**Original Research**

# A Hybrid Edge-Cloud Framework for Efficient Big Data Processing in IoT Environments

Hande Erkoç[1] and Ali Batan[2]

[1]Uşak Institute of Technology, Department of Computer Science, İstiklal Mah. 75. Cadde No:16, Uşak, Turkey.
[2]Bolu Technical University, Department of Computer Engineering, D100 Karayolu, Bolu, Turkey.

**Abstract**
The rapid proliferation of Internet of Things networks generates voluminous and heterogeneous data, requiring robust mechanisms for real-time processing and analysis. A hybrid edge-cloud framework for efficient big data management emerges as a promising architecture to support latency-sensitive services and resource-intensive computations. This framework strategically leverages edge servers to handle immediate and localized tasks, thereby alleviating the burden on centralized cloud infrastructures. By processing large data streams at or near the source of their generation, network bottlenecks can be minimized, and faster response times become possible. Meanwhile, the cloud remains essential for complex analytical workloads and high-capacity data storage. These dual-layer interactions enable adaptive workload distribution, facilitate dynamic scaling, and can be guided by advanced resource allocation principles. This approach effectively addresses the demand for low-latency processing in applications such as autonomous vehicles, connected healthcare, and industrial automation. This work explores the design of a sophisticated hybrid architecture that orchestrates data traffic and computation flows among distributed edge nodes and centralized cloud data centers. The proposed framework aims to improve communication overhead, computation accuracy, and real-time responsiveness, even in resource-constrained devices. Limitations, such as varying network conditions and potential security vulnerabilities, are also examined to highlight the challenges of deploying this approach in real-world scenarios.

## 1. Introduction

The explosive growth of distributed sensor networks, connected devices, and autonomous systems has necessitated a more nuanced approach to data processing and analytics. As new services and applications emerge, the ability to rapidly capture, filter, and act upon real-time data streams becomes ever more critical. Traditional cloud-centric solutions, while powerful in terms of computational capability, can incur substantial latency due to the distance between connected devices and remote data centers [1]. An extended round-trip network delay may be acceptable for batch processing or analytical tasks without strict timing requirements, but it becomes prohibitive for many modern Internet of Things applications. Consequently, the paradigm of pushing selected computational tasks closer to the data source, referred to as edge computing, has gained significant traction. [2]

Edge computing enables real-time decision-making by exploiting localized computing resources situated close to endpoints or within gateways. This approach can reduce data movement overhead and decrease response times for tasks requiring immediate intervention. However, edge devices often have constrained processing power, memory, and storage capacity [3]. These limitations present new challenges for effectively coordinating operations among numerous heterogeneous nodes. Moreover, there is a delicate balance between the volume of data processed at the edge and the portion sent to the cloud for further analysis [4]. When these localized computations must handle large, unstructured data streams, resource constraints become even more pronounced, creating potential inefficiencies or bottlenecks.

The advantages of a hybrid edge-cloud framework lie in the complementary capabilities of each layer. The edge layer excels at immediate data processing and filtration, where localized analytics and swift responses are necessary [5]. The cloud layer retains a pivotal role for tasks involving advanced machine learning algorithms, large-scale data aggregation, and long-term storage. By amalgamating both layers, one can exploit distributed and localized intelligence while still benefiting from the robustness of cloud infrastructures [6]. This synergy enables dynamic workload partitioning and migration, where tasks can be shifted between edge nodes and the cloud in real time, depending on resource availability, network conditions, and priority levels. The management of these tasks, however, introduces immense complexity, as orchestrating numerous devices with varying capacities and real-time requirements demands sophisticated coordination [7].

Recent hardware and networking innovations further propel the feasibility of hybrid architectures [8]. Low-power processors, accelerators, and high-speed networking interfaces enable edge nodes to serve more complex roles than ever before. Simultaneously, the evolution of containerization and virtualization technologies facilitates scalable deployments across cloud and edge environments [9]. Yet, even with these benefits, unresolved issues remain in terms of interoperability, data consistency, and failover strategies. Organizations that require strict compliance or face data sovereignty constraints must carefully evaluate where and how data is processed. [10]

Another important consideration is energy efficiency. Many IoT devices and edge nodes rely on batteries or intermittent power sources. Prolonged operation requires careful attention to energy consumption, both in computation and data communication [11, 12]. Automated decisions regarding task offloading, bandwidth allocation, and storage must account not only for performance metrics like latency but also for energy budgets. This introduces an optimization challenge that spans multiple layers, combining computational speed, memory usage, data transfer overhead, and battery constraints. [13]

Alongside these technical considerations, there is growing attention to security and privacy. As data flows between edge nodes and the cloud, it becomes vulnerable to interception or tampering unless robust security mechanisms are in place. Moreover, data processed at the edge may be of a highly localized or personalized nature, raising questions about adherence to various privacy regulations [14]. A well-designed hybrid architecture must, therefore, integrate cryptographic protocols and access control policies that address these security requirements without imposing significant latency overhead. The dynamic and distributed nature of edge computing underscores the need to continuously reevaluate security postures and patch vulnerabilities in a timely manner. [15, 16]

The increasing ubiquity of IoT endpoints and the diversity of data they generate forces us to rethink every aspect of the data processing pipeline. The ability of a hybrid edge-cloud system to capture, process, and distribute data in near real-time can bolster numerous application domains, including industrial automation, smart cities, healthcare monitoring, and autonomous driving. Despite its promise, the development and deployment of such frameworks must be approached with carefully planned strategies to cope with scalability, reliability, and manageability [17]. The ensuing sections delve deeper into the conceptual underpinnings and mathematical constructs that guide the orchestration of a hybrid edge-cloud environment for big data processing, followed by an implementation-driven performance evaluation. Limitations will also be examined to highlight realistic outcomes and potential areas where future improvements are warranted. [18]

## 2. Hybrid Edge-Cloud Architecture Overview

The conceptual design of a hybrid edge-cloud architecture for handling big data workloads in IoT environments relies on multiple interlinked layers that ensure reliability, scalability, and flexibility. At the most fundamental level, IoT sensors and devices generate continuous data streams, often with varying degrees of temporal and spatial correlation [19]. These data streams must be captured and processed in a manner that can accommodate the scale and velocity of the incoming information. The edge tier manages initial data aggregation, filtering, and real-time analytics using device-level or gateway-level resources. Depending on the available computational capacity and workload requirements, a subset of the data is relayed to the cloud tier for extended storage and more complex analysis. [20, 21]

Central to this architecture is the orchestrator, which can be conceptualized as a collection of software services running within edge nodes or cloud instances. The orchestrator monitors real-time metrics, such as CPU utilization, memory availability, network bandwidth, and latency constraints, to determine the optimal distribution of tasks [22]. When the volume of incoming data surges, the orchestrator may decide to provision additional edge nodes, if they are available, or direct more of the workload to the cloud if edge resources are overburdened. Conversely, if network conditions degrade, the orchestrator may prefer local data processing on the edge to minimize the delay caused by back-and-forth data transfers.

To ensure high availability, each edge node can be designed with multiple redundant connections or local caching mechanisms [23]. This redundancy aims to minimize single points of failure, an essential requirement for mission-critical applications where data loss or processing delays could lead to catastrophic results. A practical concern is the dynamic nature of IoT deployments [24]. Devices may periodically join or leave the network, or they may relocate physically. This dynamic behavior forces the orchestrator to use adaptive algorithms that recompute optimal task placements as system conditions evolve. Another important aspect is workload heterogeneity [25]. IoT data may consist of time-series measurements, images, video streams, or discrete event logs. Each data format demands a specialized set of preprocessing, feature extraction, and analytics functions, potentially requiring GPU-accelerated computing for tasks like computer vision. [26]

However, the adoption of a hybrid edge-cloud system poses several technical hurdles. First, the hardware profile of edge devices can vary widely, necessitating a common abstraction layer or containerization technology to ensure application portability [27]. Second, maintaining consistent data semantics across decentralized nodes can be difficult if local versions of data are processed or updated independently. Finally, security policies must be uniformly enforced without introducing bottlenecks or single points of compromise. These issues become more acute as the scale of deployment increases, particularly in industrial or urban environments where thousands or millions of devices must be managed simultaneously. [28, 29]

Resource allocation methods within this architecture often rely on intelligent partitioning of tasks. Different classes of tasks may have different latency demands, computational complexities, and degrees of parallelizability [30]. High-priority tasks that require immediate responses should be processed at the edge, whereas computationally heavy tasks might be offloaded to the cloud. The orchestrator sets thresholds for the maximum latency permitted for certain operations, thus guiding the choice of local versus remote processing. Over time, historical usage data can be harnessed to predict likely workload spikes, enabling proactive resource provisioning. [31, 32]

Another significant component is the networking layer, where edge nodes connect to the cloud through gateways or local networks. Traditional wide-area communication protocols might be too slow for latency-sensitive applications, encouraging the usage of dedicated channels or 5G-based infrastructures [33]. The capacity and reliability of these connections have a direct impact on the feasibility of migrating tasks to the cloud. One may also employ a distributed file system or object storage layer that spans edge and cloud, streamlining data sharing while maintaining consistency.

In many practical implementations, a central database or metadata repository in the cloud maintains the global state of the system, such as device configurations, credentials, and software versions [34]. Edge nodes periodically synchronize with this repository to ensure consistent configurations across the entire environment. Local states, such as intermediate processing results or short-term caches, remain confined to the edge unless flagged as relevant for long-term analysis or archiving [35]. This dual-layer storage model can mitigate bandwidth usage and reduce response times while preserving the ability to conduct aggregate analyses at the cloud level.

Despite the advanced capabilities of hybrid architectures, the associated overhead and complexity can impact performance. If many tasks are concurrently offloaded to the cloud, network congestion or cloud-side resource constraints can degrade quality of service. Conversely, overloading edge nodes with computationally heavy tasks introduces local bottlenecks and energy constraints. Therefore, dynamic balancing remains a cornerstone of efficient operation [36]. These factors underscore the importance of a framework that not only splits tasks but also estimates resource usage, anticipates network conditions, and applies predictive modeling to optimize throughput.

In the broader context, the transition to hybrid edge-cloud systems resonates with the need to ensure reliability and efficiency in data-driven decision-making processes [37, 38]. From real-time anomaly detection in industrial systems to real-time image analytics in public safety applications, the demand for instantaneous feedback has reshaped the computing landscape. While the technical underpinnings remain complex, emerging solutions focusing on orchestration, resource discovery, and adaptive allocation suggest a clear path forward. The next section dives into the underlying mathematical principles that guide these resource allocation and data processing decisions, offering a rigorous way to handle the intrinsic complexity of hybrid edge-cloud ecosystems. [39]

## 3. Mathematical Modeling of Data Processing Workflows

The optimal distribution of computational tasks between the edge and the cloud can be formulated as a constrained optimization problem, in which the system seeks to minimize total latency, maximize throughput, or balance multiple objectives subject to hardware, network, and energy constraints. Let the system have a set of tasks, denoted by $\{T_1, T_2, \ldots, T_n\}$, each requiring a certain amount of computational resources and exhibiting a time sensitivity characterized by a latency deadline $D_i$. The task latency $L_i$ can be decomposed into processing time $P_i$ and transmission time $X_i$, yielding [40]

$$L_i = P_i + X_i.$$

If a task is assigned to an edge node $j$, the processing time and local data transmission time might be smaller due to physical proximity, but the available computational power on that edge node is limited. Conversely, offloading a task to the cloud might provide more abundant computing resources but result in higher network latency [41]. We introduce a binary decision variable $a_{i,j}$ that equals 1 if task $T_i$ is assigned to node $j$ and 0 otherwise. The decision problem is

$$\min_{a_{i,j}} \sum_{i=1}^{n} L_i$$

subject to [42]

$$\sum_{j=1}^{m} a_{i,j} = 1, \quad \forall i,$$

$$L_i = f(P_i, X_i, a_{i,j}, r_j, b_j),$$

$$L_i \leq D_i, \quad \forall i,$$

$$0 \leq a_{i,j} \leq 1, \quad \forall i, j,$$

where $m$ denotes the total number of nodes (edge + cloud), $r_j$ is the processing rate of node $j$, and $b_j$ is the available bandwidth linking node $j$ to the system. The function $f$ captures the combined effects of computational load and network transfer. The constraint $\sum_{j=1}^{m} a_{i,j} = 1$ ensures that each task is assigned to exactly one node, while $L_i \leq D_i$ enforces the deadline requirement. The system also needs to respect the capacity limits of each node, which can be expressed as

$$\sum_{i=1}^{n} a_{i,j} \cdot \omega_i \leq c_j, \quad \forall j,$$

where $\omega_i$ represents the resource demand of task $T_i$, and $c_j$ denotes the capacity of node $j$ [43]. If an edge node has limited energy reserves, an additional constraint can be introduced:

$$\sum_{i=1}^{n} a_{i,j} \cdot e_i \leq E_j,$$

where $e_i$ denotes the energy consumption of task $T_i$, and $E_j$ is the total available energy at node $j$. [44]

This basic framework can be extended to incorporate multi-stage workflows, where intermediate outputs of one task become inputs for another. For instance, let each job be represented by a directed acyclic graph (DAG), with vertices denoting sub-tasks and edges specifying data dependencies. The overall latency is dictated by the critical path of this DAG [45]. If the edges in the DAG have associated data sizes, we must account for intermediate data transfer times:

$$L_{\text{job}} = \max_{\mathcal{P} \in \text{Paths of the DAG}} \sum_{k \in \mathcal{P}} P_k + \sum_{(k,l) \in \mathcal{P}} X_{k,l},$$

where $P_k$ is the processing time of sub-task $k$ and $X_{k,l}$ is the transfer time of intermediate data between tasks $k$ and $l$. Minimizing $L_{\text{job}}$ for a set of jobs can be more complex than the independent-task scenario, often necessitating integer linear programming or heuristic methods to find near-optimal solutions. Lagrangian relaxation techniques or primal-dual formulations may also be used to handle the capacity constraints and synergy among tasks. [46]

Another important dimension is the stochastic nature of network conditions and computational demands. Network bandwidth can vary over time due to background traffic or environmental conditions in wireless scenarios. Similarly, the computational load on each node can change as tasks arrive or complete [47]. Incorporating uncertainty leads to stochastic optimization models, where a common approach is to define probability distributions for bandwidth $\hat{b}_j$ and available CPU cycles $\hat{r}_j$. One can then define the expected latency for a task assigned to node $j$ as

$$\mathbb{E}[L_i] = \mathbb{E}\big[P_i(\hat{r}_j) + X_i(\hat{b}_j)\big].$$

The optimization problem becomes [48]

$$\min_{a_{i,j}} \sum_{i=1}^{n} \mathbb{E}[L_i],$$

with additional constraints that guarantee an acceptable level of performance under certain percentile conditions. For example, one might require that the probability of meeting the deadline is above a threshold $\alpha$: [49]

$$\Pr(L_i \leq D_i) \geq \alpha, \quad \forall i.$$

This leads to chance-constrained optimization formulations, which can be difficult to solve exactly but may be approached with sample-based or scenario-based approximation techniques.

To reduce the complexity, one could use hierarchical optimization, splitting the problem into two layers [50, 51]. The first layer determines how many tasks should generally be processed by the edge versus the cloud. The second layer focuses on allocating tasks among specific nodes of a given layer [52]. The synergy between these two layers can be guided by a partial Lagrangian decomposition, in which the constraints linking layers are relaxed, enabling parallel solution procedures. Alternatively, reinforcement learning approaches can be employed, where an agent dynamically learns to assign tasks based on observed system states, aiming to minimize a cumulative cost metric. The policy updates over time, refining the allocation strategies as real-world data and network states are observed. [53]

In practice, such advanced models typically require fast heuristics because of the time-critical nature of many IoT systems. Approximation algorithms, greedy schemes, or meta-heuristics like simulated annealing are frequently used [54]. These approaches offer near-optimal solutions in a fraction of the time that exact methods would require. The choice of a specific algorithm depends on system constraints, task arrival patterns, and performance requirements. For example, a scheduling heuristic may prioritize tasks with the earliest deadline first, breaking ties by also considering the current load on each node [55]. Another variation might focus on maximizing throughput for a batch of tasks, ignoring individual deadlines as a trade-off for higher overall system performance.

The theoretical models described here provide a foundation for understanding the core optimization issues in hybrid edge-cloud data processing [56]. They serve as a basis for controlling the interactions among distributed nodes, ensuring that tasks are executed in a timely manner without surpassing computational or energy capacities. In the following section, the focus shifts to a more tangible perspective, describing an actual implementation of a hybrid edge-cloud framework and empirical performance evaluations [57]. These results highlight the efficacy of the proposed strategies and also reveal certain areas where further refinements are required, especially under extreme network or resource constraints.

## 4. Implementation and Performance Evaluation

The prototype implementation of a hybrid edge-cloud framework for big data processing involves configuring a set of heterogeneous edge nodes that connect to a central cloud service. Edge nodes are outfitted with moderate-capacity CPUs and optional hardware accelerators, such as GPUs or specialized chips, depending on the target application domain [58]. The cloud environment, hosted within a large data center, consists of multiple high-performance servers interconnected by a high-bandwidth internal fabric. Containers or virtual machines are deployed on both edge and cloud environments to facilitate the seamless migration of microservices in response to fluctuating workloads. [59]

In this particular prototype, a dedicated orchestrator service runs in the cloud but maintains local monitoring agents at each edge node. These agents collect real-time metrics, including CPU utilization, memory usage, available bandwidth, and battery level where applicable. The orchestrator aggregates this information and applies the resource allocation strategies detailed in the preceding mathematical formulation [60]. Tasks arrive in a streaming manner from various IoT devices, such as temperature sensors, cameras, and industrial machinery logs. Certain tasks involve real-time anomaly detection, whereas others require periodic aggregation and analytics that are less time-sensitive. [61]

The actual data flows use a combination of message broker protocols and REST-based microservices, ensuring that tasks can be distributed with minimal overhead. To reduce network latency, an application-level gateway at each edge node preprocesses raw data, compresses it, or discards irrelevant portions before transmission to the cloud. For compute-intensive analytics, the orchestrator measures the predicted speed of performing the computation at the edge against the overhead of sending the data to the cloud [62]. Depending on the results, it forwards only the model parameters or partial results if local inference is feasible. Conversely, tasks that exceed the local capacity threshold are offloaded to the cloud [63]. The entire workflow is orchestrated by high-level schedules that adapt to real-time conditions every few seconds or minutes, depending on the criticality of the application.

Performance testing was conducted in a controlled lab environment to simulate various conditions, such as congested network links, sudden spikes in task arrivals, and reduced CPU availability [64]. When the network exhibited stable conditions and moderate traffic, the hybrid strategy demonstrated a significant improvement in latency-sensitive task completion times compared to a purely cloud-based approach. The test scenarios showed that edge-based processing could reduce average response times by up to 40 percent for those tasks requiring immediate attention. However, as the local load increased on the edge nodes, the orchestrator naturally pushed more workloads to the cloud, resulting in slightly elevated latencies but maintaining an overall system responsiveness. [65, 66]

Another experiment introduced an artificial network bottleneck between the edge and the cloud, emulating a scenario in which cellular coverage was poor or bandwidth was throttled. This stressed the importance of local processing [67]. Tasks that were deemed urgent continued to be processed locally, while tasks with less stringent latency requirements were queued for eventual cloud processing.

Although the system sustained degraded performance, it was still able to handle essential workloads without major interruptions. Data throughput metrics indicated that local buffering mechanisms and partial offloading policies played a crucial role in mitigating service disruptions. [68]

The framework was also examined under a heavy workload scenario to investigate how it behaves when the cumulative computational demand exceeds the combined capacity of all the edge nodes. In this case, many tasks waited in queues at the cloud data center, resulting in increased average completion times [69]. Nonetheless, the system maintained a consistent strategy of executing short, real-time tasks at the edge, ensuring that high-priority operations continued to meet their deadlines. This result confirms the advantage of maintaining an integrated environment that can flexibly use both local and remote resources without completely overloading either.

Energy consumption measurements were collected at the edge nodes, focusing on the trade-off between computation and network transmission [70]. In scenarios where data was large but tasks were simple, local processing was often more efficient because transmitting raw data to the cloud would have consumed considerable power. Conversely, when tasks demanded complex neural network inference, offloading them to the cloud saved energy, especially if the edge device lacked hardware acceleration [71]. The orchestrator's ability to factor in device battery states was found to be vital in preserving edge node longevity.

While these results underscore the potential of hybrid architectures, limitations were observed in instances of abrupt workload transitions and extreme resource constraints [72]. The orchestrator's scheduling logic occasionally lagged behind rapid changes in system load, particularly when multiple edge nodes simultaneously reported high utilization. This discrepancy introduced temporary inefficiencies until the next scheduling interval. Future implementations could incorporate more frequent updates or predictive models that anticipate workload changes and proactively scale resources. [73, 74]

Despite careful planning, certain security challenges surfaced during stress tests. As data moved between edge nodes and the cloud, ensuring robust encryption at every stage was nontrivial [75]. Moreover, the distributed nature of the environment complicated the patching process for vulnerabilities. Each node needed to run software upgrades in synchrony to avoid version mismatches that could lead to disrupted communication. These findings suggest that practical deployments must incorporate automated software provisioning pipelines capable of securely updating both cloud and edge tiers. [76, 77]

In a real-world smart city application, scaling beyond laboratory settings will require handling thousands of devices and multiple data types simultaneously. Meeting this challenge demands a broader set of strategies, potentially including multi-level caching, more advanced container orchestration platforms, and robust failure recovery mechanisms [78]. Emphasis on debugging and monitoring tools is equally crucial to identify performance bottlenecks and security gaps, given the complexity of distributed microservices that run on both ends.

Collectively, these performance evaluations demonstrate that a hybrid edge-cloud framework can handle diverse workloads efficiently if properly orchestrated. The synergy between local data processing and centralized analytics offers tangible benefits in latency reduction, bandwidth conservation, and energy optimization [79]. The next section concludes the discussion by summarizing the insights gained and suggesting paths for future improvements, including advanced methods for scheduling under uncertainty, tighter security provisions, and more robust approaches to handling large-scale distributed data sets.

## 5. Conclusion

A hybrid edge-cloud framework constitutes a powerful solution for the challenges inherent in processing large-scale, heterogeneous data flows within IoT environments [80]. The synergy between distributed edge intelligence and centralized cloud analytics provides a balanced approach that can accommodate the low-latency demands of many real-time applications, while still leveraging the computational depth available at large data centers. The architectures, algorithms, and performance results presented here illustrate how the orchestrated allocation of resources can reduce data transfer overhead and improve the efficiency of mission-critical tasks. [81]

Fundamental to this approach is the realization that different layers excel at distinct types of work. The edge nodes are adept at fast local computations, capable of immediate responses to critical events, while the cloud remains indispensable for complex analytics, archival storage, and large-scale model training. Mathematical models for task scheduling and resource allocation, combined with heuristic methods, demonstrate how edge and cloud resources can be strategically combined to deliver reliable performance under fluctuating network conditions and varied workload intensities [82]. In particular, local processing can significantly diminish latency for sensitive tasks, while offloading resource-intensive operations to the cloud can sustain the high throughput demanded by larger-scale analytics.

Still, it remains essential to acknowledge the limitations observed in both theoretical and empirical evaluations [83]. Rapid fluctuations in workload or abrupt changes in network conditions can disrupt system balance if scheduling decisions are not updated promptly. The introduction of highly dynamic task characteristics or extremely constrained resources may require more advanced adaptive algorithms, possibly involving predictive modeling or machine learning-based control. Implementing robust security features across a large and often geographically distributed deployment poses another considerable challenge, where encryption, identity management, and secure update mechanisms must function reliably on both low-power edge devices and powerful cloud servers. [84]

In real-world applications spanning industrial automation, healthcare monitoring, and city-wide sensor deployments, the potential exists for further refinements that exploit specific application characteristics. For example, domain-specific protocols or custom hardware acceleration could drastically optimize local computations, while hierarchical or multi-level caching might reduce the repeated transfer of similar data sets [85]. Additionally, strategies for energy conservation at the edge, particularly in battery-powered contexts, will become increasingly relevant as the number of deployed devices grows into the billions.

Scalability concerns also invite new lines of research into more robust orchestration mechanisms. Larger-scale systems may require multiple orchestrators working collaboratively, with regional clusters of edge nodes sharing partial states to collectively manage resource allocation [86]. Graph-based partitioning and distributed consensus algorithms may help maintain overall system coherence. A further extension would incorporate advanced network functions, such as dynamic routing or network slicing, to ensure that reliable, low-latency communication channels exist when needed. [87]

In conclusion, hybrid edge-cloud frameworks offer a compelling vision for tackling the constraints of big data processing in IoT ecosystems. By integrating distributed intelligence and centralized capabilities, these architectures can adapt to evolving conditions, strike a better balance between latency and throughput, and account for energy and security requirements. Ongoing research and development efforts are likely to refine these methods, making edge-cloud systems more resilient, efficient, and scalable across a growing array of real-world applications. [88]

## References

[1] M. Durresi, A. Subashi, A. Durresi, L. Barolli, and K. Uchida, “Secure communication architecture for internet of things using smartphones and multi-access edge computing in environment monitoring,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, pp. 1631–1640, 3 2018.

[2] M. Wörsdörfer, “What happened to ’big tech’ and antitrust? and how to fix them!,” *Philosophy of Management*, vol. 21, pp. 345–369, 1 2022.

[3] S. S. Jaffer, A. Hussain, M. A. Qureshi, and W. S. Khawaja, “Towards the shifting of 5g front haul traffic on passive optical network,” *Wireless Personal Communications*, vol. 112, pp. 1549–1568, 1 2020.

[4] A. Vyshedskiy, R. Netson, E. Fridberg, P. Jagadeesan, M. Arnold, S. Barnett, A. Gondalia, V. Maslova, L. de Torres, S. Ostrovsky, D. Durakovic, A. Savchenko, S. McNett, M. Kogan, I. Piryatinsky, and D. Gold, “Boston cognitive assessment (boca) - a comprehensive self-administered smartphone- and computer-based at-home test for longitudinal tracking of cognitive performance.,” *BMC neurology*, vol. 22, pp. 92–, 3 2022.

[5] C. A. Albayrak, O. Renn, and K. Teille, "Leitlinien für das menschliche handeln in einer digitalisierten welt," *HMD Praxis der Wirtschaftsinformatik*, vol. 55, pp. 1048–1064, 8 2018.

[6] S. Pal, R. Kumar, L. H. Son, K. Saravanan, M. Abdel-Basset, G. Manogaran, and P. H. Thong, "Novel probabilistic resource migration algorithm for cross-cloud live migration of virtual machines in public cloud," *The Journal of Supercomputing*, vol. 75, pp. 5848–5865, 5 2019.

[7] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Integrating polystore rdbms with common in-memory data," in *2020 IEEE International Conference on Big Data (Big Data)*, pp. 5762–5764, IEEE, 2020.

[8] C.-T. Yang, S.-T. Chen, C.-H. Chang, W. Den, and C.-C. Wu, "Implementation of an environmental quality and harmful gases monitoring system in cloud," *Journal of Medical and Biological Engineering*, vol. 39, pp. 456–469, 3 2018.

[9] R. Bertolo, A. J. Hung, F. Porpiglia, P. Bove, M. Schleicher, and P. Dasgupta, "Systematic review of augmented reality in urological interventions: the evidences of an impact on surgical outcomes are yet to come.," *World journal of urology*, vol. 38, pp. 2167–2176, 3 2019.

[10] E. Zaghloul, K. Zhou, and J. Ren, "P-mod: Secure privilege-based multilevel organizational data-sharing in cloud computing," *IEEE Transactions on Big Data*, vol. 6, pp. 804–815, 12 2020.

[11] J. Wang, "The impact of artificial intelligence (ai) on customer relationship management: A qualitative study," *International Journal of Management and Accounting*, pp. 74–90, 9 2023.

[12] M. Kansara, "Cloud migration strategies and challenges in highly regulated and data-intensive industries: A technical perspective," *International Journal of Applied Machine Learning and Computational Intelligence*, vol. 11, no. 12, pp. 78–121, 2021.

[13] E. Ippoliti and T. Nickles, "Introduction: Scientific discovery and inference," *Topoi*, vol. 39, pp. 835–839, 10 2019.

[14] S. Ma and S. Yang, "Covid-19 forecasts using internet search information in the united states.," *Scientific reports*, vol. 12, pp. 11539–, 7 2022.

[15] I. L. Olokodana, S. P. Mohanty, E. Kougianos, and O. O. Olokodana, "Real-time automatic seizure detection using ordinary kriging method in an edge-iomt computing paradigm," *SN Computer Science*, vol. 1, pp. 1–15, 8 2020.

[16] R. Avula, "Healthcare data pipeline architectures for ehr integration, clinical trials management, and real-time patient monitoring," *Quarterly Journal of Emerging Technologies and Innovations*, vol. 8, no. 3, pp. 119–131, 2023.

[17] X. Mu, Y. Du, and H. Wang, "Research on cloud-edge ac power monitoring systems based on modbus," *Journal of Physics: Conference Series*, vol. 2479, pp. 12007–012007, 4 2023.

[18] J. Arthanari and R. Baskaran, "Enhancement of video streaming analysis using cluster-computing framework," *Cluster Computing*, vol. 22, pp. 3771–3781, 3 2018.

[19] M. Chernyshev, S. Zeadally, and Z. A. Baig, "Healthcare data breaches: Implications for digital forensic readiness," *Journal of medical systems*, vol. 43, pp. 7–, 11 2018.

[20] J. M. Spector, "Smart learning futures: a report from the 3 rd us-china smart education conference," *Smart Learning Environments*, vol. 5, pp. 5–, 6 2018.

[21] M. Abouelyazid, "Forecasting resource usage in cloud environments using temporal convolutional networks," *Applied Research in Artificial Intelligence and Cloud Computing*, vol. 5, no. 1, pp. 179–194, 2022.

[22] T. W. Hazlett, "U.s. antitrust policy in the age of amazon, google, microsoft, apple, netflix and facebook," *Constitutional Political Economy*, vol. 35, pp. 73–108, 3 2023.

[23] M. Hassan and S. Bansal, "S3qlrdf: distributed sparql query processing using apache spark—a comparative performance study," *Distributed and Parallel Databases*, vol. 41, pp. 191–231, 1 2023.

[24] E. Grosholz, "On the plains and prairies of minnesota: The role of mathematical statistics in biological explanation," *Synthese*, vol. 199, pp. 5377–5393, 2 2021.

[25] A. V. Sadybekov and V. Katritch, "Computational approaches streamlining drug discovery.," *Nature*, vol. 616, pp. 673–685, 4 2023.

[26] P. Huang, Y. C. Wen, and Y. Zhang, "Does the monitoring effect of big 4 audit firms really prevail? evidence from managerial expropriation of cash assets," *Review of Quantitative Finance and Accounting*, vol. 55, pp. 739–768, 12 2019.

[27] C. Turner, "Augmented reality, augmented epistemology, and the real-world web," *Philosophy & Technology*, vol. 35, 3 2022.

[28] S. Jain, V. Tiwari, A. Thapa, R. Mangla, R. K. Jaiswal, V. Kumar, S. Tiwari, M. G. Tulbure, R. Galkate, A. K. Lohani, and K. Pandey, "Performance evaluation of google earth engine based precipitation datasets under different climatic zones over india," *Remote Sensing in Earth Systems Sciences*, vol. 5, pp. 263–276, 11 2022.

[29] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Big data in cloud computing review and opportunities," *arXiv preprint arXiv:1912.10821*, 2019.

[30] S. Smith, "Maximizing cloud computing benefits in the age of big data," *Zenodo (CERN European Organization for Nuclear Research)*, 11 2022.

[31] N. Craig, I. G. Garcia, and S. Koren, "The weak scale from weak gravity," *Journal of High Energy Physics*, vol. 2019, pp. 1–34, 9 2019.

[32] A. K. Saxena, R. R. Dixit, and A. Aman-Ullah, "An lstm neural network approach to resource allocation in hospital management systems," *International Journal of Applied Health Care Analytics*, vol. 7, no. 2, pp. 1–12, 2022.

[33] E. Humphreys, "The future landscape of isms standards," *Datenschutz und Datensicherheit - DuD*, vol. 42, pp. 421–423, 6 2018.

[34] R. Lawrence, Z. He, D. K. Chakravorty, W. Brashear, H. Liu, S. B. Nite, L. M. Perez, C. P. Francis, N. Dronamraju, X. Yang, T. Guleria, and J. Kim, "Cybersecurity and data science curriculum for secondary student computing programs," *The Journal of Computational Science Education*, vol. 14, no. 2, pp. 6–9, 2023.

[35] F. R. Ferraro, C. Pallanca, B. Lanzoni, C. crociati, E. Dalessandro, L. Origlia, R. M. Rich, S. Saracino, A. Mucciarelli, E. Valenti, D. Geisler, F. Mauro, S. Villanova, C. M. Bidin, and G. Beccari, "A new class of fossil fragments from the hierarchical assembly of the galactic bulge," *Nature Astronomy*, vol. 5, pp. 311–318, 12 2020.

[36] M. E. Maros, D. Capper, D. T. Jones, V. Hovestadt, A. von Deimling, S. M. Pfister, A. Benner, M. Zucknick, and M. Sill, "Machine learning workflows to estimate class probabilities for precision cancer diagnostics on dna methylation microarray data.," *Nature protocols*, vol. 15, pp. 479–512, 1 2020.

[37] Y. Wu, Y. D. Wei, and H. Li, "Analyzing spatial heterogeneity of housing prices using large datasets," *Applied Spatial Analysis and Policy*, vol. 13, pp. 223–256, 5 2019.

[38] M. Kansara, "A comparative analysis of security algorithms and mechanisms for protecting data, applications, and services during cloud migration," *International Journal of Information and Cybersecurity*, vol. 6, no. 1, pp. 164–197, 2022.

[39] J. Mao, D. Ifenthaler, T. Fujimoto, A. Garavaglia, and P. G. Rossi, "National policies and educational technology: a synopsis of trends and perspectives from five countries," *TechTrends*, vol. 63, pp. 284–293, 4 2019.

[40] S. Lu, Y. Yao, B. Luo, Z. Yu, D. Li, and W. Shi, "Edgeware: toward extensible and flexible middleware for connected vehicle services," *CCF Transactions on High Performance Computing*, vol. 4, pp. 339–356, 5 2022.

[41] A. Khedimi, T. Menouer, C. Cérin, and M. Boudhar, "A cloud weather forecasting service and its relationship with anomaly detection," *Service Oriented Computing and Applications*, vol. 16, pp. 191–208, 7 2022.

[42] F. Moreira, M. J. Ferreira, C. S. Pereira, A. S. Gomes, C. A. Collazos, and D. F. Escudero, "Eclectic as a learning ecosystem for higher education disruption," *Universal Access in the Information Society*, vol. 18, pp. 615–631, 7 2019.

[43] K. Sohn and G. Jang, "Ground vehicle driving by full sized humanoid," *Journal of Intelligent & Robotic Systems*, vol. 99, pp. 407–425, 12 2019.

[44] A. A. Nasr, A. T. Chronopoulos, N. A. El-Bahnasawy, G. Attiya, and A. El-Sayed, "A novel water pressure change optimization technique for solving scheduling problem in cloud computing," *Cluster Computing*, vol. 22, pp. 601–617, 11 2018.

[45] Z. Chen, "The insignificance reading of shenme revisited," *Journal of East Asian Linguistics*, vol. 30, pp. 83–107, 4 2021.

[46] A. V. Mithapalli and S. S. Joshi, "A framework for secure data storage and retrieval in cloud environment," *International Journal of Engineering and Advanced Technology*, vol. 9, pp. 2511–2520, 12 2019.

[47] X. Yu and Q. Liao, "Understanding user passwords through password prefix and postfix (p3) graph analysis and visualization," *International Journal of Information Security*, vol. 18, pp. 647–663, 4 2019.

[48] C. J. Anumba, A. Akanmu, X. Yuan, and C. Kan, "Cyber–physical systems development for construction applications," *Frontiers of Engineering Management*, vol. 8, pp. 72–87, 9 2020.

[49] J. Kileel, A. Moscovich, N. Zelesko, and A. Singer, "Manifold learning with arbitrary norms," *Journal of Fourier Analysis and Applications*, vol. 27, pp. 1–56, 9 2021.

[50] M. Attaran, "The impact of 5g on the evolution of intelligent automation and industry digitization," *Journal of ambient intelligence and humanized computing*, vol. 14, pp. 1–17, 2 2021.

[51] R. Avula *et al.*, "Data-driven decision-making in healthcare through advanced data mining techniques: A survey on applications and limitations," *International Journal of Applied Machine Learning and Computational Intelligence*, vol. 12, no. 4, pp. 64–85, 2022.

[52] Y. Liu and B. Dong, "Modeling urban scale human mobility through big data analysis and machine learning," *Building Simulation*, vol. 17, pp. 3–21, 8 2023.

[53] L. Truong-Hong, F. Jia, E. Che, S. Emamgholian, D. Laefer, and A. V. Vo, "Preface: 3d geoinfo 2021," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. VIII-4/W2-2021, pp. 1–4, 10 2021.

[54] S. Kang, R. Jin, X. Deng, and R. S. Kenett, "Challenges of modeling and analysis in cybermanufacturing: a review from a machine learning and computation perspective.," *Journal of intelligent manufacturing*, vol. 34, pp. 1–14, 8 2021.

[55] A. Razagui, K. Abdeladim, S. Semaoui, A. H. Arab, and S. Boulahchiche, "Modeling the forecasted power of a photovoltaic generator using numerical weather prediction and radiative transfer models coupled with a behavioral electrical model," *Energy Reports*, vol. 6, no. 1, pp. 57–62, 2020.

[56] F. Barreiro, D. Benjamin, T. Childers, K. De, J. Elmsheuser, A. Filipcic, A. Klimentov, M. Lassnig, T. Maeno, D. Oleynik, S. Panitkin, and T. Wenaus, "The future of distributed computing systems in atlas: Boldly venturing beyond grids," *EPJ Web of Conferences*, vol. 214, pp. 03047–, 9 2019.

[57] Y. F. Huang, J. L. Ng, K. F. Fung, T. K. Weng, N. AlDahoul, A. N. Ahmed, M. Sherif, B. Chaplot, K. L. Chong, and A. Elshafie, "Space–time heterogeneity of drought characteristics in sabah and sarawak, east malaysia: implications for developing effective drought monitoring and mitigation strategies," *Applied Water Science*, vol. 13, 9 2023.

[58] B. J. Lauck, A. M. Sinnott, A. W. Kiefer, D. A. Padua, J. R. Powell, H. R. Sledge, and J. P. Mihalik, "Association between head impact biomechanics and physical load in college football.," *Annals of biomedical engineering*, vol. 50, pp. 1437–1443, 8 2022.

[59] J. S. Yoo, D. Boulware, and D. Kimmey, "Parallel co-location mining with mapreduce and nosql systems," *Knowledge and Information Systems*, vol. 62, pp. 1433–1463, 8 2019.

[60] S. Gupta, M. Sharma, and R. Chawla, "Fpga implementation of r-fifo-based high-speed data acquisition iot architecture model," *SN Applied Sciences*, vol. 2, pp. 1–8, 3 2020.

[61] R. Maheshwari, S. Yadav, and P. Vidyarthic, "Cloud computing: A boon for education system," *SSRN Electronic Journal*, 2019.

[62] H. Gholami, C. K. Chang, P. Aduri, A. Ma, and B. Rekabdar, "A data-driven situation-aware framework for predictive analysis in smart environments," *Journal of Intelligent Information Systems*, vol. 59, pp. 679–704, 6 2022.

[63] N. Zincir-Heywood, R. Birke, E. Bou-Harb, G. Casale, K. El-Khatib, T. Inoue, N. Kumar, H. Lutfiyya, D. Puthal, A. Shami, N. Stakhanova, and F. Zulkernine, "Guest editorial: Special issue on machine learning and artificial intelligence for managing networks, systems, and services—part i," *IEEE Transactions on Network and Service Management*, vol. 19, no. 4, pp. 3988–3994, 2022.

[64] G. I. Simari, M. V. Martinez, F. R. Gallo, and M. A. Falappa, "The big-2/rose model of online personality," *Cognitive Computation*, vol. 13, pp. 1198–1214, 7 2021.

[65] R. Dillerup, T. Witzemann, S. Schacht, and L. Schaller, "Planung im digitalen zeitalter," *Controlling & Management Review*, vol. 63, pp. 46–53, 4 2019.

[66] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Context-aware query performance optimization for big data analytics in healthcare,” in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, pp. 1–7, 2019.

[67] T. Yi, I. Pan, S. Collins, F. Chen, R. Cueto, B. Hsieh, C. Hsieh, J. L. Smith, L. Yang, W. Liao, L. H. Merck, H. X. Bai, and D. Merck, “Dicom image analysis and archive (diana): an open-source system for clinical ai applications.,” *Journal of digital imaging*, vol. 34, pp. 1–9, 11 2021.

[68] D. Zhang and M. R. Kabuka, “Distributed relationship mining over big scholar data,” *IEEE Transactions on Emerging Topics in Computing*, vol. 9, pp. 354–365, 1 2021.

[69] L. Longchamps, B. Tisseyre, J. Taylor, L. Sagoo, A. Momin, S. Fountas, L. Manfrini, Y. Ampatzidis, J. K. Schueller, and R. Khosla, “Yield sensing technologies for perennial and annual horticultural crops: a review,” *Precision Agriculture*, vol. 23, pp. 2407–2448, 8 2022.

[70] S. M. Lee, D. Lee, and Y. S. Kim, “The quality management ecosystem for predictive maintenance in the industry 4.0 era,” *International Journal of Quality Innovation*, vol. 5, pp. 1–11, 3 2019.

[71] H. Labbaci, B. Medjahed, and Y. Aklouf, “A social network approach for recommending interoperable web services,” *Distributed and Parallel Databases*, vol. 38, pp. 927–961, 8 2020.

[72] T. C. Lim, “Use of the mchargian lusa in agricultural research and decision-making in the age of non-stationarity and big earth observation data,” *Socio-Ecological Practice Research*, vol. 1, pp. 297–324, 8 2019.

[73] N. Ogbuke, Y. Y. Yusuf, A. Gunasekaran, N. Colton, and D. Kovvuri, “Data-driven technologies for global healthcare practices and covid-19: opportunities and challenges,” *Annals of Operations Research*, 7 2023.

[74] R. Avula, “Applications of bayesian statistics in healthcare for improving predictive modeling, decision-making, and adaptive personalized medicine,” *International Journal of Applied Health Care Analytics*, vol. 7, no. 11, pp. 29–43, 2022.

[75] N. Dutta, S. Kar, and R. S. Sobel, “What influences entrepreneurship among skilled immigrants in the usa? evidence from micro-data,” *Business Economics*, vol. 56, pp. 146–154, 5 2021.

[76] M. Javaid, A. Haleem, R. P. Singh, S. Rab, R. Suman, and I. H. Khan, “Evolutionary trends in progressive cloud computing based healthcare: Ideas, enablers, and barriers,” *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 124–135, 2022.

[77] M. Kansara, “A structured lifecycle approach to large-scale cloud database migration: Challenges and strategies for an optimal transition,” *Applied Research in Artificial Intelligence and Cloud Computing*, vol. 5, no. 1, pp. 237–261, 2022.

[78] H. Bouzary and F. F. Chen, “Service optimal selection and composition in cloud manufacturing: a comprehensive survey,” *The International Journal of Advanced Manufacturing Technology*, vol. 97, pp. 795–808, 4 2018.

[79] H. A. Al-Jundi and E. Y. Tanbour, “Design and evaluation of a highfidelity virtual reality manufacturing planning system,” *Virtual Reality*, vol. 27, pp. 677–697, 8 2022.

[80] H. Yuan, M. Zhou, Q. Liu, and A. Abusorrah, “Fine-grained resource provisioning and task scheduling for heterogeneous applications in distributed green clouds,” *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1380–1393, 2020.

[81] L. Rompis, “A logic circuit simulation for finding identical or redundant files using counter and register,” *Advances in Image and Video Processing*, vol. 6, pp. 26–26, 6 2018.

[82] J. Ahmed, A. Malik, M. Ilyas, and J. S. Alowibdi, “Instance launch-time analysis of openstack virtualization technologies with control plane network errors,” *Computing*, vol. 101, pp. 989–1014, 5 2018.

[83] P. Kanagala and R. Jayaraman, “Effective encryption approach to improving the secure cloud framework through fuzzy-based encrypted cryptography,” *Soft Computing*, 4 2023.

[84] S. T. Milan, L. Rajabion, A. M. Darwesh, M. Hosseinzadeh, and N. J. Navimipour, “Priority-based task scheduling method over cloudlet using a swarm intelligence algorithm,” *Cluster Computing*, vol. 23, pp. 663–671, 6 2019.

[85] A. M. González-Zapata, E. Tlelo-Cuautle, I. Cruz-Vega, and W. D. Leon-Salas, “Synchronization of chaotic artificial neurons and its application to secure image transmission under mqtt for iot protocol,” *Nonlinear Dynamics*, vol. 104, pp. 4581–4600, 5 2021.

[86] S. Shah, A. Mulahuwaish, K. Z. Ghafoor, and H. S. Maghdid, “Prediction of global spread of covid-19 pandemic: a review and research challenges.,” *Artificial intelligence review*, vol. 55, pp. 1–22, 7 2021.

[87] V. K. Pathak, C. Nayak, R. Singh, M. K. Dikshit, and T. Sai, "Optimizing parameters in surface reconstruction of transtibial prosthetic socket using central composite design coupled with fuzzy logic-based model," *Neural Computing and Applications*, vol. 32, pp. 15597–15613, 4 2020.

[88] A. M. Jacob, "Small molecules, big impact: a tale of hydrides past, present, and future," *Astrophysics and Space Science*, vol. 368, 9 2023.