, enable efficient deployment strategies, and support the balancing of operational trade-offs [45]. They also promote reproducible experimental analyses and facilitate the systematic modeling and simplification of system architectures [11], enhancing the transparency and comparability of edge-cloud research and practice.

Reference	Year	Focus topic	Technological criteria	Organizational criteria	Environmental criteria	Metrics defined
[53]	2018	simulation environment	✓	(✓)	(✓)	-
[30]	2018	trust	(✓)	(✓)	-	✓
[5]	2020	performance evaluation	✓	(✓)	✓	✓
[2]	2021	***holistic***	✓	-	(✓)	-
[25]	2021	network	(✓)	-	-	✓
[37]	2021	simulation framework	✓	-	(✓)	-
[39]	2021	resource management	(✓)	-	(✓)	-
[59]	2021	***holistic***	✓	-	(✓)	✓
[45]	2022	***holistic***	✓	-	(✓)	✓
[4]	2023	autonomy of IoT	(✓)	(✓)	-	-
[11]	2023	system modeling	✓	-	-	✓
[15]	2023	security	-	(✓)	-	✓
[20]	2024	AI-based ECS	✓	-	(✓)	✓
[19]	2024	QoS of *Mobile Edge Computing* (MEC)	(✓)	(✓)	(✓)	✓
[41]	2025	simulation framework for AI models	(✓)	-	(✓)	-
[9]	2025	AI / ML	✓	(✓)	(✓)	-
[10]	2025	hardware & AI models	(✓)	(✓)	-	✓
[64]	2025	trust	(✓)	(✓)	-	✓

Tab. 1: Systematic analysis of related work

Many related studies define evaluation criteria, however, only about half of them derive metrics or measurements, see Table 1. Further, related works often highlight a specific topic, for example, on AI or *Machine Learning* (ML) [9, 10, 20, 41]. Consequently, these studies overlook important evaluation aspects. For instance, the emphasis of Canpolat Şahin et al. [10] lies on hardware and AI models. They refer to the ISO/IEC 25010 norm for software product quality assessment, which does not account for factors such as network properties and environmental aspects. On the other hand, there are studies that explore the development of a comprehensive efficiency model for the network component [25], but do not take the entire edge-cloud infrastructure into account.

Overall, performance criteria represent the primary focus of most studies. For example, although Iftikhar et al. [20] adopt a holistic approach, the emphasis on AI-based fog and edge computing results in limited consideration of organizational and environmental perspectives. Some studies propose simulation frameworks like EdgeCloudSim, PureEdgeSim and EdgeAISim [37, 41, 53]. These frameworks are well suited for performance evaluation. While they may address factors such as cost and energy consumption, they are not designed to assess organizational or environmental aspects. Environmental aspects are often only partly covered. While many studies include the aspect of energy consumption [39, 41], which is widely recognized as a highly important topic, other environmental aspects remain insufficiently explored. Organizational aspects are the least frequently addressed in academic literature, and no study was found that provides a comprehensive analysis of this dimension. Aslanpour et al. [5] provide one of the most comprehensive analyses, offering extensive coverage of technical and environmental criteria, including performance evaluation metrics. However, organizational aspects such as interoperability and user experience are only partially addressed, and the study has become partly outdated due to rapid advancements in ECS since 2020. Therefore, we identified a research gap in this area and address it by developing a holistic perspective that integrates all relevant criteria and metrics for ECS.

3 Evaluation Criteria of Edge-Cloud Systems

Drawing on the insights from practice and research during our task force sessions, we develop a framework of evaluation criteria for ECS (see Figure 1). The framework comprises 16 evaluation criteria as well as according metrics and measurements structured into the dimensions of technology (see Sect. 3.1), organization (see Sect. 3.2) and environment (see Sect. 3.3). The framework presents a holistic overview about the criteria to be considered when evaluating ECS. During ECS evaluation, the framework supports the identification of criteria and metrics pertinent to a specific ECS configuration in its respective application domain. In the following, we discuss the evaluation criteria, which are depicted within the inner ring, starting with the technology dimension.

3.1 Technological Criteria

Of all three dimensions, technological criteria for ECS evaluation have been most widely covered by research in information technology. The technological criteria comprise characteristics essential to the technical functioning of an ECS for a given application scenario. Such criteria have dedicated metrics, which can usually be quantified or estimated through specific measurement methods. Core to an ECS are the *reliability, availability, resilience and speed* with which data processing tasks can be carried out. High demands in these areas usually go hand-in-hand with shifting workloads towards the edge. ECS integrate hardware components with varying computing and communication requirements, interconnected through networks that shape overall functionality, performance, and resilience. As these *networks* are often shared with other systems, their practical usability may be constrained. Hence, properties such as latency and bandwidth should be assessed or simulated to ensure sufficient performance for ECS implementation. Also computing devices might be shared among different applications and sufficient computing capacity, particularly at the edge, needs to be ensured. Therefore, it is crucial to assess *resource usage and utilization* to distribute computing demands across resources. More and more, AI and analytics services are becoming essential components of ECS. Dedicated criteria such as model accuracy and training efficiency are needed to measure the performance of AI components. Finally, the *scalability* of an ECS refers to its ability to adapt to fluctuating payload requirements, such as a rising number of end-user devices, which is crucial for long-term success.

3.2 Organizational Criteria

Organizational evaluation criteria refer to, e.g., the human, structural, cultural and cost aspects in the establishment of applications based on ECS. In the absence of robust technical measurement approaches, these criteria are often assessed through theoretical constructs such as surveys and observations, or, in practice, even by intuition. From a human perspective, two central criteria are the *maintenance and administrative efforts* and skills required by providers and system integrators of ECS as well as the *user experience* that is perceived by the natural persons interacting with an ECS for a specific task, which might be highly subjective. As ECS increasingly involve sharing potentially sensitive data across organizational actors, data providers must be able to trust the functioning of the technical system and the actors involved, whereas data users must also trust the data provided. ECS must typically be integrated with already existing infrastructures and applications. Consequently, one must ensure a system's interoperability. While we structure interoperability into the organizational dimension, it also comprises technical, semantic and legal aspects. Lastly, costs represent a key organizational decision criterion for ECS, with design choices often guided by cost–benefit considerations. The distributed nature of ECS complicates the identification and evaluation of relevant cost factors. While operating costs typically dominate in cloud systems, edge deployments entail higher capital expenditures.

3.3 Environmental Criteria

In our work, environmental criteria mainly target the assessment of political, legal and ecological concerns related to ECS. From a political and legal perspective, applications based on ECS are subject to increasing regulatory requirements. In the European Union, providers, deployers and users of ECS must ensure they comply with legal

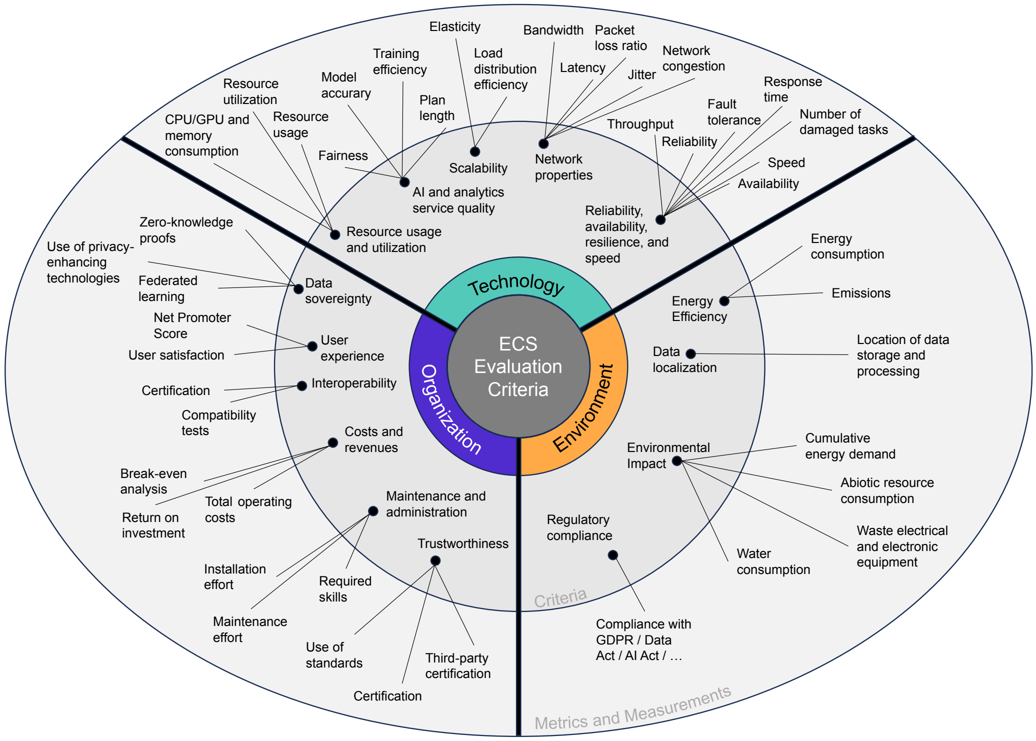

Fig. 1: Evaluation criteria for Edge-Cloud Systems (ECS)

provisions such as the General Data Protection Regulation, the AI Act, the Data Act and further, potentially domain-specific obligations. Also other jurisdictions, such as California, continue to advance regulations related to data protection and AI. Data localization is a criterion that partially relates to legal obligations but is also often required by users of ECS, especially in the public sector or critical infrastructures. Lastly, ecological effects of ECS need to be considered in the evaluation process. In the literature, these comprise first and foremost the energy efficiency of ECS. Recent advances, however, point to the need to assess the environmental impact more holistically by a) considering the cumulative energy demand across the whole life cycle of an ECS application, and b) considering the consumption of further abiotic resources like water and air, which are essential for the functioning of data centers.

4 Use Cases

This section illustrates the practical application of the previously described criteria in selected use cases stemming from the authors' research projects. In Sect. 4.1, network-related aspects are analyzed as a fundamental layer of the continuum. Sect. 4.2 focuses on the role of AI approaches within ECS, while Sect. 4.3 highlights industrial applications emphasizing sustainability and CO_2 assessment.

4.1 Network Parameter Prediction in ECS

The network interconnecting hardware components in an ECS not only provides communication functionality but also has a decisive impact on the systems performance, reliability, and resilience. Depending on the use case, different network types are employed. Sensor networks (such as Zigbee or Bluetooth Low Energy) gather data from distributed sensors and forward it to edge gateways. Local networks (WLAN / LAN) enable short-range communication between edge devices, while mobile networks (4G/5G) provide wide-area, high-speed connectivity for mobile or remote nodes. Selecting the appropriate network technology requires balancing application-specific demands such as latency, bandwidth, reliability, and mobility.
Network performance is commonly characterized by parameters such as bandwidth and latency. Bandwidth, typically measured in gigabits per second (Gb/s), indicated the maximum transmission capacity. Latency refers to the time required to transmit a data packet or data set from one network node to another and is usually specified in milliseconds (ms) or microseconds (µs). Latency

depends heavily on network utilization, due to this it is useful to specify, for example, the minimum, average, and maximum latency for a specific data set and application. Application domains differ widely in their tolerance for delay - from real-time requirements to several minutes. Industrial automation systems often impose strict real-time constraints. According to ISO/IEC 2382:2015 real-time processing denotes data processing that meets the timing requirements of an external process [21]. The key aspect of real-time behavior is deterministic response timing to external or internal events.

Furthermore several additional functional network parameters can be measured and evaluated to characterize communication performance:

- Reliability (packet loss ratio): Reliability describes the packet delivery rate between two endpoints within a specified maximum transmission duration [17].
- Throughput: In data transmission, network throughput refers to the actual amount of data successfully transmitted between two network nodes within a given period. Network throughput is usually measured in bits per second (bps), such as megabits per second (Mbps) or gigabits per second (Gbps) [5]. In contrast, nominal bandwidth indicates the theoretical maximum transmission capacity of the network.
- Network congestion: Network congestion occurs when the data load exceeds the capacity of a node or link, resulting in reduced quality of service. It can be assessed by measuring the packet loss ratio.
- Runtime variance (jitter): Runtime variance refers to the deviation in latency during packet transmission [65]. In industrial applications with real-time constraints, it is crucial for runtime variance to remain within specified tolerances to ensure deterministic behavior.

In general, the network parameters described in the previous section for assessing communication performance are measurable; however, obtaining accurate measurements often requires substantial effort and may itself affect network performance. Additionally, measurement can impose an extra load on the network, leading to performance degradation. In some ECS, constant observation of network conditions is necessary to determine how services are distributed along the ECC as well as to assess the duration, effort, and reliability of data transmission..

To address these challenges, AI-based estimation of network performance metrics offers an attractive alternative to direct measurements. An AI model designed for estimating performance parameters can be trained on data obtained from prior measurements and subsequently used to infer performance parameters without requiring continuous monitoring during operation. Since networks in ECS often exhibit dynamic topologies, so-called *Graph Neural Networks* (GNNs) [46] have proven particularly effective for modeling the problem and estimating performance parameters. GNNs are specifically designed to process graph-structured data and are invariant to permutations of nodes and edges, allowing them to generalize across graphs of varying size and structure, i.e., making accurate predictions for previously unseen graphs. Consequently, they can also adapt to topology changes and are therefore particularly suitable for use in the context of data transmission in communication networks [55]. For training, the GNN is provided with the current network topology within the ECC along with additional traffic-related parameters such as packet sizes and inter-packet intervals. Once trained, the inference model can estimate the key performance metrics directly from these inputs, eliminating the need for continous measurement during operation (see Figure 2).[1]

4.2 Artificial Intelligence Applications

In the following, AI methods applied within ECS are discussed. Sect. 4.2.1 introduces *Federated Learning* (FL) as a collaborative ML paradigm, Sect. 4.2.2 outlines the use of AI planning for distributed and adaptive decision-making, and Sect. 4.2.3 presents *Case-Based Reasoning* (CBR) as a knowledge-driven reasoning approach for analytics and control in the ECC.

4.2.1 Federated Learning

FL [36] is a collaborative ML framework in which a shared model is trained in a decentralized manner across multiple clients without exchanging their raw data. Each client, ranging from edge devices and production lines to entire organizations, performs local training and periodically shares model parameters with a central server. The server aggregates them to a global model and redistributes it to the clients.

This paradigm is particularly valuable when individual datasets are limited to train reliable models independently, and data sharing is restricted by privacy or communication constraints. Applications that satisfy these conditions represent natural candidates for FL, as the method allows

1 A detailed description of the GNN-based prediction of performance metrics can be found in [32]

Fig. 2: Graph Neural Network (GNN) topology and used metrics

sharing of learned knowledge without disclosing sensitive information. As a leading domain, healthcare [12] adopts FL to allow hospitals and research institutions collaboratively train diagnostic or predictive models without exchanging patient records. Similarly, IoT infrastructures such as smart grids [44] or vehicular networks [6] benefit from FL, since locally collected sensor data can be exploited for anomaly detection or demand forecasting safely and efficiently. From an industrial perspective, FL is particularly relevant for predictive maintenance, defect detection, and quality optimization across distributed production sites. It enables companies to integrate knowledge from heterogeneous environments and devices into stronger global models, thereby improving model performance and resilience while safeguarding proprietary data.

A major challenge of FL in practical deployments is heterogeneity inherent in distributed ECC environments, which manifests at multiple levels. Data heterogeneity is caused by non-independent and identically distributed (non-IID) data, unbalanced samples sizes, and varying label availability across clients, leading to biased models and slower convergence. System heterogeneity arises from differences in computational capacity, memory, network conditions and latency of participating nodes, which create stragglers and increase communication and energy costs. Model heterogeneity presents an additional challenge when clients have different model architectures or optimization objectives because of their data or computing capacity, creating tension between global generalization and local personalization. These various dimensions of heterogeneity directly affect the performance and efficiency of FL.

To address the mentioned challenges of heterogeneity, diverse solution strategies have been developed. Data heterogeneity can be mitigated through approaches such as personalized FL, in which the local training becomes more personalized by adopting methods such as meta learning, or clustered FL, where clients with similar data distributions are grouped for training. Data augmentation and generation provide further means of compensating for unbalanced or sparse datasets. System heterogeneity can be addressed by hierarchical FL, where intermediate aggregations are performed at the edge before the construction of a final global model in the cloud, thus reducing communication costs and accommodating resource-limited devices. Asynchronous training and selective client participation additionally help to reduce the impact of stragglers and improve efficiency while maintaining inclusiveness. Model heterogeneity is targeted through approaches such as knowledge distillation and split learning. Both methods allow models with different architectures to exchange knowledge without requiring strict alignment of parameters. These developing methods illustrate how FL can be adapted to heterogeneous industrial environments while mitigating the impact of these limitations.

An ECC industrial scenario of applying FL is cross-factory anomaly detection. Each production line deploys a local model for time series data collected by its own sensors, which differ in label availability, preprocessing, product types and sensor modalities, reflecting strong data heterogeneity. Edge devices perform local training and transmit model parameters to the cloud, where they are aggregated to a shared global model. By redistributing the improved model to all clients and repeating the loop, FL enables robust anomaly detection that continuously adapts to new data while preserving data protection and reducing transmission costs.

The evaluation of FL in an ECC setting needs to reflect both the benefits and trade-offs of the method. Model quality remains the primary benchmark, typically measured by accuracy or related performance indicators, but it must also account for robustness under heterogeneous and non-IID data distributions. Training efficiency is another key criterion, as convergence speed depends on the balance between local computation and communication costs across the continuum. Energy efficiency must be considered as well, since frequent updates increase transmission costs while prolonged local training raises the energy footprint of edge devices. Finally, fairness plays a crucial role, as evaluation must capture whether participants with different data volumes and computational capacities contribute equitably and benefit proportionally from the collaboratively trained model.

4.2.2 AI Planning

AI planning addresses state transformation problems in which a discrete world model has to be transitioned from an initial state to a desired goal state through a sequence of actions [13, 16]. The initial state represents a discrete state of the planning environment, defined by a set of facts, while the goal state describes the desired changes, such as newly manufactured products. To enable the transition, a set of operators is required, each specifying a possible transformation within the environment, for example drilling a hole into a workpiece.

The application domain of AI planning in the context of an ECC lies particularly in *autonomously acting systems*. Such systems must be capable of adapting their behavior dynamically to changing conditions in order to maintain functionality and efficiency. AI planning provides the mechanisms to generate adaptive action sequences and thus to support both internal and external processes. Internally, it enables automated resource allocation by distributing computations and data across network nodes, which is crucial in large-scale or real-time scenarios with legal or organizational constraints. Externally, planning techniques can be applied to manufacturing-related control processes of industrial partners. In this domain, planning methods are already well established in the field of business process management, where they enhance both flexibility and the degree of automation [35, 51]. The planning problem in such cases consists of deriving an executable sequence of manufacturing actions that leads to the desired product.

A major limitation of AI planning lies in its high computational complexity, which scales unfavorably with increasing problem size. Edge-cloud infrastructures help to mitigate these limitations by distributing workloads across nodes or delegating them to more powerful devices. Hierarchical planning can further reduce the complexity by first constructing abstract plans that are iteratively refined. This allows different layers of the continuum to address subproblems of corresponding granularity, thereby reducing the overall computational burden. However, the distribution of planning tasks also involves important trade-offs: classical planning problems are not easily decomposed, and the adaptation effort as well as the integration of partial results may outweigh the benefits of reduced computation time or lower plan costs. Thus, the use of distributed approaches requires a careful assessment of when the additional overhead is justified.

Another promising strategy is to combine AI planning with complementary methods such as CBR. In hybrid approaches, the planning problem can be decomposed into subproblems, each addressed by different techniques operating on distinct ECC nodes. This integration exploits the strengths of multiple AI methods while leveraging the distributed nature of the continuum. Such combinations can further improve adaptability and reduce computational load, especially when case knowledge can be reused to accelerate the generation of feasible plans.

The evaluation of AI planning in an ECC setting requires systematic approaches. On the one hand, the throughput time of the planning phase and the reduction of energy consumption are central benchmarks, as they directly affect both responsiveness and sustainability. On the other hand, the quality of the generated plans must be considered. Quality can be approximated by the length of the plan (i.e., number of steps) or by the costs of the planned actions. Plan costs can thereby serve as a proxy for planning quality, since more efficient sequences typically lead to lower operational expenditure. For the evaluation of the planning process itself, time, energy consumption, and plan costs are therefore relevant metrics that enable a differentiated comparison of centralized, distributed, and hybrid execution strategies.

A practical illustration of these concepts is found in production planning. An edge device within an ECC may detect that a process cannot proceed as intended due to unforeseen conditions identified by data analysis. It then initiates adaptive replanning. If local resources are insufficient, it may delegate the generation of a coarse plan to another edge device using CBR, before refining the result and passing the executable plan back to the workflow engine. This example highlights how ECC infrastructures enable distributed, adaptive, and efficient planning in complex industrial environments while also allowing performance to be assessed along multiple dimensions such as speed, energy efficiency, and plan quality.

4.2.3 Case-Based Reasoning

CBR is a knowledge-driven AI method that addresses problems by reusing solutions from previously encountered experiences. In contrast to data-driven approaches of traditional ML, CBR relies on experience-based knowledge stored in the form of a case consisting of a problem description and its corresponding solution. When a new problem arises, the system searches for similar cases, evaluates their suitability, and either directly reuses the solution or adapts it to the current context. Newly created solutions are then stored as cases themselves, resulting in a continuously evolving knowledge base [7, 8]. Because of this reliance on experiential knowledge, CBR requires comparatively little training data, which makes it especially valuable in domains where only limited data sets are available. The application domain of CBR within an ECC can be broadly divided into analytics, and control systems. On the analytics side, CBR is used to investigate and interpret data generated in production or within the continuum itself. A prominent example is predictive maintenance, where sensor data streams are analyzed to detect anomalies and forecast failures before they occur, or the identification of data quality issues [34]. Since IoT data streams are often complex and extensive, but labeled failure cases are scarce [50], CBR is particularly well suited. By reusing existing knowledge, CBR enables accurate classification and diagnosis even with limited amounts of training data [52]. On the control systems side, CBR can directly support decision-making and control processes, for instance by guiding resource allocation or by adapting workflows dynamically to new circumstances.

The evaluation of CBR in distributed ECC environments requires systematic criteria. Central criteria are the quality of the solution, the speed of reasoning, and the energy efficiency of the approach. Quality refers to the appropriateness of the retrieved or adapted solutions and can be measured against the outcomes in real-world scenarios. Speed captures the computational duration required to find and adapt a suitable case, which is particularly relevant for real-time or near-real-time applications. Energy efficiency is measured by the computational resources required and can be compared with the energy footprint of centralized cloud-based analysis. Thus, key metrics for the evaluation of CBR include solution quality, computation time, and energy consumption. In an ECC, distributing CBR computations across multiple nodes provides significant advantages. Large case bases can be partitioned among different edge devices, enabling local similarity searches and thereby reducing both communication overhead and latency. Error detection and analysis can take place directly at the nodes concerned, transforming the problem into smaller subproblems that can be solved locally. This reduces not only the complexity of the individual reasoning tasks but also the network load, since only the results or relevant subsets of data need to be transmitted to other nodes or to the cloud. A further benefit is increased robustness of edge-cloud distributed CBR approaches, as case knowledge is stored and applied close to where it is most relevant.

In process control, CBR can also be applied in combination with AI planning (see Sect. 4.2.2) for the scheduling and coordination of processes [33]. By combining planning with CBR the complexity of planning tasks can be reduced. In this approach, cases serve as control knowledge or as a basis for identifying similar solutions. The efficiency gain of such a hybrid method can be measured not only through energy consumption but also by the time required to generate complete plans. Within an ECC, this synergy is particularly powerful: different nodes can host different methods, with CBR providing coarse solutions that are refined through planning algorithms executed elsewhere in the continuum. A practical example illustrates these advantages. In a predictive maintenance scenario [48], edge devices monitor machine data streams for anomalies. Using CBR, they compare the current data with known failure patterns and classify emerging fault progressions. Each machine executes this reasoning locally, while the cloud analyzes potential error cases that depend on data from multiple machines. In this way, the reasoning process is both distributed and context-specific: case knowledge is only stored and applied where it is relevant. The reduction of data transfer, together with the decomposition of complex problems into smaller local tasks, leads to improved efficiency and enables near real-time analysis. Ultimately, this demonstrates how CBR in combination with ECC infrastructures supports adaptive, energy-efficient, and robust solutions for complex industrial scenarios.

4.3 Industrial Applications

Sustainability is becoming a central concern in industrial and logistics operations, driving the need for precise and transparent methods to assess energy consumption and associated CO_2e emissions. Edge computing enables real-time analysis of high-frequency measurements, which allows for task-specific CO_2e assessments, captures short-term fluctuations that would otherwise be lost in aggregation, and reduces the volume of data transmitted to central systems. Consequently, edge-enabled monitoring improves the accuracy of environmental reporting and supports optimization in production and logistics processes [18].

Several approaches exist to determine CO_2e emissions using edge devices. Evaluations typically focus on two main aspects: energy efficiency and emissions. Other relevant criteria include latency, computational resource requirements, network load, and data availability [43]. Common indicators are energy consumption per process or task (kWh), product or process-related CO_2e footprints (g CO_2e per unit or per km), and the temporal resolution of measurements [28]. The demands of the network and storage influence both the feasibility and the operational costs. Standardization ensures comparability and traceability: production footprints follow ISO 14067 [22], while logistics applications align with ISO 14083 [24]. Both approaches combine measured energy and material flows with emission factors such as kg CO_2e per kWh or per kg of fuel.

Local, high-frequency processing at the edge provides clear advantages over cloud-only approaches. Low-frequency telemetry, often sampled in minutes or hours, can obscure dynamic events such as acceleration, braking, production peaks, or sudden leaks. Edge devices capture these events directly, enabling emissions to be more precisely assigned to products, orders, or transport segments. At the same time, the processed results can be aggregated and transmitted efficiently to central systems, reducing network load and mobile data costs while maintaining insight into operational dynamics [61].

A typical edge-based CO_2e accounting pipeline comprises four stages. First, sensor and controller data are collected locally with timestamps and relevant metadata. Second, preprocessing filters noise, compresses data streams, and computes derived variables such as instantaneous power or distance. Third, emission calculations apply standard formulas and emission factors to generate CO_2e values at event or order-level. Finally, summaries, alerts, or condensed datasets are transmitted instead of raw high-frequency streams. Figure 3 shows a logistics example, five-minute aggregated distance data produce stepwise emission curves that underestimate true CO_2e values, while second-level edge processing captures peaks and produces accurate order-level footprints. This granularity supports targeted optimization, including smoother driving and route planning.

Despite these benefits, edge-based accounting faces technical and operational constraints. High-frequency data transfer is often limited by bandwidth or telecommunication costs. Many edge devices have limited processing power and memory, requiring lightweight algorithms or hardware acceleration to maintain real-time performance. Integrating heterogeneous data sources such as OPC-UA, Modbus TCP, or CAN networks demands consistent metadata and careful engineering [60]. In addition, results depend on the local electricity mix and the defined system boundaries, which means that operational footprints must be interpreted in context or complemented with life-cycle data [56].

Evaluating edge-enabled CO_2e accounting effectively requires clearly defined system boundaries, appropriate sampling rates, and the use of recognized emission factors with documented sources. Automated preprocessing and quality checks at the edge reduce the transmission of low-value data. Reporting uncertainty and performing sensitivity analyzes,for example, concerning sampling intervals or emission factors, improve transparency and support robust decision-making. In summary, edge-based CO_2e monitoring provides a practical, accurate, and efficient solution for capturing tasks-related emissions while minimizing communication overhead.

5 Challenges and Directions for Future Research

The preceding sections outlined potential criteria and exemplary approaches for evaluating ECS. Building on insights from the task force discussions between practitioners and researchers, the following section highlights key challenges and proposes a brief agenda for future research on ECS evaluation. We focus on two major areas: (1) establishing the foundations for ECS evaluation and (2) planning and conducting evaluation activities. Table 2 summarizes the identified challenges and future research questions for the evaluation of ECS.

5.1 Foundations for the Evaluation of Edge-Cloud Systems

Usually, assessment methods and procedures originate in the research environment, and practical implementations are missing. As ECS are complex systems that may behave differently based on their individual composition and usage context, *some aspects can be simulated very well under laboratory conditions while others cannot be sufficiently predicted*. For example, the extent to which the network load for data transmission or the energy consumption of certain applications in a real scenario differs from a scenario under laboratory conditions is often unclear and difficult to estimate in advance, especially when no real-world data is available. Researchers should therefore further analyze which parameters and under what conditions results of laboratory experiments and simulations can be effectively transferred into practice.

Additional research should be dedicated to removing *barriers to entry* for ECS evaluation. As the implementation and evaluation of ECS continue to pose challenges, especially for *Small and medium enterprises* (SME), accessible ECS evaluation approaches are needed. Research should develop new ways for ECS evaluation suitable for SME and other enterprises with limited resources. One example could be workshop formats to evaluate the potentials of edge-cloud applications previous to investing large sums into prototypes. Building upon the identified evaluation areas and criteria in this article, research can

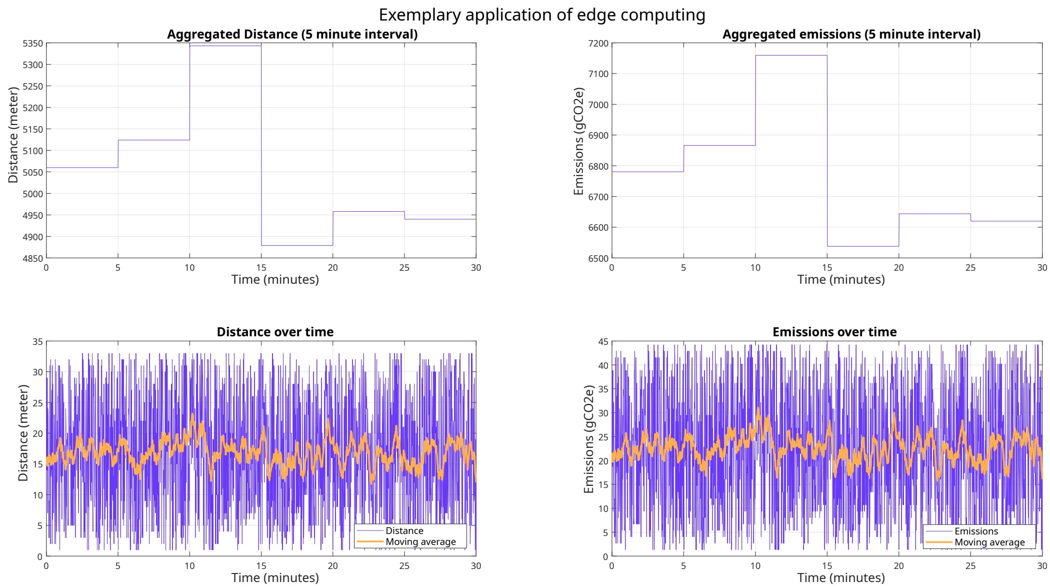

Fig. 3: Additional insights through granular data in CO_2e calculations

build a maturity model for ECS evaluation that helps organizations assess their status quo and show paths for further development of evaluation capabilities.

Furthermore, the currently *limited availability of system and resource descriptions* hampers the transparency, monitoring, and evaluation of ECS. Several factors contribute to this issue. Uniform standards for describing the performance parameters of ECS components are missing and relevant data cannot be queried by system components at runtime. Additionally, resource utilization and time-critical parameters such as latencies are not well documented, which leads to additional efforts for system integrators, who must determine these parameters themselves. Also, adopting AI technologies in combination with ECS requires trust, and consequently, AI explainability. Yet, AI results are often difficult to understand and lack transparency of the model's decision-making process. Researchers should therefore pursue the development of unified and modular approaches suitable for describing all components of ECS thereby enabling the scalable and automated monitoring of ECS. Potentially, such approaches could build upon existing research in the context of semantic technologies and digital product passports. As the instantiation of such semantic models is a highly labor-intensive step, research should investigate methods for the automation of this task, potentially drawing on AI technologies. Further research should also improve the explainability of AI in ECS. Explainable AI should not only present the correct answer but also how this answer was created, relying on decisions made locally at the edge and in the cloud infrastructure.

Finally, there is a general *lack of the needed qualifications* among systems engineers and edge computing users such as plant operators. Such persons, while responsible for driving edge-cloud projects, have typically little experience in network technologies or computer science needed for successful ECS implementation. Therefore, they often collaborate with external experts. However, barriers such as stringent firewalls or limited physical access to resources make such collaborations difficult. Therefore, there is a need for the dedicated development of skills needed for ECS deployment and evaluation for practitioners. Future research should therefore identify the skills needed for successful ECS deployment and evaluation and can subsequently develop qualification curricula for practitioners involved with setting up ECS to enable more practitioners in the future.

5.2 Planning and Execution of Edge-Cloud Systems Evaluation

While being critical factors in ECS adoption, the *assessment of "soft" criteria such as data privacy and data sovereignty is challenging*. Currently, established frameworks for evaluating these criteria are missing, which makes it difficult to objectively assess different approaches to data sovereignty and privacy in practical scenarios, lead-

Area	Current issues	Exemplary research questions
Foundations for the evaluation of ECS	Limited applicability of research results in practice	– For which parameters can laboratory results and simulations be effectively transferred into practice?
	High barriers to entry for ECS evaluation	– What are suitable low-effort methods for ECS evaluation? – How could a maturity model for ECS look like?
	Limited transparency of system descriptions	– How could a unified and modular approach for describing components of ECS look like? – What are ways to efficiently create semantic models for ECS components? – How can explainability of distributed AI in ECS be provided?
	Limited qualification of system integrators and users	– What are the skills needed to perform a holistic evaluation of ECS? – How could a curriculum for ECS evaluation look like?
Planning and execution of ECS evaluation	Complicated assessment of soft criteria like data privacy and sovereignty	– What are possible metrics and measurement approaches for data privacy and sovereignty in ECS? – How can the compliance of ECS to relevant regulatory provisions be ensured by technical means?
	Tensions between ECS evaluation complexity, cost, and performance benefits	– What are cost-efficient methods for ECS evaluation? – What criteria should be assessed as part of an efficient "minimum viable assessment"?
	Application-specific optimization targets and evaluation factors selection	– How could context-specific approaches for ECS evaluation look like? – How to design organization-specific frameworks for ECS evaluation?

Tab. 2: Challenges and research agenda for Edge-Cloud Systems (ECS) evaluation

ing to uninformed decisions and non-optimal solutions. When the conformance to related regulatory provisions needs to be assessed, organizations further rely on dedicated and costly external legal consulting. Future research should address these problems by developing usable metrics, measurement approaches and design knowledge for establishing data sovereignty and privacy in ECS. Such design knowledge may, for example, consider zero-trust architectures, which offer the opportunity to reduce security risks and improve protection against threats, especially in dynamic and scaling IT landscapes. Novel architectural approaches should be evaluated in light of emerging regulatory frameworks, such as the EU AI Act and the EU Data Act, considering the extent to which they technically ensure compliance with these legal provisions.

When planning and performing evaluations of ECS *tensions between the possible performance benefits and the efforts and costs associated with the evaluation procedures* emerge. While additional or more detailed assessments allow for further optimization of ECS, the additional costs can at some point be no longer justified. Practitioners are therefore longing for methods that allow for the most holistic and detailed assessment at the lowest costs. Researchers should develop cost-effective procedures for ECS evaluation and provide guidance on conducting cost-efficient assessments. This may include the development of a minimum viable assessment procedure, which comprises the main criteria to be assessed, potentially including the core value propositions of edge computing such as low latency and bandwidth.

Finally, and most importantly, the high heterogeneity of ECS and their applications requires *optimization targets and evaluation scenarios for ECS to be tailored to the specific context*. For example, ECS may include specialized hardware for demanding environments. Similarly, each application scenario comes with its own business requirements with regards to latency, scalability or security, among other criteria. This leads to high complexity in the evaluation process and low transferability of existing evaluation results. To address these challenges, researchers should develop context-specific frameworks and guidance for ECS evaluation. For example, typical ECS use cases could be clustered and their pertaining criteria, metrics and measurement approaches could be derived. Moreover, in large industrial enterprises with numerous branches and production sites testing various application scenarios, standardized governance approaches are needed to ensure consistent metrics, evaluation rules, and the reusability of assessment outcomes.

6 Conclusion

In this article, we presented the findings of the task force "Evaluation of Edge-Cloud Systems - Approaches, tools, and metrics" aiming to provide a holistic picture of contemporary approaches to the evaluation of ECS. We presented an overview of relevant evaluation criteria and measurements, and offered practical examples of how certain parts of ECS can be evaluated in practice. Finally, we elaborated on the current challenges and potential paths for future research in the field of ECS evaluation. We found that our contribution is valuable for both researchers and practitioners in the field of planning, implementing and assessing ECS and their components.

Our research enhances the conversation on performance benchmarking of edge and cloud computing architectures. While comparable previous approaches [5, 59] mainly draw on the existing body of literature, our approach directly builds upon the insights of practice and applied research. Therefore, we are able to extend existing findings by incorporating more contemporary criteria such as compliance with novel regulatory provisions, data sovereignty or environmental impacts along the whole system life cycle, providing a more holistic overview about relevant criteria. Furthermore, our practice-inspired research agenda for ECS evaluation seeks to foster future studies with strong practical impact.

For the practitioner community, our framework of evaluation criteria can serve as a valuable foundation for conducting ECS evaluations in a structured manner. With our framework, practitioners gain an overview of the generally relevant criteria. In a second step, they could use this overview to select the relevant criteria and metrics for their individual endeavors. Further, practitioners could use our criteria overview to create organizational guidelines for the assessment of each criterion, and to structure previous measurements within a measurements catalog, thereby enhancing the organizational reusability of assessment results.

Our study is not without limitations. It primarily draws on insights from ECS projects in Germany and Austria, and the task force did not include practitioners from sectors such as mobility, retail, or consumer electronics. Including participants from additional regions and industries could have provided further perspectives. Moreover, our study does not assess the relative importance or interrelationships of the identified criteria. Future research could explore these aspects, particularly regarding emerging criteria such as data sovereignty, environmental impacts, and novel regulatory requirements. Finally, our set of questions for open research is neither exhaustive nor complete. Further challenges for research will emerge with the increasing diffusion of ECS and the continuously changing competitive landscape. We therefore hope that our work motivates and facilitates future research in the area of ECS evaluation.

Acknowledgment: The authors gratefully acknowledge all task force participants for their commitment, insightful contributions, and feedback that substantially informed the development of this article.

Funding: This work is partially funded by the German Federal Ministry for Economic Affairs and Energy under grant No. 01MD22002C *EASY* [49], grant No. 01MD22001C *EDNA* [28], and co-funded by the Austrian Research Promotion Agency under grant No. FO999897905

Author contributions: All authors contributed equally to this work.

References

[1] Alt, B., Bast, S., Babel, M., Buitmann, J., Creutz, L., Dartmann, G., Ehaus, M., Engel, T., Gast, F., Großegesse, N., Guldner, A., Jahnke, N., Jilg, D., Körner, M.F., Leskow, J., Mundorf, J., Naumann, S., Pehlken, A., Schacht, M., Schick, L., Schlagenhauf, T., Schultheis, A., Schörner, K., Straub, S., Strüker, J., Traphöner, R., Weiher, M.A.: Evaluation of edge-cloud systems: Approaches, tools and metrics (2025), https://doi.org/10.5281/zenodo.17233607

[2] Angel, N.A., Ravindran, D., Vincent, P.M.D.R., Srinivasan, K., Hu, Y.C.: Recent advances in evolving computing paradigms: Cloud, edge, and fog technologies. Sensors (Basel, Switzerland) **22**(1) (2021), https://doi.org/10.3390/s22010196

[3] Apat, H.K., Nayak, R., Sahoo, B.: A comprehensive review on internet of things application placement in fog computing environment. Internet of Things **23**, 100866 (2023), https://doi.org/10.1016/j.iot.2023.100866

[4] Ashraf, Q.M., Tahir, M., Habaebi, M.H., Isoaho, J.: Toward autonomic internet of things: Recent advances, evaluation criteria, and future research directions. IEEE Internet of Things Journal **10**(16), 14725–14748 (2023), https://doi.org/10.1109/JIOT.2023.3285359

[5] Aslanpour, M.S., Gill, S.S., Toosi, A.N.: Performance evaluation metrics for cloud, fog and edge computing: A review, taxonomy, benchmarks and standards for future research. Internet of Things **12**, 100273 (2020), https://doi.org/10.1016/j.iot.2020.100273

[6] Babbar, H., Rani, S., Shabaz, M.: Federated learning with enhanced cryptographic security for vehicular cyber-physical systems. Scientific Reports **15**(1), 28593 (2025), https://doi.org/10.1038/s41598-025-14341-0

[7] Bergmann, R.: Experience Management: Foundations, Development Methodology, and Internet-Based Applications, LNCS, vol. 2432. Springer (2003), https://doi.org/10.1007/3-540-45759-3

[8] Bergmann, R., Althoff, K.D., Minor, M., Reichle, M., Bach, K.: Case-Based Reasoning: Introduction and Recent Developments. Künstliche Intell. **23**(1), 5–11 (2009), https://

www.researchgate.net/publication/236951398_Case-Based_Reasoning_-_Introduction_and_Recent_Developments

[9] Cajas Ordóñez, S.A., Samanta, J., Suárez-Cetrulo, A.L., Carbajo, R.S.: Intelligent edge computing and machine learning: A survey of optimization and applications. Future Internet **17**(9), 417 (2025), https://doi.org/10.3390/fi17090417

[10] Canpolat Şahin, M., Kolukısa Tarhan, A.: Evaluation and selection of hardware and ai models for edge applications: A method and a case study on uavs. Applied Sciences **15**(3), 1026 (2025), https://doi.org/10.3390/app15031026

[11] Čolaković, A.: Iot systems modeling and performance evaluation. Computer Science Review **50**, 100598 (2023), https://doi.org/10.1016/j.cosrev.2023.100598

[12] Dayan, I., Roth, H.R., Zhong, A., Harouni, A., Gentili, A., Abidin, A.Z., Liu, A., Costa, A.B., Wood, B.J., Tsai, C.S., et al.: Federated learning for predicting clinical outcomes in patients with covid-19. Nature medicine **27**(10), 1735–1743 (2021), https://doi.org/10.1038/s41591-021-01506-3

[13] Green, C.: Application of Theorem Proving to Problem Solving. In: 1st IJCAI. p. 219–240. IJCAI'69, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1969), https://dl.acm.org/doi/10.5555/1624562.1624585

[14] Grover, V., Chiang, R.H., Liang, T.P., Zhang, D.: Creating strategic business value from big data analytics: A research framework. Journal of Management Information Systems **35**(2), 388–423 (2018), https://doi.org/10.1080/07421222.2018.1451951

[15] Guo, Z., Lu, Y., Tian, H., Zuo, J., Lu, H.: A security evaluation model for multi-source heterogeneous systems based on iot and edge computing. Cluster Computing **26**(1), 303–317 (2023), https://doi.org/10.1007/s10586-021-03410-4

[16] Haslum, P.: Admissible Heuristics for Automated Planning. Ph.D. thesis, Linköping University, Sweden (2006), https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-6042

[17] Herlich, M., Pfeiffenberger, T., Du, J., Dorfinger, P.: 5G-MLab: Developing a Wireless Reliability Measurement Framework. Tech. rep., Salzburg Research (2017), https://www.salzburgresearch.at/wp-content/uploads/2017/11/Herlich-5G-MLab-Paper.pdf, Accessed: 2025-09-30

[18] Huang, J., Zhou, S., Li, G., Shen, Q.: Real-time monitoring and optimization methods for user-side energy management based on edge computing. Scientific reports **15**(1), 24890 (2025), https://doi.org/10.1038/s41598-025-07592-4

[19] Huang, J., Liu, F., Zhang, J.: Multi-dimensional qos evaluation and optimization of mobile edge computing for iot: A survey. Chinese Journal of Electronics **33**(4), 859–874 (2024), https://doi.org/10.23919/cje.2023.00.264

[20] Iftikhar, S., Gill, S.S., Song, C., Xu, M., Aslanpour, M.S., Toosi, A.N., Du, J., Wu, H., Ghosh, S., Chowdhury, D., Golec, M., Kumar, M., Abdelmoniem, A.M., Cuadrado, F., Varghese, B., Rana, O., Dustdar, S., Uhlig, S.: Ai-based fog and edge computing: A systematic review, taxonomy and future directions. Internet of Things **21**, 100674 (2023), https://doi.org/10.1016/j.iot.2022.100674

[21] International Organization for Standardization (ISO): Iso/iec 2382:2015: Information technology — vocabulary (2015), https://www.iso.org/standard/63598.html, accessed: 2025-11-03

[22] International Organization for Standardization (ISO): Iso 14067:2018: Greenhouse gases — carbon footprint of products — requirements and guidelines for quantification (2018), https://www.iso.org/standard/71206.html, accessed: 2025-11-03

[23] International Organization for Standardization (ISO): Iso/iec tr 23188:2020: Information technology — cloud computing — edge computing landscape (2020), https://www.iso.org/standard/74846.html, accessed: 2025-11-03

[24] International Organization for Standardization (ISO): Iso 14083:2023: Greenhouse gases — quantification and reporting of greenhouse gas emissions arising from transport chain operations (2023), https://www.iso.org/standard/78864.html, accessed: 2025-11-03

[25] Jin, S., Ou, Q., Feng, Y., Zhang, N., Cong, L., Wang, Y., Yu, P.: An efficiency evaluation method for cloud-edge collaborative network. In: 2021 International Wireless Communications and Mobile Computing (IWCMC). pp. 51–56. IEEE (2021), https://doi.org/10.1109/IWCMC51323.2021.9498744

[26] Kong, X., Wu, Y., Wang, H., Xia, F.: Edge computing for internet of everything: A survey. IEEE Internet of Things Journal **9**(23), 23472–23485 (2022), https://doi.org/10.1109/JIOT.2022.3200431

[27] Koubaa, A., Ammar, A., Kanhouch, A., AlHabashi, Y.: Cloud versus edge deployment strategies of real-time face recognition inference. IEEE Transactions on Network Science and Engineering **9**(1), 143–160 (2022), https://doi.org/10.1109/TNSE.2021.3055835

[28] Leskow, J., Greiser, S., Gregorc, L., Damalas, S.: Monte carlo simulation to analyse the impact of uncertain telemetry data on the calculation of co2e emissions for trailer traffic. Wirtschaftsinformatik 2024 Proceedings (2024), https://aisel.aisnet.org/wi2024/35

[29] LF Edge: State of the edge (2024), https://stateoftheedge.com/reports/state-of-the-edge-report-2024/, accessed: 2025-11-03

[30] Ma, X., Li, X.: Trust evaluation model in edge computing based on integrated trust. In: Proceedings of the 2018 International Conference on Algorithms, Computing and Artificial Intelligence. pp. 1–6. ACM, New York, NY, USA (2018), https://doi.org/10.1145/3302425.3302491

[31] Maheshwari, S., Raychaudhuri, D., Seskar, I., Bronzino, F.: Scalability and performance evaluation of edge cloud systems for latency constrained applications. In: 2018 IEEE/ACM Symposium on Edge Computing (SEC). pp. 286–299 (2018), https://doi.org/10.1109/SEC.2018.00028

[32] Maier, C., Großegesse, N., Strohmeier, F.: Predicting performance metrics in edge-cloud networks using graph neural networks. In: Proceedings of the 3rd Workshop on Machine Learning in Networking (MaLeNe), co-located with the 6th International Conference on Networked Systems (NetSys 2025) (2025), https://opus.bibliothek.uni-augsburg.de/opus4/frontdoor/index/index/docId/125254

[33] Malburg, L., Brand, F., Bergmann, R.: Adaptive Management of Cyber-Physical Workflows by Means of Case-Based Reasoning and Automated Planning. In: 26th EDOC Workshops. LNBIP, vol. 466, pp. 79–95. Springer (2023), https://doi.org/10.1007/978-3-031-26886-1_5

[34] Malburg, L., Schultheis, A., Bergmann, R.: Modeling and Using Complex IoT Time Series Data in Case-Based Reasoning: From Application Scenarios to Implementations. In:

31st ICCBR Workshop Proc. vol. 3438, pp. 81–96 (2023), https://ceur-ws.org/Vol-3438/paper_07.pdf

[35] Marrella, A.: What Automated Planning Can Do for Business Process Management. In: BPM 2017 Workshops. LNBIP, vol. 308, pp. 7–19. Springer (2017), https://doi.org/10.1007/978-3-319-74030-0_1

[36] McMahan, B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: Artificial intelligence and statistics. pp. 1273–1282. PMLR (2017), http://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf

[37] Mechalikh, C., Taktak, H., Moussa, F.: Pureedgesim: A simulation framework for performance evaluation of cloud, edge and mist computing environments. Computer Science and Information Systems **18**(1), 43–66 (2021), https://doi.org/10.2298/CSIS200301042M

[38] Mertens, D.J.M., Haskamp, T., vom Brocke, J.: Exploring digital sustainability: A multidimensional framework for is scholarship. In: Proceedings of the 2025 International Conference on Information Systems (ICIS 2025). Digital Innovation and Entrepreneurship, vol. 4 (2025), https://aisel.aisnet.org/icis2025/diginnoventren/diginnoventren/4/

[39] Mijuskovic, A., Chiumento, A., Bemthuis, R., Aldea, A., Havinga, P.: Resource management techniques for cloud/fog and edge computing: An evaluation framework and classification. Sensors (Basel, Switzerland) **21**(5) (2021), https://doi.org/10.3390/s21051832

[40] Misra, N.N., Dixit, Y., Al-Mallahi, A., Bhullar, M.S., Upadhyay, R., Martynenko, A.: Iot, big data, and artificial intelligence in agriculture and food industry. IEEE Internet of Things Journal **9**(9), 6305–6324 (2022), https://doi.org/10.1109/JIOT.2020.2998584

[41] Nandhakumar, A.R., Baranwal, A., Choudhary, P., Golec, M., Gill, S.S.: Edgeaisim: A toolkit for simulation and modelling of ai models in edge computing environments. Measurement: Sensors **31**, 100939 (2024), https://doi.org/10.1016/j.measen.2023.100939

[42] Noghabi, S.A., Cox, L., Agarawal, S., Ananthanarayanan, G.: The emerging landscape of edge computing. GetMobile: Mobile Computing and Communications **23**, 11–20 (2019), https://doi.org/10.1145/3400713.3400717

[43] Porambage, P., Okwuibe, J., Liyanage, M., Ylianttila, M., Taleb, T.: Survey on multi-access edge computing for internet of things realization. IEEE Communications Surveys & Tutorials **20**(4), 2961–2991 (2018), https://doi.org/10.1109/COMST.2018.2849509

[44] Rajesh, M., Ramachandran, S., Vengatesan, K., Dhanabalan, S.S., Nataraj, S.K.: Federated learning for personalized recommendation in securing power traces in smart grid systems. IEEE Transactions on Consumer Electronics **70**(1), 88–95 (2024), https://doi.org/10.1109/TCE.2024.3368087

[45] Rosendo, D., Costan, A., Valduriez, P., Antoniu, G.: Distributed intelligence on the edge-to-cloud continuum: A systematic literature review. Journal of Parallel and Distributed Computing **166**, 71–94 (2022), https://doi.org/10.1016/j.jpdc.2022.04.004

[46] Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: The Graph Neural Network Model. IEEE Transactions on Neural Networks **20**(1), 61–80 (2009), https://doi.org/10.1109/TNN.2008.2005605

[47] von Scherenberg, F., Hellmeier, M., Otto, B.: Data sovereignty in information systems. Electron. Mark. **34**(1) (2024), https://doi.org/10.1007/s12525-024-00693-4

[48] Schultheis, A.: Exploring a Hybrid Case-Based Reasoning Approach for Time Series Adaptation in Predictive Maintenance. In: 32nd ICCBR Workshops. vol. 3708, pp. 230–235. CEUR-WS.org (2024), https://ceur-ws.org/Vol-3708/paper_19.pdf

[49] Schultheis, A., Alt, B., Bast, S., Guldner, A., Jilg, D., Katic, D., Mundorf, J., Schlagenhauf, T., Weber, S., Bergmann, R., Bergweiler, S., Creutz, L., Dartmann, G., Malburg, L., Naumann, S., Rezapour, M., Ruskowski, M.: EASY: Energy-Efficient Analysis and Control Processes in the Dynamic Edge-Cloud Continuum for Industrial Manufacturing. Künstliche Intelligenz (2024), https://doi.org/10.1007/s13218-024-00868-3

[50] Schultheis, A., Bertrand, Y., Grüger, J., Malburg, L., Bergmann, R., Serral Asensio, E.: Case-based Data Quality Management for IoT Logs: A Case Study Focusing on Detection of Data Quality Issues. IoT **6**(4) (2025), https://doi.org/10.3390/iot6040063

[51] Schultheis, A., Jilg, D., Malburg, L., Bergweiler, S., Bergmann, R.: Towards Flexible Control of Production Processes: A Requirements Analysis for Adaptive Workflow Management and Evaluation of Suitable Process Modeling Languages. Processes **12**(12) (2024), https://doi.org/10.3390/pr12122714

[52] Seiger, R., Schultheis, A., Bergmann, R.: Case-Based Activity Detection from Segmented Internet of Things Data. In: 33rd ICCBR Proc. LNCS, vol. 15662, pp. 438–453. Springer (2025), https://doi.org/10.1007/978-3-031-96559-3_29

[53] Sonmez, C., Ozgovde, A., Ersoy, C.: Edgecloudsim: An environment for performance evaluation of edge computing systems. Transactions on Emerging Telecommunications Technologies **29**, e3493 (2018), https://doi.org/10.1002/ett.3493

[54] Strnadl, C.F.: End-to-end-architekturen zur datenmonetarisierung im industrial internet of things (iiot). In: Trauth, D., Bergs, T., Prinz, W. (eds.) Monetarisierung von technischen Daten, pp. 169–206. Springer Berlin Heidelberg, Berlin, Heidelberg (2021), https://doi.org/10.1007/978-3-662-62915-4{_}10

[55] Suárez-Varela, J., Almasan, P., Ferriol-Galmés, M., Rusek, K., Geyer, F., Cheng, X., Shi, X., Xiao, S., Scarselli, F., Cabellos-Aparicio, A., Barlet-Ros, P.: Graph Neural Networks for Communication Networks: Context, Use Cases and Opportunities. IEEE Network **37**(3), 146–153 (2023), https://doi.org/10.1109/MNET.123.2100773

[56] Trihinas, D., Thamsen, L., Beilharz, J., Symeonides, M.: Towards energy consumption and carbon footprint testing for ai-driven iot services. In: Taheri, J. (ed.) 2022 IEEE International Conference on Cloud Engineering. pp. 29–35. IEEE, Piscataway, NJ (2022), https://doi.org/10.1109/IC2E55432.2022.00011

[57] Vailshery, L.S.: Number of internet of things (iot) connections worldwide from 2022 to 2023, with forecasts from 2024 to 2034 (Jun 2025), https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/?srsltid=

AfmBOoqjDlO3TZVipLmgpOdEfVOIgywUvDI9axUJMI-3kZmV0kGlfIQT
[58] Varghese, B., de Lara, E., Ding, A.Y., Hong, C.H., Bonomi, F., Dustdar, S., Harvey, P., Hewkin, P., Shi, W., Thiele, M., Willis, P.: Revisiting the arguments for edge computing research. IEEE Internet Computing **25**(5), 36–42 (2021), https://doi.org/10.1109/MIC.2021.3093924
[59] Varghese, B., Wang, N., Bermbach, D., Hong, C.H., Lara, E.D., Shi, W., Stewart, C.: A survey on edge performance benchmarking. ACM Comput. Surv. **54**(3) (2021), https://doi.org/10.1145/3444692
[60] Verde Romero, D.A., Villalvazo Laureano, E., Jiménez Betancourt, R.O., Navarro Álvarez, E.: An open source iot edge-computing system for monitoring energy consumption in buildings. Results in Engineering **21**, 101875 (2024), https://doi.org/10.1016/j.rineng.2024.101875
[61] Wang, F., Zhang, M., Wang, X., Ma, X., Liu, J.: Deep learning for edge computing applications: A state-of-the-art survey. IEEE Access **8**, 58322–58336 (2020), https://doi.org/10.1109/ACCESS.2020.2982411
[62] Willner, A., Gowtham, V.: Toward a reference architecture model for industrial edge computing. IEEE Communications Standards Magazine **4**(4), 42–48 (2020), https://doi.org/10.1109/MCOMSTD.001.2000007
[63] Yang, C., Lan, S., Wang, L., Shen, W., Huang, G.G.Q.: Big data driven edge-cloud collaboration architecture for cloud manufacturing: A software defined perspective. IEEE Access **8**, 45938–45950 (2020), https://doi.org/10.1109/ACCESS.2020.2977846
[64] Yang, H., Guo, L., Wang, T., Lv, C.: A novel lightweight dynamic trust evaluation model for edge computing. IEEE Transactions on Network and Service Management **22**(4), 3542–3554 (2025), https://doi.org/10.1109/TNSM.2025.3570290
[65] Zhang, T., Li, Y., Wang, M., et al.: Time-sensitive networking (tsn) for industrial automation: Current practice and future trends. ACM Computing Surveys **57**(2), 30:1–30:33 (2024), https://doi.org/10.1145/3695248