soloncouncil : Pages:13–26

Original Research

Energy-Efficient Resource Management Techniques for Big Data Workloads in Cloud Data Centers

Carrie Vander Peterson[1]

[1]Researcher at Universidad Univer.

Abstract

This paper addresses the critical challenge of managing energy consumption in cloud data centers running big data workloads. As organizations continue to migrate large-scale applications to the cloud, the demand for efficient resource provisioning, scheduling, and scaling becomes essential for both cost and sustainability reasons. Energy expenditures in data centers can be substantial, and inefficiencies often arise from underutilized resources, ineffective workload placement, and suboptimal scheduling algorithms. In this work, a comprehensive framework is proposed to reduce power consumption without degrading application performance or violating service-level requirements. The discussion encompasses novel mathematical models that capture the intricacies of workload characteristics, computational capacity, network overhead, and cooling demands. Advanced scheduling methods, including heuristic-based and optimization-driven techniques, are explored with the aim of balancing trade-offs between power savings and throughput. Furthermore, simulations and real-world tests demonstrate the feasibility of these approaches and highlight critical factors that influence practical performance, including workload heterogeneity, network constraints, and consolidation strategies. Finally, the paper addresses potential areas where the proposed model encounters limitations, emphasizing the impact of scale, dynamic workload fluctuations, and hardware diversity on optimization outcomes. Through its combination of rigorous theoretical modeling and empirical validation, this research offers valuable insights into designing more energy-aware cloud infrastructures.

1. Introduction

The emergence of large-scale computational workloads, facilitated by technological advances in data analytics, machine learning, and high-throughput applications, has intensified the need to manage cloud resources with greater care [1]. Data centers operating at scale increasingly encounter the dual pressure of maintaining strict performance guarantees while reducing overall energy consumption [2]. This tension has motivated the development of innovative methods to align resource availability with dynamic workload demands, thereby enhancing both operational efficiency and sustainability.

The significance of efficient resource management techniques becomes apparent when considering the enormous energy requirements of complex data center infrastructure [3]. The proliferation of highly parallelized big data frameworks has increased the strain on servers, networking elements, and cooling systems. Traditional scheduling algorithms, often designed under the assumption of relatively stable workloads, are now rendered insufficient, as they do not seamlessly accommodate the dynamic, bursty, and heterogeneous nature of modern big data applications [4]. This scenario has triggered research into the multi-faceted problem of reducing power usage without sacrificing quality of service, which remains a non-trivial objective given the intricate interdependencies among hardware resources, workload profiles, and physical constraints.

A pressing concern in energy-aware scheduling is the fundamental trade-off between performance metrics such as latency or throughput and the costs associated with underutilized or over-provisioned resources [5]. At certain utilization thresholds, server performance can degrade sharply, but powering down or consolidating workloads might elevate the risk of violating service-level objectives. This indicates the need for holistic solutions that model not just the computing aspects but also factors such as cooling overheads, hardware heterogeneity, and real-time traffic fluctuations [6]. The coupling of these elements renders a purely heuristic-based approach insufficient for capturing the complexities of energy usage in live operational environments.

A promising development has been the application of advanced optimization algorithms, which frame resource allocation and scheduling as optimization problems with well-defined objective functions [7]. Such approaches integrate performance metrics alongside energy considerations, enabling fine-grained management of how tasks are dispatched and servers are powered up or put into low-power states. Progress has been made by leveraging mathematical formulations that treat power as a function of resource usage and system load, often including linear, nonlinear, or convex terms that reflect hardware-specific behaviors [8]. These formulations can in turn be tackled by deterministic or metaheuristic methods, including simulated annealing or genetic algorithms. However, the complexity can become intractable at large scales or under stringent real-time constraints. [9]

Another layer of complexity arises from the shift toward distributed frameworks for big data processing, including map-reduce systems and streaming analytics platforms. The need to dynamically partition, replicate, or reassign tasks imposes an additional overhead on resource consumption, which must be factored into global scheduling decisions [10]. Communication delays and network constraints further complicate the objective of energy efficiency, especially under high throughput demands. Techniques to co-locate related tasks and exploit data locality can curtail unnecessary data transfers, thus reducing both bandwidth requirements and system-wide power usage [11, 12].

To address these multi-dimensional constraints, researchers have explored a combination of proactive and reactive scheduling policies. Proactive strategies use predictive models to anticipate demand spikes and plan resource allocations accordingly [13]. Reactive policies dynamically adjust servers, CPU frequencies, or container placements based on real-time monitoring. Integrating these approaches offers more robust performance in volatile environments, but balancing them remains a challenge, demanding sophisticated intelligence for continuously optimizing resource usage. [14]

In the sections that follow, a rigorous mathematical formulation for the resource management problem in big data cloud environments is introduced. This formulation captures the relationship between energy usage, workload characteristics, and hardware limitations [15]. Subsequently, a scheduling framework derived from this formulation is detailed, followed by an in-depth experimental evaluation using both simulated and real-world scenarios. Performance metrics are analyzed, highlighting the significant savings in power consumption that can be realized with minimal compromise on throughput [16]. This study underscores the potential for advanced modeling in designing energy-efficient data center solutions while acknowledging the practical impediments and limitations that arise, including scale, heterogeneous hardware, and rapidly shifting workload demands. Ultimately, it presents a comprehensive analysis of how these challenges can be mitigated or further researched to enable the next generation of sustainable cloud computing infrastructures. [17]

2. Mathematical Formulation for Resource Management

In this section, a mathematical formulation of the energy-efficient resource management problem is presented, capturing the essential interplay between computing elements, workload characteristics, and infrastructure constraints. The formulation proceeds by defining key variables, parameters, and constraints that collectively model the complexities of scheduling big data tasks in cloud data centers. [18, 19]

Consider a cloud data center composed of a set of physical servers [20]. Each server is characterized by its maximum processing capacity and a power function that translates utilization levels into energy consumption. Let there be N servers indexed by $i = 1, 2, \ldots, N$ [21]. Each server hosts one or more virtual machines (VMs) or containers. The total computational capacity of server i is denoted C_i [22]. Let there be M tasks indexed by $j = 1, 2, \ldots, M$, each requiring a certain amount of computational workload denoted W_j. The objective is to assign tasks to servers and manage their operational states such that the total power consumed across all servers is minimized, subject to performance constraints. [23]

Let x_{ij} be a binary decision variable indicating whether task j is assigned to server i or not:

$$x_{ij} = \begin{cases} 1 & \text{if task } j \text{ is assigned to server } i, \\ 0 & \text{otherwise.} \end{cases}$$

The power consumption of server i can be modeled as a convex function of its total utilization. Let u_i be the utilization of server i, computed as: [24]

$$u_i = \frac{\sum_{j=1}^{M} W_j x_{ij}}{C_i}.$$

In many practical scenarios, the power consumption can be approximated by a function of the form:

$$P_i(u_i) = P_{i,\text{idle}} + \alpha_i (u_i)^{\beta},$$

where $P_{i,\text{idle}}$ is the idle power consumption of server i, α_i is a coefficient indicating how fast power usage scales with utilization, and β is a parameter describing the nonlinearity of the power curve. To capture server states, define another binary variable s_i that indicates whether server i is active ($s_i = 1$) or powered down ($s_i = 0$). [25]

The total energy cost can then be formulated as:

$$\text{Energy Cost} = \sum_{i=1}^{N} s_i \left(P_{i,\text{idle}} + \alpha_i \left(\frac{\sum_{j=1}^{M} W_j x_{ij}}{C_i} \right)^{\beta} \right).$$

The primary optimization objective is to minimize this quantity: [26]

$$\min_{x_{ij}, s_i} \quad \sum_{i=1}^{N} s_i \left(P_{i,\text{idle}} + \alpha_i \left(\frac{\sum_{j=1}^{M} W_j x_{ij}}{C_i} \right)^{\beta} \right).$$

This objective is subject to several constraints. First, each task j must be assigned to exactly one server: [27]

$$\sum_{i=1}^{N} x_{ij} = 1 \quad \forall j.$$

Second, if no task is assigned to server i, that server can be powered down ($s_i = 0$), but if at least one task is assigned, the server must be active ($s_i = 1$). This can be expressed as: [28]

$$\sum_{j=1}^{M} x_{ij} \leq M\, s_i \quad \forall i.$$

Additionally, server capacity must not be exceeded:

$$\sum_{j=1}^{M} W_j x_{ij} \leq C_i \quad \forall i.$$

This basic model can be extended to account for cooling overhead, network bandwidth constraints, or memory usage by introducing additional variables and constraints [29, 30]. For instance, in many systems the cooling energy can be approximated by a function dependent on the total heat dissipated, which is in turn related to the total power consumption of active servers. Incorporating such relationships can transform the problem into a more complex, but also more realistic, multi-variable optimization. [31]

In certain high-level frameworks, a continuous version of the resource allocation variables is considered, allowing partial assignment of tasks to different servers. This might arise in container-based architectures or data-parallel processing tasks where splitting is feasible [32]. The corresponding problem can then be addressed using Lagrangian relaxation or the Karush-Kuhn-Tucker conditions for optimality. The partial derivatives of the objective with respect to resource allocation fractions can guide gradient-based search algorithms or iterative procedures. [33, 34]

The complexity of the above formulations often grows combinatorially with the number of tasks and servers, making exact solutions computationally prohibitive for large-scale deployments. Therefore, in practice, heuristic and metaheuristic approaches are crucial for generating near-optimal solutions within acceptable time bounds [35]. Nonetheless, the mathematical model provides an essential theoretical underpinning, guiding the design of algorithms and giving insight into how best to balance server utilization and power usage in big data environments.

3. Proposed Scheduling Framework

Building upon the mathematical model, this section introduces a scheduling framework designed to operationalize energy-efficient resource management for big data workloads [36]. The framework aims to mitigate the combinatorial challenges inherent in large-scale optimization while accommodating practical issues such as workload bursts, node failures, and heterogeneous infrastructure capabilities.

The foundation of the proposed framework is an intelligent controller that periodically executes an optimization process [37]. At each scheduling interval, the controller collects data on resource usage, temperature levels, queue lengths, pending jobs, and network statistics. These data points provide a snapshot of the data center's state, which is then fed into the optimization module to generate updated scheduling decisions [38]. Specifically, the module attempts to reassign tasks to servers or adjust container placements in a manner that reduces power consumption, subject to latency constraints and throughput requirements. [39]

To navigate the complexity of the combinatorial problem, the framework employs a multi-stage optimization algorithm. The first stage uses a coarse-grained, often greedy approach to partition workloads into broad categories according to their resource profiles, such as CPU-intensive or memory-intensive tasks [40]. Once partitioned, the tasks within each category are assigned to clusters of servers that are best suited to handle their specific demands. This partition-based strategy reduces the dimensionality of the problem, allowing the second stage to focus on finer-grained optimizations within each cluster. [41]

In the second stage, a local search or metaheuristic algorithm is applied to each cluster. Approaches like simulated annealing, genetic algorithms, or particle swarm optimization can be deployed [42]. These algorithms iteratively refine task assignments and server power states in order to minimize the power function described in the mathematical model. By confining the search space to a smaller cluster of nodes, the local search process can explore configurations more exhaustively, potentially converging to near-optimal solutions in a fraction of the time it would take for a data center-wide approach. [43]

The framework also integrates dynamic consolidation techniques for idle resources. If server utilization in a particular cluster drops below a threshold, the consolidation module re-evaluates task placements to power down underutilized servers [44]. Through live migration of VMs or containers, tasks can be relocated to a smaller set of active servers, thus reducing the idle power overhead. Here, a cost function balancing migration overhead against energy savings can be included [45]. Migrations might cause temporary performance degradation or network traffic surges, so an optimal consolidation plan minimizes disturbance while achieving significant energy savings.

Another key component is the adaptive frequency scaling policy [46]. Modern processors allow voltage and frequency adjustments at runtime, thereby controlling the power drawn by each server. The scheduling framework uses feedback from performance counters to adjust CPU frequencies [47]. When a server is lightly loaded, the controller reduces its frequency to match the current workload requirements.

Conversely, when critical tasks with strict latency constraints arrive, the framework can temporarily boost the frequency to ensure timely completion [48]. This dynamic approach effectively modulates power usage in direct response to changing demands, complementing the broader task allocation strategy.

Practical realizations of the proposed framework must handle transient overloads, node failures, and sudden shifts in workload composition [49]. To address these uncertainties, a proactive module forecasts upcoming load using historical data and machine learning models, often employing regression or neural network approaches to predict the volume of incoming tasks. This forecast influences both resource allocation decisions and consolidation strategies, thereby reducing the risk of unexpected system overload [50]. For instance, if the forecast indicates a surge in CPU-intensive tasks, the framework might delay certain consolidation actions or spin up additional nodes preemptively.

The underlying system architecture can be thought of as a layered design [51]. The bottom layer consists of physical servers and network links, each monitored by software agents that report metrics to the higher-level controller. The middle layer is the optimization engine, which encapsulates the modules for coarse-grained partitioning, local search, consolidation, and adaptive frequency scaling [52]. The top layer provides interfaces for administrators and users to specify performance objectives, such as maximum latency or minimum throughput guarantees. Periodically, the controller reconciles these objectives with the optimization outputs, ensuring that the system operates within acceptable performance bounds while minimizing energy consumption. [53]

To evaluate the complexity of the scheduling framework, note that coarse partitioning dramatically reduces search space size. Subsequent local searches, while still potentially expensive for large clusters, proceed in parallel across different partitions [54]. Techniques to accelerate convergence include caching evaluations of similar assignment configurations and applying partial gradient information derived from approximate cost functions [55]. These methods help ensure that the scheduling framework can operate within practical time limits, even in massive data center environments with thousands of servers and tens of thousands of tasks.

Taken together, the framework blends theoretical rigor from the mathematical model with the pragmatic requirements of large-scale data center operations [56]. Its modular design allows for incremental improvements and the introduction of specialized policies tailored to specific workloads or hardware. In the next section, this framework is evaluated via both synthetic simulations and real-world testbeds running typical big data applications, including batch analytics, streaming pipelines, and machine learning workloads. [57, 58]

4. Experimental Evaluation and Performance Analysis

In order to validate the proposed scheduling framework and assess its effectiveness in energy-efficient resource management, a series of experiments was conducted under both simulated conditions and real-world testbed environments. The evaluation spanned diverse workload types, including high-throughput batch jobs, latency-sensitive streaming pipelines, and mixed workloads characteristic of modern big data platforms [59]. This variety aimed to stress-test the scheduling algorithm against different usage patterns and performance constraints.

In the simulated setting, a custom-built workload generator modeled user arrival rates, job sizes, and resource usage profiles based on historical traces from production systems [60, 61]. The underlying data center topology, number of servers, and power consumption parameters were systematically varied. Each configuration was run multiple times with different random seeds to capture stochastic variations [62]. Baseline scenarios employed conventional methods, such as round-robin scheduling or a simple first-fit approach, allowing for direct comparison with the proposed solution.

Performance metrics focused on two main categories: energy consumption and Quality of Service (QoS) outcomes [63]. Energy consumption measurements derived from the computational model of server utilization and were aggregated over each scheduling interval. QoS outcomes included task completion times, average latency, and throughput [64]. An additional metric was the rate of task migration, which served as a measure of the overhead introduced by consolidation and dynamic reassignments. The experiments recorded these metrics at fine-grained time intervals, capturing transient behavior during workload spikes or when new scheduling decisions took effect. [65, 66]

Simulation results indicated that the proposed framework achieved between 15 and 25 percent reduction in overall energy consumption compared to baseline strategies. The extent of the reduction varied with workload intensity and the proportion of CPU-intensive tasks [67]. In particular, the local search phase in the multi-stage optimization effectively grouped tasks onto fewer servers at lower load levels, enabling the powering down of underutilized machines. This consolidation strategy was seen to yield significant energy savings [68]. However, the cost of migration was non-negligible in some scenarios, occasionally causing short-lived latency spikes for certain tasks. Nonetheless, the average QoS remained within the specified service-level thresholds, showing minimal degradation in throughput or completion times. [69]

A key factor influencing the level of power reduction was the curvature of the power function. When β was set to values reflecting a strong nonlinearity in the power curve, such that incremental utilization increases caused disproportionately higher energy consumption, the framework was more aggressive in load balancing [70]. This led to more frequent consolidation moves and higher migration overhead. Conversely, for a relatively linear power curve, the algorithm allocated tasks in a more balanced manner, resulting in moderate consolidation but fewer migrations [71]. These findings highlight the importance of correctly modeling hardware power characteristics to tune the aggressiveness of the optimization strategy.

In real-world tests, a cluster of 50 physical servers was set up, each equipped with instrumentation for power monitoring [72]. The cluster hosted containerized big data applications, including map-reduce batch jobs and a streaming analytics platform processing real-time sensor data. The proposed scheduling controller was integrated with a container orchestration system that exposed APIs for moving containers between servers and adjusting CPU frequency settings [73]. During the experiments, load generators provided synthetic traffic in patterns that ranged from stable to highly volatile.

Empirical results were consistent with simulation outcomes, revealing an energy savings of around 20 percent on average when compared to the default scheduling policies [74]. The consolidation feature proved particularly effective during off-peak hours, when the workload was light but continuous [75]. On the other hand, during peak loads, the framework largely avoided migrations to prevent performance degradation. In this regime, adaptive frequency scaling contributed to power reduction by modulating CPU frequencies in response to real-time utilization [76]. The synergy between consolidation and frequency scaling ensured that power usage was continually optimized without risking violation of critical deadlines.

Performance analysis showed that the new scheduling policy could occasionally create hotspots in the network when tasks with high data transfer requirements were concentrated on the same servers [77]. Although data locality was generally improved by co-locating tasks that share common input datasets, there were cases in which the volume of inter-node traffic surged momentarily. Additional refinements to the scheduling logic, particularly with regard to network-aware placement strategies, appear to be warranted [78, 79].

Scalability tests were undertaken by increasing the number of tasks to the limits of the real cluster, running thousands of containers with both short and long durations. While the scheduling framework continued to function effectively, its optimization cycle times grew with the system size [80]. Parallelizing the local search stage across multiple cores and nodes partially mitigated this slowdown, but the overhead remained noticeable. This underscores the importance of careful tuning of the scheduling interval: setting it too short risks excessive overhead due to constant decision-making, whereas too long an interval may allow suboptimal configurations to persist, thereby eroding potential energy gains. [81]

Overall, the evaluation demonstrated that the proposed scheduling framework yields tangible energy savings in big data workloads. The synergy between partition-based assignment, local search optimizations, and dynamic management features such as frequency scaling underscores the value of a holistic approach [82]. Nonetheless, results also exposed certain limitations and areas for future development, emphasizing the trade-offs that arise in balancing energy efficiency with system throughput and stability in large-scale cloud environments.

5. Challenges and Limitations

Despite the promising results obtained from simulations and real-world experiments, several challenges and limitations remain, highlighting areas for improvement and further investigation [83]. One such challenge is the inherent complexity of the optimization formulation. For very large data centers, with thousands of physical servers and an even higher number of tasks, the combinatorial explosion can overwhelm even sophisticated heuristic methods [84]. Although multi-stage strategies mitigate some of these computational burdens, they do not completely eliminate the overhead associated with frequent re-optimizations, particularly during peak or bursty workloads. Future research may explore more advanced approximation algorithms or specialized hardware accelerators capable of running large-scale optimization routines faster. [85, 86]

An additional concern is the accuracy of power consumption models. While polynomial or convex functions provide a tractable representation of power usage, real servers exhibit more complex behaviors influenced by temperature, component aging, and non-uniform resource usage across CPU cores [87]. Deviations between modeled and actual energy consumption can lead the optimization to make suboptimal decisions, or in worse cases, risk thermal overload in unmonitored components. This limitation calls for continuously refined models that can incorporate real-time feedback from sensors and dynamic thermal conditions [88]. Adaptive calibration techniques, in which the power model is periodically updated based on actual usage data, may address this gap.

Another challenge is ensuring reliable performance under rapidly fluctuating workloads [89]. While dynamic consolidation and frequency scaling can achieve considerable energy savings, they may also introduce performance instability if migrations or frequency adjustments are triggered too frequently. For latency-sensitive tasks, such as real-time analytics or online transaction processing, even momentary slowdowns can be detrimental [90]. The controller must therefore incorporate robust safety margins in its decision-making, possibly by applying conservative thresholds for triggering migrations or frequency scaling events. Balancing the desire for aggressive power savings with the need for stable performance remains an open research question. [91, 92]

Heterogeneous hardware environments introduce further complications. Modern data centers may host machines with different generations of CPUs, specialized accelerators such as GPUs, or memory modules with varying bandwidth capacities [93]. Assigning tasks to machines that do not align with their computational or memory demands can degrade application performance and offset any potential energy savings [94]. Extending the basic optimization formulation to encompass heterogeneous resources requires more elaborate decision variables and constraints, which further increases computational complexity. Finding techniques to handle resource heterogeneity at scale is an important frontier for future work. [95]

Network-aware scheduling, though partially addressed through data locality considerations, also warrants deeper investigation. Data-intensive applications often incur high communication costs if tasks are placed on physically distant servers, leading to increased network power consumption and potential bottlenecks [96]. While the proposed framework attempts to co-locate tasks that share input datasets, an optimized solution might account for network link constraints, latency requirements, and the fluctuating state of intermediate routing hardware. Incorporating a model that combines network power usage, link capacity, and topology awareness would be a logical step toward a more holistic representation of data center energy consumption. [97]

On the operational side, limitations arise from the overhead of reconfigurations. Live migrations of containers or VMs, while effective for consolidation, can momentarily spike CPU and network usage [98]. They can also introduce complexities when combined with other system-level events, such as security patching or hardware maintenance. Designing orchestration platforms that seamlessly integrate energy-aware scheduling with these operational tasks remains challenging [99]. Similarly, there may be organizational constraints, such as strict service-level agreements or compliance protocols, that limit the extent to which servers can be powered down or reallocated.

Finally, although the experimental evaluation provides valuable insight, it remains difficult to fully replicate the dynamic conditions of a production environment [100]. Real-world workloads exhibit long-tail behaviors, unexpected surges, and usage patterns influenced by external factors. The reliability of the proposed techniques in these scenarios requires either extended pilot deployments or robust simulation frameworks that emulate a wider range of unpredictable events [101]. Further, quantifying the return on investment for adopting such advanced scheduling methods—especially in smaller data centers—may necessitate detailed cost-benefit analyses, factoring in energy tariffs, cooling configurations, and hardware depreciation rates.

In summary, these challenges and limitations indicate that while the proposed resource management framework offers tangible benefits, it is not a panacea [102]. Its successful deployment in large-scale commercial or research data centers will depend on ongoing refinements in optimization methods, more accurate power and thermal models, and improved orchestration capabilities. As data center scales and technologies continue to evolve, energy efficiency will remain a critical area of research, requiring ongoing innovation in both algorithmic design and system engineering. [103]

6. Conclusion

This paper has examined energy-efficient resource management techniques for big data workloads in cloud data centers, presenting a thorough mathematical model alongside a multi-stage scheduling framework. By capturing the interplay between server utilization, nonlinear power consumption, and workload heterogeneity, the proposed model offers a foundation for theoretical exploration and algorithmic innovation [104]. The multi-stage scheduling framework integrates coarse partitioning, local search optimization, dynamic consolidation, and adaptive frequency scaling to reduce overall power usage while maintaining adherence to performance requirements.

Experimental evaluations, performed both under simulated workloads and on a cluster running real big data applications, demonstrated energy savings of approximately 15 to 25 percent compared to traditional or more static scheduling methods [105]. These gains were achieved without significant compromise to throughput or latency, though certain overheads arose from the increased complexity of migration and the adaptive control logic. The experiments also highlighted the influence of factors such as power function nonlinearity, workload composition, and hardware homogeneity on the efficacy of the proposed solutions. [106]

Despite its potential benefits, the framework faces limitations, including the computational expense of large-scale optimization, the intricacies of precisely modeling power usage, and the risks of performance instability under volatile workloads. Additionally, heterogeneous hardware and network constraints add layers of complexity that warrant further refinement of the scheduling approach [107]. The experiments provided evidence that advanced, data-driven strategies hold promise in dynamically adjusting resource usage and server states, but they also underscored the need for continuous modeling improvements and more extensive real-world validation.

In conclusion, energy-aware scheduling for big data workloads represents an increasingly significant challenge in large-scale cloud environments [108]. Through a combination of mathematical rigor and practical experimentation, this work contributes to the broader discourse on sustainable computing. The momentum behind energy reduction initiatives remains strong, and the strategies outlined here offer a viable path forward [109]. Future research efforts will likely focus on more sophisticated modeling of hardware and network interactions, improved orchestration mechanisms, and adaptive algorithms that can maintain stable performance under highly variable and demanding conditions. Such progress is essential to ensuring that cloud infrastructures can meet rising computational demands with minimal environmental impact. [110]

References

[1] K. Wang, "An adaptive graph sampling framework for graph analytics," *Social Network Analysis and Mining*, vol. 14, 12 2023.

[2] X. Yin and A. T. Chen, "A survey on service level components in big-cloud-iot systems with hybrid meta-heuristic techniques," *International Journal of Advanced Information and Communication Technology*, pp. 95–101, 7 2020.

[3] S. K. Chettri, D. Debnath, and P. Devi, "Leveraging digital tools and technologies to alleviate covid-19 pandemic," *SSRN Electronic Journal*, 2020.

[4] M. A. LoPresti, A. M. Alhajjat, K. C. Ott, F. Scorletti, E. E. Rowell, J. Bolden, X. Pombar, M. Coglan, N. Hroma, J. Ito, A. Boat, A. F. Shaaban, R. Bowman, V. M. Lu, A. Sáenz, S. Banh, A. L. David, J. Deprest, M. Z. Tahir, D. Thompson, F. V. Calenbergh, R. Devlieger, L. D. Catte, L. Lewi, H. Maes, P. D. Vloo, G. Esposito, P. Palma, G. Mosiello, B. D. Iacobelli, G. Lucignani, C. Marras, and D. Pang, "28th congress of the european society for pediatric neurosurgery (espn) rome-italy, 7-10 may 2023," *Child's Nervous System*, vol. 39, pp. 1369–1443, 4 2023.

[5] M. Talal, A. A. Zaidan, B. B. Zaidan, O. S. Albahri, M. A. Alsalem, A. S. Albahri, A. Alamoodi, L. M. Kiah, F. M. Jumaah, and M. Alaa, "Comprehensive review and analysis of anti-malware apps for smartphones," *Telecommunication Systems*, vol. 72, pp. 285–337, 5 2019.

[6] B. R. Chabowski, P. Gabrielsson, G. T. M. Hult, and F. V. Morgeson, "Sustainable international business model innovations for a globalizing circular economy: a review and synthesis, integrative framework, and opportunities for future research," *Journal of International Business Studies*, vol. 56, pp. 383–402, 10 2023.

[7] T. Griffin and J. Lowenberg-DeBoer, "Modeling local terrain attributes in landscape-scale site-specific data using spatially lagged independent variable via cross regression," *Precision Agriculture*, vol. 21, pp. 937–954, 11 2019.

[8] L. Luciano, I. Kiss, P. W. Beardshear, E. Kadosh, and A. B. Hamza, "Wise: A computer system performance index scoring framework," *Journal of Cloud Computing*, vol. 10, pp. 1–15, 1 2021.

[9] H. Doad and T. Gupta, "What the future upholds," *International Journal of Advanced Research in Science, Communication and Technology*, pp. 371–384, 10 2021.

[10] J.-P. Pinelli, M. Esteva, E. M. Rathje, D. B. Roueche, S. J. Brandenberg, G. Mosqueda, J. E. Padgett, and F. L. Haan, "Disaster risk management through the designsafe cyberinfrastructure," *International Journal of Disaster Risk Science*, vol. 11, pp. 719–734, 11 2020.

[11] D. Miehle, B. Häckel, S. Pfosser, and J. Übelhör, "Modeling it availability risks in smart factories," *Business & Information Systems Engineering*, vol. 62, pp. 323–345, 8 2019.

[12] M. Muniswamaiah, T. Agerwala, and C. Tappert, "Data virtualization for analytics and business intelligence in big data," in *CS & IT Conference Proceedings*, vol. 9, CS & IT Conference Proceedings, 2019.

[13] A. Minhajuddin, M. K. Jha, C. R. C. Fatt, T. L. Mayes, J. D. Berry, M. H. Trivedi, J. Dennison, C. Volmar, J. A. Timmons, C. Wahlestedt, A. A. Keiser, T. Dong, E. A. Kramár, C. Butler, D. P. Matheos, L. Tong, N. Berchtold, S. Chen, M. Samad, J. Beardwood, S. Shanur, A. M. Rodriguez, P. Baldi, C. W. Cotman, M. A. Wood, M. Ananth, D. A. Talmage, V. K. Martinez, B. McAlpin, R. Mahalingam, I. Mahant, A. Kavelaars, and C. J. Heijnen, "Acnp 60th annual meeting: Poster abstracts p1 - p275.," *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*, vol. 46, pp. 72–217, 12 2021.

[14] S. Yoo, "Investigating important urban characteristics in the formation of urban heat islands: a machine learning approach," *Journal of Big Data*, vol. 5, pp. 2–, 1 2018.

[15] T. Mungai, "Cloud computing in managing big data," *European Journal of Engineering and Technology Research*, vol. 1, pp. 30–33, 7 2018.

[16] H. Shao, W. Li, W. Kang, and S. J. Rey, "When spatial analytics meets cyberinfrastructure: an interoperable and replicable platform for online spatial-statistical-visual analytics," *Journal of Geovisualization and Spatial Analysis*, vol. 4, pp. 1–16, 6 2020.

[17] I. Bello, H. Chiroma, U. A. Abdullahi, A. Y. Gital, F. Jauro, A. Khan, J. O. Okesola, and S. M. Abdulhamid, "Detecting ransomware attacks using intelligent algorithms: recent development and next direction from deep learning and big data perspectives," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 8699–8717, 11 2020.

[18] A. Delgoshaei, R. Beighizadeh, M. K. A. B. M. Arffin, Z. B. Leman, and A. Ali, "Forecast innovative development level in green supply chains using a comprehensive fuzzy algorithm," *International Journal of Fuzzy Systems*, vol. 25, pp. 880–895, 12 2022.

[19] M. Abouelyazid and C. Xiang, "Architectures for ai integration in next-generation cloud infrastructure, development, security, and management," *International Journal of Information and Cybersecurity*, vol. 3, no. 1, pp. 1–19, 2019.

[20] A. Obukhova, E. Merzlyakova, I. Ershova, and K. Karakulina, "Introduction of digital technologies in the enterprise," *E3S Web of Conferences*, vol. 159, pp. 04004–, 3 2020.

[21] P. K. Pradhan, M. L. R. Liberato, V. Kumar, S. V. B. Rao, J. A. Ferreira, and T. Sinha, "Simulation of mid-latitude winter storms over the north atlantic ocean: impact of boundary layer parameterization schemes," *Climate Dynamics*, vol. 53, pp. 6785–6814, 9 2019.

[22] C. Sur, "Genatseq gan with heuristic reforms for knowledge centric network with browsing characteristics learning, individual tracking and malware detection with website2vec," *SN Computer Science*, vol. 1, pp. 1–18, 7 2020.

[23] A. Weingram, Y. Li, H. Qi, D. Ng, L. Dai, and X. Lu, "xccl: A survey of industry-led collective communication libraries for deep learning," *Journal of Computer Science and Technology*, vol. 38, pp. 166–195, 3 2023.

[24] M. T. Vafea, E. Atalla, J. Georgakas, F. Shehadeh, E. K. Mylona, M. Kalligeros, and E. Mylonakis, "Emerging technologies for use in the study, diagnosis, and treatment of patients with covid-19.," *Cellular and molecular bioengineering*, vol. 13, pp. 1–9, 6 2020.

[25] A. Telenti and X. Jiang, "Treating medical data as a durable asset.," *Nature genetics*, vol. 52, pp. 1005–1010, 9 2020.

[26] S. Malik, R. Rouf, K. Mazur, and A. Kontsos, "The industry internet of things (iiot) as a methodology for autonomous diagnostics in aerospace structural health monitoring," *Aerospace*, vol. 7, pp. 64–, 5 2020.

[27] X. Solé-Beteta, J. Navarro, D. Vernet, A. Zaballos, R. Torres-Kompen, D. Fonseca, and A. Briones, "Automatic tutoring system to support cross-disciplinary training in big data," *The Journal of Supercomputing*, vol. 77, pp. 1818–1852, 5 2020.

[28] R. K. Vemuri, P. C. S. Reddy, B. S. P. Kumar, J. Ravi, S. Sharma, and S. Ponnusamy, "Deep learning based remote sensing technique for environmental parameter retrieval and data fusion from physical models," *Arabian Journal of Geosciences*, vol. 14, pp. 1–10, 6 2021.

[29] C. Luna-Nevarez, "Neuromarketing, ethics, and regulation: An exploratory analysis of consumer opinions and sentiment on blogs and social media," *Journal of Consumer Policy*, vol. 44, pp. 559–583, 8 2021.

[30] 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.

[31] H. Goldspiel, B. Barr, J. Badding, and D. Kuehn, "Snapshots of nature-based recreation across rural landscapes: Insights from geotagged photographs in the northeastern united states.," *Environmental management*, vol. 71, pp. 234–248, 10 2022.

[32] W. L. Filho, P. Yang, J. H. P. P. Eustachio, A. M. Azul, J. C. Gellers, A. Gielczyk, M. A. P. Dinis, and V. Kozlova, "Deploying digitalisation and artificial intelligence in sustainable development research.," *Environment, development and sustainability*, vol. 25, pp. 4957–4988, 3 2022.

[33] S. Wolfram, "Are all fish the same shape if you stretch them? the victorian tale of on growth and form," *The Mathematical Intelligencer*, vol. 40, pp. 39–61, 9 2018.

[34] A. K. Saxena, "Evaluating the regulatory and policy recommendations for promoting information diversity in the digital age," *International Journal of Responsible Artificial Intelligence*, vol. 11, no. 8, pp. 33–42, 2021.

[35] S. G. Farrag, N. Sahli, Y. El-Hansali, E. M. Shakshuki, A. Yasar, and H. Malik, "Stimf: a smart traffic incident management framework," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 85–101, 1 2021.

[36] R. Mehmood, A. Sheikh, C. Catlett, and I. Chlamtac, "Editorial: Smart societies, infrastructure, systems, technologies, and applications," *Mobile Networks and Applications*, vol. 28, pp. 598–602, 5 2022.

[37] R. Colpari, N. Sajjad, A. Kiran, M. Chakraborty, V. Tripathi, P. Baranwal, B. Janardhana, D. Stepanova, and D. Wischert, "Conceptual analysis for a technology demonstration mission of the ion beam shepherds," *CEAS Space Journal*, vol. 15, pp. 567–584, 7 2022.

[38] A. Acker and A. Kreisberg, "Social media data archives in an api-driven world," *Archival Science*, vol. 20, pp. 105–123, 9 2019.

[39] X. Luo, "Call for papers," *IEEE Network*, vol. 36, no. 3, pp. 47–47, 2022.

[40] R. Priyadarshini, R. K. Barik, H. Dubey, and B. K. Mishra, "A survey of fog computing-based healthcare big data analytics and its security," *International Journal of Ambient Computing and Intelligence*, vol. 12, pp. 53–72, 4 2021.

[41] J. A. Warrick, A. C. Ritchie, K. M. Schmidt, M. E. Reid, and J. B. Logan, "Characterizing the catastrophic 2017 mud creek landslide, california, using repeat structure-from-motion (sfm) photogrammetry," *Landslides*, vol. 16, pp. 1201–1219, 3 2019.

[42] R. Searles, S. Herbein, T. Johnston, M. Taufer, and S. Chandrasekaran, "Creating a portable, high-level graph analytics paradigm for compute and data-intensive applications," *International Journal of High Performance Computing and Networking*, vol. 13, no. 1, pp. 105–105, 2019.

[43] B. A. Konomi, E. L. Kang, A. Almomani, and J. Hobbs, "Bayesian latent variable co-kriging model in remote sensing for quality flagged observations," *Journal of Agricultural, Biological and Environmental Statistics*, vol. 28, pp. 423–441, 2 2023.

[44] Z. Moon, R. Horne, A. Phillips, G. Özakinci, A. Sobota, M. Johnston, K. Guastaferro, H. Anne, M. P. Schaub, H. López-Pelayo, N. Boumparis, Z. Khadjesari, M. Blankers, H. Riper, L. Pas, L. Gelberg, A. Ghosh, A. Gonzalez, and A. Gual, "17th international congress of behavioral medicine," *International Journal of Behavioral Medicine*, vol. 30, pp. 1–165, 7 2023.

[45] K. Alsubhi, Z. Imtiaz, A. Raana, M. U. Ashraf, and B. Hayat, "Meacc: an energy-efficient framework for smart devices using cloud computing systems," *Frontiers of Information Technology & Electronic Engineering*, vol. 21, pp. 917–930, 7 2020.

[46] S. M. Chester, W. Landry, J. Liu, D. Poland, D. Simmons-Duffin, N. Su, and A. Vichi, "Carving out ope space and precise o(2) model critical exponents," *Journal of High Energy Physics*, vol. 2020, pp. 142–, 6 2020.

[47] C. Woods, H. Yu, and H. Huang, "Predicting the success of entrepreneurial campaigns in crowdfunding: a spatio-temporal approach," *Journal of Innovation and Entrepreneurship*, vol. 9, pp. 1–23, 7 2020.

[48] S. Ullah, M. D. Awan, and M. S. H. Khiyal, "Big data in cloud computing: A resource management perspective," *Scientific Programming*, vol. 2018, pp. 1–17, 2018.

[49] A. E. Kamal, L. Han, L. Lu, and S. Jabbar, "Guest editorial: Special issue on software defined networking: Trends, challenges, and prospective smart solutions," *Peer-to-Peer Networking and Applications*, vol. 12, pp. 291–294, 1 2019.

[50] H. C. Ates, P. Q. Nguyen, L. Gonzalez-Macia, E. Morales-Narváez, F. Güder, J. J. Collins, and C. Dincer, "End-to-end design of wearable sensors.," *Nature reviews. Materials*, vol. 7, pp. 887–907, 7 2022.

[51] K. Bi, D. Lin, Y. Liao, C.-H. Wu, and P. Parandoush, "Additive manufacturing embraces big data," *Progress in Additive Manufacturing*, vol. 6, pp. 181–197, 3 2021.

[52] S. Mehrban, M. A. Khan, M. W. Nadeem, M. Hussain, M. M. Ahmed, O. Hakeem, S. Saqib, L. M. Kiah, F. Abbas, and M. Hassan, "Towards secure fintech: A survey, taxonomy, and open research challenges," *IEEE Access*, vol. 8, pp. 23391–23406, 2020.

[53] P. Rouzrokh, B. Khosravi, S. Vahdati, M. Moassefi, S. Faghani, E. Mahmoudi, H. Chalian, and B. J. Erickson, "Machine learning in cardiovascular imaging: A scoping review of published literature.," *Current radiology reports*, vol. 11, pp. 34–45, 12 2022.

[54] G. Rathee, M. Balasaraswathi, K. P. Chandran, S. D. Gupta, and C. S. Boopathi, "A secure iot sensors communication in industry 4.0 using blockchain technology," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 533–545, 4 2020.

[55] P. Delva, Z. Altamimi, A. Blazquez, M. Blossfeld, J. Böhm, P. Bonnefond, J.-P. Boy, S. Bruinsma, G. Bury, M. Chatzinikos, A. Couhert, C. Courde, R. Dach, V. Dehant, S. Dell'Agnello, G. Elgered, W. Enderle, P. Exertier, S. Glaser, R. Haas, W. Huang, U. Hugentobler, A. Jäggi, O. Karatekin, F. G. Lemoine, C. L. Poncin-Lafitte, S. Lunz, B. Männel, F. Mercier, L. Métivier, B. Meyssignac, J. Müller, A. Nothnagel, F. Perosanz, R. Rietbroek, M. Rothacher, H. Schuh, H. Sert, K. Sosnica, P. Testani, J. Ventura-Traveset, G. Wautelet, and R. Zajdel, "Genesis: co-location of geodetic techniques in space," *Earth, Planets and Space*, vol. 75, 1 2023.

[56] A. Hussain, M. Aleem, A. Khan, M. A. Iqbal, and M. A. Islam, “Ralba: a computation-aware load balancing scheduler for cloud computing,” *Cluster Computing*, vol. 21, pp. 1667–1680, 3 2018.

[57] M. Qiu, Y. Xiang, and Y. Zhang, “Guest editor’s introduction to the special section on social network security,” *IEEE Transactions on Dependable and Secure Computing*, vol. 15, pp. 549–550, 7 2018.

[58] R. Avula, “Addressing barriers in data collection, transmission, and security to optimize data availability in healthcare systems for improved clinical decision-making and analytics,” *Applied Research in Artificial Intelligence and Cloud Computing*, vol. 4, no. 1, pp. 78–93, 2021.

[59] B. B. Merga, M. B. Moisa, D. A. Negash, Z. Ahmed, and D. O. Gemeda, “Land surface temperature variation in response to land-use and land-cover dynamics: A case of didessa river sub-basin in western ethiopia,” *Earth Systems and Environment*, vol. 6, pp. 803–815, 3 2022.

[60] C. K. Wikle, “Comparison of deep neural networks and deep hierarchical models for spatio-temporal data,” *Journal of Agricultural, Biological and Environmental Statistics*, vol. 24, pp. 175–203, 3 2019.

[61] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, “Automatic visual recommendation for data science and analytics,” in *Advances in Information and Communication: Proceedings of the 2020 Future of Information and Communication Conference (FICC), Volume 2*, pp. 125–132, Springer, 2020.

[62] T. Sakamoto, T. Goto, M. Fujiogi, and A. K. Lefor, “Machine learning in gastrointestinal surgery.,” *Surgery today*, vol. 52, pp. 1–13, 9 2021.

[63] R. Liu, “An edge-based algorithm for tool wear monitoring in repetitive milling processes,” *Journal of Intelligent Manufacturing*, vol. 34, pp. 2333–2343, 3 2022.

[64] P. Pagel, E. Portmann, and K. Vey, “Cognitive computing,” *Informatik-Spektrum*, vol. 41, pp. 1–4, 3 2018.

[65] B.-K. D. Park, J. Park, B. C. Lee, and B. Lee, “Automated system for evaluating consistency between cad model and 3d scan of vehicle seat,” *Journal of Mechanical Science and Technology*, vol. 36, pp. 869–876, 2 2022.

[66] 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.

[67] P. Patel, “Advancements and challenges in data harmonization: A comprehensive review,” *Journal of Artificial Intelligence & Cloud Computing*, pp. 1–5, 6 2022.

[68] S. Arya, K. S. Sandhu, J. Singh, and S. kumar, “Deep learning: as the new frontier in high-throughput plant phenotyping,” *Euphytica*, vol. 218, 3 2022.

[69] S. Felszeghy, S. Pasonen-Seppänen, A. Koskela, P. Nieminen, K. Härkönen, K. M. A. Paldanius, S. Gabbouj, K. Ketola, M. Hiltunen, M. Lundin, T. Haapaniemi, E. Sointu, E. B. Bauman, G. E. Gilbert, D. A. Morton, and A. Mahonen, “Using online game-based platforms to improve student performance and engagement in histology teaching.,” *BMC medical education*, vol. 19, pp. 273–273, 7 2019.

[70] T. Hayajneh, K. N. Griggs, M. Imran, and B. J. Mohd, “Secure and efficient data delivery for fog-assisted wireless body area networks,” *Peer-to-Peer Networking and Applications*, vol. 12, pp. 1289–1307, 1 2019.

[71] P. S. Prabha and S. M. Kumar, “Impacted cyber attacks assessment in wide range of big data security systems,” *International Journal of Recent Technology and Engineering*, vol. 8, pp. 625–629, 10 2019.

[72] K. D. Gupta, D. Dasgupta, and Z. Akhtar, “Determining sequence of image processing technique (ipt) to detect adversarial attacks,” *SN Computer Science*, vol. 2, pp. 1–20, 7 2021.

[73] V. Stephanie, M. A. P. Chamikara, I. Khalil, and M. Atiquzzaman, “Privacy-preserving location data stream clustering on mobile edge computing and cloud,” *Information Systems*, vol. 107, pp. 101728–, 2022.

[74] J. Huang, P. Beling, L. Freeman, and Y. Zeng, “Trustworthy ai for digital engineering transformation,” *Journal of Integrated Design and Process Science*, vol. 25, pp. 1–7, 4 2022.

[75] Y. Chen, “Iot, cloud, big data and ai in interdisciplinary domains,” *Simulation Modelling Practice and Theory*, vol. 102, pp. 102070–, 2020.

[76] I. Reinhartz-Berger, W. Guédria, and P. Bera, “Guest editorial for emmsad’2017 special section,” *Software & Systems Modeling*, vol. 18, pp. 1809–1811, 8 2018.

[77] J. W. Harvey, "Deconstructing sunlight – a community enterprise," *Solar Physics*, vol. 295, pp. 1–37, 6 2020.

[78] W. Chen and S. Srinivasan, "Going digital: implications for firm value and performance," *Review of Accounting Studies*, vol. 29, pp. 1619–1665, 3 2023.

[79] A. K. Saxena and A. Vafin, "Machine learning and big data analytics for fraud detection systems in the united states fintech industry," *Emerging Trends in Machine Intelligence and Big Data*, vol. 11, no. 12, pp. 1–11, 2019.

[80] M. Seyyedattar, S. Zendehboudi, and S. Butt, "Technical and non-technical challenges of development of offshore petroleum reservoirs: Characterization and production," *Natural Resources Research*, vol. 29, pp. 2147–2189, 9 2019.

[81] K. A. Alattas and A. Mardani, "A novel extended internet of things (iot) cybersecurity protection based on adaptive deep learning prediction for industrial manufacturing applications," *Environment, Development and Sustainability*, vol. 24, pp. 9464–9480, 4 2022.

[82] Y. Liu, T. Zhang, X. Wang, G. Yu, and T. Li, "New development of cognitive diagnosis models," *Frontiers of Computer Science*, vol. 17, 7 2022.

[83] D. James, "Anti-development impacts of tax-related provisions in proposed rules on digital trade in the wto," *Development*, vol. 62, pp. 58–65, 9 2019.

[84] X. Zhu, J. Shi, F. Xie, and R. Song, "Pricing strategy and system performance in a cloud-based manufacturing system built on blockchain technology," *Journal of Intelligent Manufacturing*, vol. 31, pp. 1985–2002, 2 2020.

[85] D. J. Crichton, L. Cinquini, H. Kincaid, A. Mahabal, A. Altinok, K. Anton, M. Colbert, S. Kelly, D. Liu, C. Patriotis, S. Lombeyda, and S. Srivastava, "From space to biomedicine: Enabling biomarker data science in the cloud.," *Cancer biomarkers : section A of Disease markers*, vol. 33, pp. 479–488, 4 2022.

[86] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Approximate query processing for big data in heterogeneous databases," in *2020 IEEE international conference on big data (big data)*, pp. 5765–5767, IEEE, 2020.

[87] A. Brighen, H. Slimani, A. Rezgui, and H. Kheddouci, "Listing all maximal cliques in large graphs on vertex-centric model," *The Journal of Supercomputing*, vol. 75, pp. 4918–4946, 2 2019.

[88] T. Nickles, "Alien reasoning: Is a major change in scientific research underway?," *Topoi*, vol. 39, pp. 901–914, 3 2018.

[89] D. H. Blagoeva, P. D. Ørberg Jensen, and H. Merchant, "Services in international business studies: A replication and extension of merchant and gaur (2008)," *Management International Review*, vol. 60, pp. 427–457, 6 2020.

[90] J. Beierle, M. Algorri, M. Cortés, N. S. Cauchon, A. Lennard, J. P. Kirwan, S. Oghamian, and M. J. Abernathy, "Structured content and data management-enhancing acceleration in drug development through efficiency in data exchange.," *AAPS open*, vol. 9, pp. 11–, 5 2023.

[91] A. Farnsworth, Y. T. E. Lo, P. J. Valdes, J. R. Buzan, B. J. W. Mills, A. S. Merdith, C. R. Scotese, and H. R. Wakeford, "Climate extremes likely to drive land mammal extinction during next supercontinent assembly," *Nature Geoscience*, vol. 16, pp. 901–908, 9 2023.

[92] 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.

[93] S. Guha and S. Kumar, "Emergence of big data research in operations management, information systems, and healthcare: Past contributions and future roadmap," *Production and Operations Management*, vol. 27, pp. 1724–1735, 9 2018.

[94] C. Jaeik, N. Chilamkurti, and S.-J. Wang, "Editorial of special section on enabling technologies for industrial and smart sensor internet of things systems," *The Journal of Supercomputing*, vol. 74, pp. 4171–4172, 9 2018.

[95] S. B. Balaji, M. N. Krishnan, M. Vajha, V. Ramkumar, B. Sasidharan, and P. V. Kumar, "Erasure coding for distributed storage: an overview," *Science China Information Sciences*, vol. 61, pp. 100301–, 9 2018.

[96] H. Min and B.-R. Lea, "Motivators and inhibitors for business analytics adoption from the cross-cultural perspectives: A data mining approach," *Information Systems Frontiers*, vol. 26, pp. 1041–1062, 5 2023.

[97] S. Lamichhane, A. Thapa, M. S. Thapa, S. Panthi, and A. J. Giordano, "Understanding human-leopard conflict in the 'mid-hill' region of western nepal," *Journal of Mountain Science*, vol. 20, no. 12, pp. 3464–3475, 2023.

[98] H. M. Makrani, H. Sayadi, N. Nazari, S. M. P. Dinakarrao, A. Sasan, T. Mohsenin, S. Rafatirad, and H. Homayoun, "Adaptive performance modeling of data-intensive workloads for resource provisioning in virtualized environment," *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, vol. 5, pp. 1–24, 12 2020.

[99] E. Ayaburi, M. Maasberg, and J. Lee, "Decision framework for engaging cloud-based big data analytics vendors," *Journal of Cases on Information Technology*, vol. 22, no. 4, pp. 60–74, 2020.

[100] S. K. Tripathi, "Sdn: shape of things to come," *CSI Transactions on ICT*, vol. 8, pp. 3–5, 5 2020.

[101] C. Rong, J. Geng, T. J. Hacker, H. Bryhni, and M. G. Jaatun, "Openiac: open infrastructure as code - the network is my computer," *Journal of Cloud Computing*, vol. 11, 5 2022.

[102] J. D. White, A. Ortega-Castrillon, H. Matthews, A. A. Zaidi, O. Ekrami, J. Snyders, Y. Fan, A. J. Penington, S. V. Dongen, M. D. Shriver, and P. Claes, "Meshmonk: Open-source large-scale intensive 3d phenotyping," *Scientific reports*, vol. 9, pp. 6085–6085, 4 2019.

[103] R. Iacovazzi, Q. Liu, X. Zhou, S. Kireev, N. Sun, and S. peng Ho, "Cosmic-2 soundings impacts on a ro-based noaa microwave satellite data quality monitoring system," *Terrestrial, Atmospheric and Oceanic Sciences*, vol. 33, 3 2022.

[104] S. Roy, S. Ghosh, C. Giri, D. K. Kole, and D. Sarkar, "Mining the human networks and identification of group activities using the crime scraping engine," *SN Computer Science*, vol. 4, 7 2023.

[105] L. Zuo, "Data transfers using bandwidth reservation through multiple disjoint paths of dynamic hpns," *Journal of Network and Systems Management*, vol. 29, pp. 1–24, 3 2021.

[106] T. M. Manderso, Y. A. Mekonnen, T. A. Worku, I. Ahmed, and S. G. Setegn, "Identification of groundwater potential zones using integrated gis-based analytical hierarchy process and multicriteria decision analysis methods in jedeb watershed, ethiopia," *Sustainable Water Resources Management*, vol. 10, 12 2023.

[107] C. Yuan and S. S. Agaian, "A comprehensive review of binary neural network," *Artificial Intelligence Review*, vol. 56, pp. 12949–13013, 3 2023.

[108] P. K. Singh, B. Bhargava, W.-C. Hong, and P. Angin, "1174: futuristic trends and innovations in multimedia systems using big data, iot and cloud technologies (ftims)," *Multimedia Tools and Applications*, vol. 81, pp. 34439–34445, 9 2022.

[109] J. Chen and H. Wang, "Guest editorial: Big data infrastructure ii," *IEEE Transactions on Big Data*, vol. 4, pp. 148–149, 6 2018.

[110] M. K. Hosen, M. S. Alam, T. Chakraborty, and M. R. Golder, "Monitoring spatiotemporal and seasonal variation of agricultural drought in bangladesh using modis-derived vegetation health index," *Journal of Earth System Science*, vol. 132, 11 2023.