

# Multi-Agent AI Systems for Autonomous and Context-Aware Data Orchestration in Hybrid Cloud Platforms

Aditya Gupta<sup>1</sup>, Sai Kiran Oruganti<sup>2</sup>

<sup>1</sup>Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya, Selangor Darul Ehsan, 47301, Malaysia.

<sup>2</sup>Faculty of Engineering and Built Science, Lincoln University College, Petaling Jaya, Selangor Darul Ehsan, 47301, Malaysia.

## Article Info

### Article history:

Received January 08, 2025

Revised February 15, 2026

Accepted February 21, 2026

### Keywords:

Context-aware

Data orchestration

Markov decision process

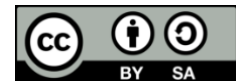
Workload balance

Decision making

## ABSTRACT

Hybrid cloud platforms face challenges in data orchestration due to dynamic resource allocation and workload changes. The framework uses multiple reinforcement learning agents equipped with context-awareness to autonomously manage data orchestration tasks. This investigation aims to develop an artificial intelligence (AI)-based data orchestration model using a Flexible Binary Spider Wasp Algorithm-enriched Double Deep Q-Network with Markov Decision Process (FBSWA-DDQN-MDP) to autonomously manage and optimize data placement, migration, and processing in hybrid cloud platforms. Data is collected from simulated hybrid cloud environments with varying workloads and resource availability. To ensure the dataset is prepared for modeling tasks, it has been preprocessed to eliminate missing values, normalize continuous features, and encode categorical variables. Principal Component Analysis (PCA) was used for feature extraction to improve computational efficiency. Using Python, simulations showed that the FBSWA-DDQN-MDP model outperformed traditional techniques with average energy consumption (AEC) (17.0 J), and average renting charge (ARC) (0.0006 cost) obtained at 1.8 Weight  $\omega_2$ , normalized reward mean and standard deviation (SD) (0.97±0.01) values achieved at 1800 no of training episodes with adaptive response times under dynamic workloads. The proposed multi-agent AI system significantly improves data orchestration in hybrid cloud environments.

This is an open access article under the [CC BY-SA](#) license.



**Corresponding Author:** Aditya Gupta (e-mail: [gupta.aditya56@gmail.com](mailto:gupta.aditya56@gmail.com))

## 1. INTRODUCTION

The architecture integrates cloud, edge, and Internet of Things (IoT) resources to enable seamless and continuous digital service delivery. To explore a unified cloud framework that ensures interoperability among distributed components through semantic integration. Additionally, Artificial Intelligence (AI)-driven processing enhances automation, decision-making, and efficiency across the hybrid infrastructure [1]. Cyber-Physical Systems (CPSs) integrate human interaction, sensing, actuation, and computation to manage real-world processes. Edge Intelligence (EI) reduces latency by giving out IoT-generated data closer to the source. The Osmotic Computing-based Cloud-Edge AI microservice construction for dynamic AI deployment [2]. Multi-Agent Systems (MAS) and Agent-Oriented Programming (AOP) enable autonomous decision-making in Distributed Artificial Intelligence (DAI). To highlight the overlooked real-time limits in MAS, emphasizing the need for agents to act within time limits. To inspect Real-Time Agents (RTA), aligning them with real-time system principles for reliability and reaction [3]. Smart grids (SGs) integrate electrical and digital technologies to optimize energy generation, delivery, and consumption while reducing reliance on fossil fuels. Multi-Agent Systems (MAS) enhance Simulation Graph (SG) security, reliability, and efficiency. Smart agents play a crucial role in modern energy organization and communication networks [4]. Hyper-personalization enhances marketing by tailoring content to user favorites but raises privacy concerns despite boosting engagement. To propose a MAS using Virtual Identities (VIs) to balance personalization and

privacy. The approach improves data transparency and user control while interpreting hyper-personalization patterns [5]. Large Language Models (LLMs) enable innovative language understanding but struggle with complex reasoning and intent alignment. A multi-dimensional classification to enhance knowledge and optimization in LLM-powered multi-agent systems [6].

The aim is to develop an AI-based data orchestration model for hybrid cloud platforms facing dynamic resource allocation and workload challenges. The proposed Flexible Binary Spider Wasp Algorithm-enriched Double Deep Q-Network with Markov Decision Process (FBSWA-DDQN-MDP) framework combines the ability to enable intelligent, autonomous decision-making. To optimize data placement, migration, and processing through multi-agent reinforcement learning, improving resource utilization, reducing delay, and enhancing overall system adaptability.

The research was organized as follows: Section 2 provides relevant literature, Section 3 describes the methodology, Section 4 contains the findings, and Section 5 provides the conclusion.

## 2. LITERATURE REVIEW

The research [7] proposed a dynamic orchestration approach for containerized applications in Edge-Cloud systems to address mobility, network heterogeneity, and resource constraints. A cross-layer technique with context awareness enhances orchestration flexibility. The approach was validated through a proof-of-concept implementation on an embedded testbed. To propose [8] an AI-powered conversational orchestration framework for 6G-based everything as a Service (XaaS) in Industry 5.0, enabling real-time service automation. By integrating cognitive and collaborative AI, the approach enhanced network and cloud service orchestration. Experimental results were effective in improving 6G automation, flexibility, and scalability across Autonomous Vehicle (AV) networks. To lower end-to-end latency, to enhance serverless function provisioning in edge networks [9].

To dynamically distribute computing resources and wireless bandwidth, a context-aware learning framework was suggested. Handling high concurrency and resource heterogeneity presents difficulties, however. Simulations demonstrated a 95% convergence time decrease with significant end-to-end delay. To examine the Cognitive Internet paradigm, emphasizing its advantages, transformative power, and how AI infrastructures support cognitive services [10]. To facilitate the transition, it suggested hybrid edge-cloud (HEC) platforms that prioritize privacy and decision-making autonomy. The difficulties of integrating disparate systems and guaranteeing adaptability were among the limitations. Case studies illustrate the paradigm's potential and practical uses. The Internet of Artificial Intelligence (IoAI) emphasized autonomy as a key factor for error-free AI applications [11]. To enhance daily life, industry, and environmental compatibility while facing challenges in transitioning from human supervision to full autonomy. The findings suggested IoAI gradually evolve from human-controlled systems to fully autonomous settlements.

The Container Deployment Coordination (CODECO) framework optimizes next-gen Internet applications across the Edge-Cloud continuum using decentralized AI for infrastructure selection [12]. To address challenges, like sporadic connectivity and node failures, through dynamic orchestration. Results showed CODECO superiority in real-world application deployments. They proposed a joint orchestration of 5G and Robot Operating System (ROS) to optimize energy usage in robots [13]. Through field experiments, the system offloads computational tasks to the 5G edge and manages sensors to save energy. Results show a 15% reduction in energy consumption, extending battery life, and improving resource utilization. Limitations include the specific testbed setup, which affects generalizability. To overcome problems in workload distribution in centralized cloud systems, investigate Hybrid Edge Cloud (HEC) [14]. By examining traditional and agentic workloads, HEC was able to reduce costs by 80% and save 75% on energy. However, resource limitations in dynamic workloads can affect processing in real time. The platform integrates IoT, cloud, edge computing, and AI for real-time decision-making and predictive analytics [15]. The examination of integration challenges like scalability and interoperability while enhancing low-latency data processing. The findings advance pervasive environmental intelligence for smarter, sustainable infrastructure. To enhance real-time decision-making in Internet of Drones (IoD) using LLMs and RAG, achieving high accuracy and low latency [16].

## 3. METHOD

The methodology starts with simulating hybrid cloud settings featuring dynamic workloads and resource variability. Data was collected to train context-aware reinforcement knowledge agents using the FBSWA-DDQN-MDP model. Data Orchestration is Modern data orchestration combines rule-based management and format transformation. This ensures whole data access, analysis, and automation. While operation varies by use case, a typical data pipeline follows similar stages, as shown in Figure 1.

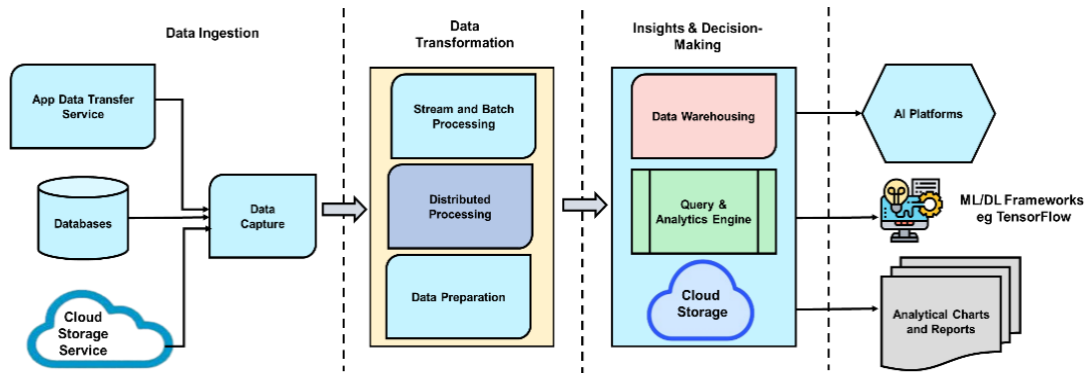


Figure 1. Hybrid Cloud Data Processing and Analytics Workflow.

**Data Ingestion:** The ingestion stage collects and organizes data from various sources, including legacy systems, flat files, and cloud storage, to handle organized and semi-structured data, preparing it for the next pipeline stages. This phase follows specific processes to ensure data readiness.

- Enhancing incoming data by integrating it with existing knowledge
- Assigning relevant labels and adding contextual annotations
- Conducting integrity verification and validation checks for accuracy

**Data transformation:** The transformation stage regulates ingested data from various sources into a consistent format for analysis. Also known as the cleansing stage, it ensures data is registered for internal systems. This phase involves specific developments based on the data type.

- Using different programming languages
- Matching and organizing data
- Removing duplicate data
- Handling tasks step by step in a flow

**Insights and Decision Making:** The decision-making stage utilizes a unified data pool for breakdown through big data or analytics platforms. To apply business logic to derive key visions for users and services. This crucial stage confirms that the data was actionable for decision-making.

- Removing repeated or unnecessary data
- Processing data based on specific needs or scenarios
- Analyzing data and creating reports to find useful insights

### 3.1. Problem formulation

**Data Flow Challenges:** Ensuring that data flows seamlessly and securely across various cloud environments and on-premises systems is a big challenge. **Scalability and Performance:** As data volumes grow, efficient data processing and analysis across hybrid environments need scalable and performant data orchestration technologies. **Security and Compliance:** In hybrid cloud settings, data security and compliance are critical, necessitating powerful data orchestration systems that can enforce security regulations and maintain compliance across many locations. **Lack of Visibility and Control:** Without effective orchestration, organizations may be unable to see their data landscape, making data flow management and control challenging. **Data movement and Lifecycle Management:** Hybrid cloud environments require data movement across multiple cloud environments as well as data lifecycle management (retention, deletion).

### 3.2. Data collection

The data was collected from an open source Kaggle website: <https://www.kaggle.com/datasets/programmer3/cloud-data-orchestration-dataset>. The cloud data orchestration dataset includes 7,365 records that mimic how intelligent agents would behave in a hybrid cloud setting. Using reinforcement learning techniques, it was intended to simulate autonomous, context-aware decision-making for activities such as data migration, replication, and deletion. Every record in the collection contains: Timestamps with regular intervals, System metrics, such as CPU and memory use, network bandwidth, and latency. Agent actions: Workload levels: low, medium, and high. Environmental context (peak time flag, source and destination nodes). Metrics derived include data transmission time, resource utilization score, and computed rewards. The dataset has been preprocessed to remove missing values, normalize continuous features, and encode categorical variables, making it suitable for modeling tasks. Principal Component Analysis (PCA) was used

to minimize dimensionality while conserving variance across system parameters, resulting in improved computing efficiency and model performance.

### 3.3. Multi-Agent AI system

A multi-agent AI system uses multiple intelligent agents to manage data orchestration tasks collaboratively. The proposed methodology integrated FBSWA for optimization and DDQN with MDP for intelligent decision-making. FBSWA-DDQN-MDP enables adaptive learning and dynamic resource allocation, improving delay, delivery, and system performance.

### 3.4. Flexible Binary Spider Wasp Algorithm-enriched Double Deep Q-Network

To improve data orchestration on hybrid cloud platforms, the FBSWA-DDQN-MDP optimizes the handling of data, migration, and placement. While DDQN minimizes overappraisal bias and confirms adaptive decision-making, FBSWA effectively chooses the best course of action. Reinforcement learning (DDQN) was used to control capacity in real time while modeling data orchestration as a Markov Decision Process (MDP), and was also used in the decision-making process.

### 3.5. Double Deep Q-Network (DDQN)

DDQN uses a single max estimator for action selection and evaluation, leading to overoptimistic action values. Double Q-Learning addresses this by decoupling selection and estimation into two estimators, reducing overestimation bias. In DDQN, the target network updates its limits periodically for stable learning. To approach results in reasonable action value estimation and improved dynamic allocation of resources.

$$Z(s) = H(s) + \mu q(t', \operatorname{argmax}_q(t', b^*; \theta(s)); \theta'(s)) \quad (1)$$

$$Z(s) = H(s) + (t', \operatorname{argmax}_q(t', b^*; \theta_j); \theta_{j-1}) \quad (2)$$

Where  $Z(s)$  represents the optimized data orchestration value,  $H(s)$  is the immediate system reward, and  $\mu$  controls update intensity.  $(t')$  models' capacity impact  $b^*$  was the optimal action, and  $\theta(s), \theta'(s)$ , were policy and aim network parameters for stable learning by equations (1) and (2).

### 3.6. Flexible Binary Spider Wasp Algorithm (FBSWA)

The Binary Spider Wasp Algorithm (BSWA) was a bio-inspired optimization technique based on the hunting strategy of spider wasps. BSWA operates in binary space to effectively solve feature selection and classification problems. BSWA balances exploration and exploitation, ensuring accurate and efficient optimization in complex environments. The FBSWA was introduced to enhance resource allocation in a hybrid cloud. It operates in two phases: first, binary improvement was applied, followed by the integration of Spider Wasp Optimization (SWO) with a Genetic Algorithm (GA). A flat crossover operator replaces the traditional GA crossover to improve the selection of optimal features while enhancing classification accuracy. Since FBSWA generates continuous values, a two-step transfer function is used to convert them into binary values, with S1 being the most effective among the tested functions. The fitness function was designed to balance accuracy and feature reduction, ensuring optimal classification with minimal feature selection. By employing these improvements, FBSWO achieves higher detection accuracy and efficient feature selection in hybrid cloud security with data orchestration in cloud platforms.

$S(N_j^i(s))$  represents the success probability of data delivery, where  $f$  was a network factor, and  $i, j, s$  denote node indices and system state, optimizing decision making and delivery by equation (3).

$$S(N_j^i(s)) = 1/(1 + f^{-2}) \quad (3)$$

$E(t)$  denotes data delivery efficiency,  $\alpha, \beta$ , where weights were a quality factor, and  $\frac{|Q|}{|M|}$  balances line up data with memory to enhance delay to the hybrid cloud and delivery by equation (4).

$$E(t) = \alpha * FQQ + \beta * \frac{|Q|}{|M|} \quad (4)$$

$E(s)$  optimizes dynamically allocated resources in hybrid cloud platforms by maximizing accuracy, where  $ACC(t)$  accuracy  $T_e$  was a tuning factor, and  $\frac{K_e}{K_s}$  balances explored and selected structures to reduce delay and enhance data orchestration using the FBSWA model, as shown in equation (5).

$$E(s) = \text{maximize} \left( ACC(t) + T_e * \left( 1 - \frac{K_e}{K_s} \right) \right) \quad (5)$$

### 3.7. Markov Decision Process (MDP)

An MDP was a mathematical framework for decision-making in dynamic environments. It is defined by  $(S, A, R, P, \gamma)$ , where  $S$  represents states,  $A$  represents actions,  $R$  is the reward function,  $P$  is the transition probability, and  $\gamma$  is the discount factor. In the proposed system, MDP enables optimal decision-making for dynamic offloading by learning the best actions to maximize rewards over time.

**State Space:** The state space was designed to represent large datasets, Central Processing Unit (CPU) capacity, and network delays, influencing processing time. To enable informed decision-making for optimizing system performance at each time interval. These variables define the system's state at the start of each condition, crucial for decision-making in dynamic offloading. T- Time stage for event tracking, Data Traffic (DT)- Incoming file size affecting processing load, Utilization Time (UT)- CPU capacity determining computational power; Bandwidth Time (BT) – Network delay impacting communication efficiency.

**Action Space:** The action space defines decision-making for executing applications locally ( $A = 0$ ) or offloading them remotely ( $A = 1$ ). To enable optimal task allocation based on system conditions.

**Rewards:** The reward function Response Time (RT) evaluates the efficiency of local vs cloud processing based on computing cost. The system optimizes decision-making by selecting between local execution ( $C_{cloud\_T}$ ) and cloud offloading ( $C_{cloud\_T}$ ). The reward was calculated in equation (6).

$$QS = - \frac{(C_{algorithms} - C_{locals})}{(C_{locals} + C_{locals})} \quad (6)$$

Where  $QS$  quantifies the performance improvement,  $C_{algorithms}$  represents the execution cost using the proposed orchestration algorithm, and  $C_{locals}$  depicts the execution cost of non-optimized processing. To improve the classification in the cloud, FBSWA combined binary improvement, SWO, and a GA with Flat crossover. The design improved delay, data delivery, and orchestration while dynamically allocating resources when combined with DDQN and MDP. In hybrid cloud platforms, adaptive decision-making, effective resource management, and high detection accuracy were guaranteed by custom fitness and reward mechanisms, as shown by algorithm 1.

#### Algorithm 1: FBSWA-DDQN-MDP

Initialization:

Initialize  $\theta, \theta'$  → policy and target network parameters  
 Initialize replay memory  $D \leftarrow \emptyset$   
 Initialize spider wasp population  $P \leftarrow \{W1, W2, \dots, Wn\}$   
 Define state space  $S = \{T, DT, UT, BT\}$   
 Define action space  $A = \{0: \text{local}, 1: \text{offload}\}$   
 Set discount factor  $\gamma \in [0,1]$   
 Set  $\mu \in R$   
 Set learning rate  $\alpha$

For each episode  $e = 1$  to  $E$  do:

Initialize state  $s_0 \leftarrow \{T_0, DT_0, UT_0, BT_0\}$

For each timestep  $t = 1$  to  $T$  do:

With probability  $\epsilon$ :

Select a random action  $a_t \in A$

Else:

$a_t \leftarrow \text{argmax}_a Q(st, a; \theta)$

Observe reward  $R_t$

Observe next state  $st+1$

Store transition:  
 $D \leftarrow D \cup (st, at, Rt, st+1)$

Sample minibatch:  
 $B \leftarrow \{(s, a, r, s')\} \subset D$

For each  $(s, a, r, s')$  in  $B$  do:  
 $b^* \leftarrow \operatorname{argmax}_a Q(s', a; \theta)$   
 $Z(s) \leftarrow H(s) + \mu * Q(s', b^*; \theta')$

Update  $\theta$  by minimizing loss:  
 $L = E[(Z(s) - Q(s, a; \theta))^2]$

If  $t \bmod \tau == 0$  then:  
 $\theta' \leftarrow \theta$

Update state:  
 $st \leftarrow st+1$

End for

For each spider wasp  $W_i \in P$  do:  
 Evaluate fitness:  
 $F(W_i) = \alpha * FQQ + \beta * (|Q| / |M|)$

Apply Binary Improvement Phase:  
 $W_i \leftarrow S1(FBSWA(W_i)) \rightarrow$  binary value transformation

Apply Flat Crossover and Genetic Update:  
 $Crossover(W_i, W_j) \rightarrow W_{new}$   
 $Mutation(W_{new})$

Select elite individuals based on:  
 $E(s) = \operatorname{maximize}(ACC(t) + T_e * (1 - K_e / K_s))$

Update spider wasp population:  
 $P \leftarrow \operatorname{top}_k(W_1, W_2, \dots, W_n)$

End for

Return:  
 Optimized Q-policy  $\theta^*$   
 Optimized spider wasp population  $P^*$   
 Delivery success:  
 $S(N_j^i(s)) = 1 / (1 + f^{(-2)})$

#### 4. EXPERIMENTAL RESULTS

This section explains the effectiveness of the proposed FBSWA-DDQN-MDP method with the help of system configuration and comparison phase with average energy efficiency, renting charge, and normalized reward. Table 1 displays the experimental setup of the proposed technique.

Table 1. Experimental setup

Parameter	Details
Programming Language	Python 3.8
DL Framework	TensorFlow 2.5
Simulation environment	OpenAI Gym
Hardware	16 GB RAM, NVIDIA GTX 1080 GPU

Average energy consumption (Joules (J)) was the measure of electrical energy used by a system to perform specific tasks; in this case, it reflects the energy required for data placement, migration, and processing within hybrid cloud environments. Average renting charge (cost) refers to the cost incurred for using cloud infrastructure resources, such as storage, processing power, and bandwidth, over time in a data orchestration. Normalized reward is a metric in reinforcement learning that reflects how effectively an agent achieves its objectives compared to the best possible outcome. A value close to the ideal indicates that the model makes consistently accurate and efficient decisions. The objective is to improve data orchestration, reduce delay, and enhance data delivery in hybrid cloud environments.

The comparison between the proposed FBSWA-DDQN-MDP model and the cooperative multi-agent deep reinforcement learning (CMADRL) based framework [17] with average energy consumption, renting charge, and normalized reward. The existing model utilizes the cooperative twin delayed DDPG (CMATD3) method.

Table 2 and Figure 2 show the numerical values and pictorial representation for average energy consumption between CMATD3 and FBSWA-DDQN-MDP. The proposed FBSWA-DDQN-MDP model achieves lower average energy consumption (17.0 J) than CMATD3 (19J) at 1.8 Weight  $\omega_2$ . The FBSWA-DDQN model was effective at maximizing resource use in hybrid cloud systems, as seen by its persistent decrease in energy consumption.

Table 2. CMATD3 vs. FBSWA-DDQN for average energy consumption (Joule (J))

Weight $\omega_2$	CMATD3[17]	FBSWA-DDQN-MDP [Proposed]
0.2	6	5.5
0.4	8	7.5
0.6	10	9.2
0.8	12	11.0
1.0	14	12.8
1.2	16	14.3
1.4	17	15.2
1.6	18	16.0
1.8	19	17.0

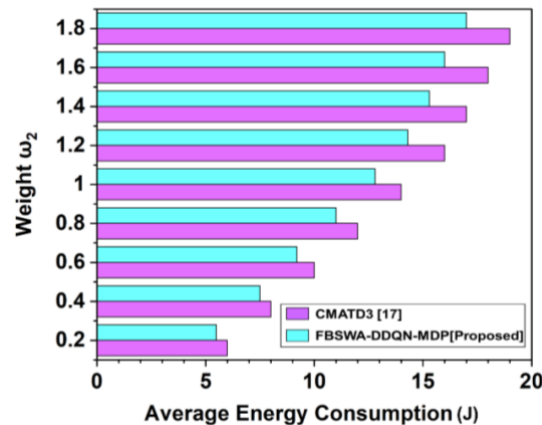


Figure 2. Graphical representation of average energy consumption.

Figure 3 and Table 3 demonstrate the outcome of the average renting charge in a hybrid cloud system based on Weight  $\omega_2$ . The proposed FBSWA-DDQN-MDP obtains a lower average renting charge (0.0006 cost) at 1.8 Weight  $\omega_2$ . The FBSWA-DDQN model consistently incurs lower renting costs, highlighting its cost-effectiveness in hybrid cloud resource management.

Table 4 and Figure 4 illustrate the normalized reward values between the proposed FBSWA-DDQN-MDP and CMATD3 based on the number of training episodes. The FBSWA-DDQN-MDP achieves higher mean and lower standard deviation values ( $0.97 \pm 0.01$ ) than CMATD3 ( $0.95 \pm 0.01$ ) at 1800 training episodes.

Table 3. Performance Comparison of the average renting charge (cost)

Weight $\omega_2$	CMATD3[17]	FBSWA-DDQN-MDP [Proposed]
0.2	0.0038	0.0035
0.4	0.0030	0.0028
0.6	0.0022	0.0020
0.8	0.0016	0.0015
1.0	0.0012	0.0011
1.2	0.0010	0.0009
1.4	0.0009	0.0008
1.6	0.0008	0.0007
1.8	0.0007	0.0006

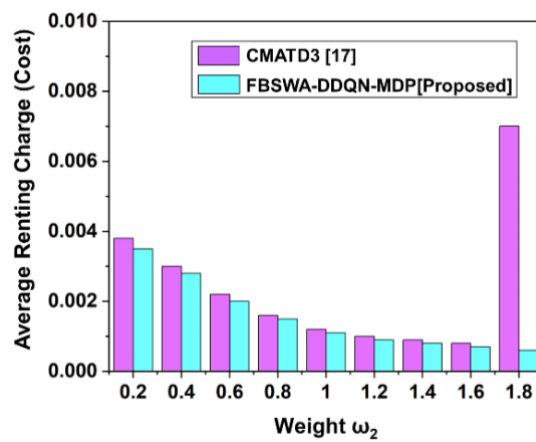
Figure 3. Average Renting Charge based on weight ( $\omega_2$ )

Table 4. Normalized Reward values for CMATD3 vs. FBSWA-DDQN.

No. of Training Episodes	CMATD3[17]	FBSWA-DDQN-MDP [proposed]
200	0.80 $\pm$ 0.07	0.82 $\pm$ 0.06
600	0.86 $\pm$ 0.05	0.88 $\pm$ 0.04
1000	0.90 $\pm$ 0.03	0.92 $\pm$ 0.02
1400	0.93 $\pm$ 0.01	0.95 $\pm$ 0.01
1800	0.95 $\pm$ 0.01	0.97 $\pm$ 0.01

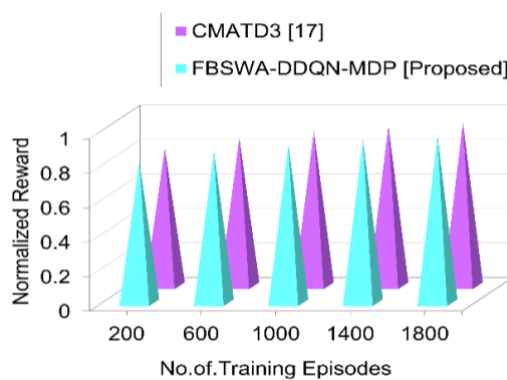


Figure 4. Graphical representation of normalized reward.

The proposed FBSWA-DDQN-MDP model efficiently manages dynamic data orchestration in hybrid cloud environments using reinforcement learning agents. The agents autonomously handle tasks like data assignment and migration by learning from structure states and rewards through an MDP framework. The traditional orchestration method CMATD3 [17] frequently falls short in dynamic settings due to restricted adaptability and a lack of intelligent resource management. The proposed system incorporates multiple reinforcement learning agents equipped with context-awareness, enabling autonomous decision-making for data placement, migration, and processing to overcome the limits. The integration of the FBSWA-DDQN-MDP enhances decision-making and exploration. Simulation results demonstrate a reduction in data transfer time and an increase in resource utilization. Overall, the model shows strong adaptability and performance optimization under varying workloads.

## 5. CONCLUSION

An AI-driven data orchestration framework was implemented to improve efficiency in hybrid cloud environments utilizing an FBSWA-DDQN-MDP. The cloud data orchestration dataset consists of 7,365 records that represent how intelligent agents will act in a hybrid cloud environment. Using reinforcement learning techniques, it was meant to imitate autonomous, context-aware decision-making for tasks such as data migration, replication, and deletion. To prepare it for modeling tasks, the dataset has undergone preprocessing that includes encoding categorical variables, normalizing continuous features, and removing missing values. PCA was used to minimize dimensionality and preserve variance across system metrics for feature extraction, improving model performance and computing efficiency. The proposed FBSWA-DDQN-MDP model demonstrated strong performance in hybrid cloud orchestration, achieving reduced average energy consumption (17.0 J) and renting charges (0.0006 cost) at 1.8 weight  $\omega_2$ , and a high mean and lower SD value of  $(0.97 \pm 0.01)$  at 1800 training episodes, outperforming the existing CMATD3 method. While the model shows outstanding adaptability and efficiency in simulated environments, its real-time deployment may face challenges under unpredictable conditions and extreme workloads. While the current approach assumes ideal system conditions, future enhancements would comprise fault tolerance and real-time adaptability across heterogeneous cloud infrastructures to further improve scalability and resilience.

## DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are available in the Kaggle repository, <https://www.kaggle.com/datasets/programmer3/cloud-data-orchestration-dataset>

## CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest to this work.

## REFERENCES

- [1] S. Alberternst *et al.*, "Orchestrating Heterogeneous Devices and AI Services as Virtual Sensors for Secure Cloud-Based IoT Applications," *Sensors*, vol. 21, no. 22, p. 7509, Nov. 2021, <https://doi.org/10.3390/s21227509>.
- [2] P. Trakadas *et al.*, "A Reference Architecture for Cloud-Edge Meta-Operating Systems Enabling Cross-Domain, Data-Intensive, ML-Assisted Applications: Architectural Overview and Key Concepts," *Sensors*, vol. 22, no. 22, p. 9003, Nov. 2022, <https://doi.org/10.3390/s22229003>.
- [3] D. Calvaresi, Y. Dicente Cid, M. Marinoni, A. F. Dragoni, A. Najjar, and M. Schumacher, "Real-time multi-agent systems: rationality, formal model, and empirical results," *Autonomous Agents and Multi-Agent Systems*, vol. 35, no. 1, p. 12, Apr. 2021, <https://doi.org/10.1007/s10458-020-09492-5>.
- [4] S. S. Binyamin and S. Ben Slama, "Multi-Agent Systems for Resource Allocation and Scheduling in a Smart Grid," *Sensors*, vol. 22, no. 21, p. 8099, Oct. 2022, <https://doi.org/10.3390/s22218099>.
- [5] S. Darshan, "Multi-Agent Orchestration for Autonomous Data Pipelines: A Systems Architecture for Self-Healing, Context-Aware, and Resilient Data Processing," *International Journal of Emerging Research in Engineering and Technology*, vol. 7, no. 1, pp. 47–53, 2026, <https://doi.org/10.63282/3050-922X.IJERET-V7I1P107>.
- [6] H. Chen, "Intelligent Agent-Based Market Research: Cloud-Orchestrated Large Language Models as Financial Analysts," *Cloud Computing and Data Science*, pp. 161–168, Jan. 2026, <https://doi.org/10.37256/ccds.7120269128>.
- [7] R. C. Sofia, D. Dykeman, P. Urbanetz, A. Galal, and D. A. Dave, "Dynamic, Context-Aware Cross-Layer Orchestration of Containerized Applications," *IEEE Access*, vol. 11, pp. 93129–93150, 2023, <https://doi.org/10.1109/ACCESS.2023.3307026>.
- [8] S. Shahzadi, N. R. Chaudhry, and M. Iqbal, "A Novel 6G Conversational Orchestration Framework for Enhancing Performance and Resource Utilization in Autonomous Vehicle Networks," *Sensors*, vol. 23, no. 17, p. 7366, Aug. 2023, <https://doi.org/10.3390/s23177366>.
- [9] F. F. Salerno and A. C. G. Maçada, "Data orchestration as an emerging phenomenon: a systematic literature review on its intersections with data governance and data strategy," *Management Review Quarterly*, Oct. 2025,

- <https://doi.org/10.1007/s11301-025-00558-w>.
- [10] M. Weber, A. Hein, J. Weking, and H. Krcmar, "Orchestration logics for artificial intelligence platforms: From raw data to industry-specific applications," *Information Systems Journal*, vol. 35, no. 3, pp. 1015–1043, May 2025, <https://doi.org/10.1111/isj.12567>.
- [11] S. Banaeian Far and A. Imani Rad, "Internet of Artificial Intelligence (IoAI): the emergence of an autonomous, generative, and fully human-disconnected community," *Discover Applied Sciences*, vol. 6, no. 3, p. 91, Feb. 2024, <https://doi.org/10.1007/s42452-024-05726-3>.
- [12] M. Shahin *et al.*, "DatApollo: Orchestration of Serverless Functions for Scalable Data Mining," *IEEE Access*, vol. 13, pp. 142813–142828, 2025, <https://doi.org/10.1109/ACCESS.2025.3591712>.
- [13] M. Groshev, L. Zanzi, C. Delgado, X. Li, A. de la Oliva, and X. Costa-Pérez, "Energy-Aware Joint Orchestration of 5G and Robots: Experimental Testbed and Field Validation," *IEEE Transactions on Network and Service Management*, vol. 22, no. 4, pp. 3046–3059, Aug. 2025, <https://doi.org/10.1109/TNSM.2025.3555126>.
- [14] S. Limkar, M. A. Hossain, S. T. Amin, and Y. Ahmad, "Dynamic Resource Orchestration for Computing, Data, and IoT in Networked Systems: A Data-Centric Approach," in *Current and Future Cellular Systems*, Wiley, 2025, pp. 185–207, <https://doi.org/10.1002/9781394256075.ch10>
- [15] I. Ficili, M. Giacobbe, G. Tricomi, and A. Puliafito, "From Sensors to Data Intelligence: Leveraging IoT, Cloud, and Edge Computing with AI," *Sensors*, vol. 25, no. 6, p. 1763, Mar. 2025, <https://doi.org/10.3390/s25061763>.
- [16] A. Sezgin, "Scenario-Driven Evaluation of Autonomous Agents: Integrating Large Language Model for UAV Mission Reliability," *Drones*, vol. 9, no. 3, p. 213, Mar. 2025, <https://doi.org/10.3390/drones9030213>.
- [17] J. Chen, P. Chen, X. Niu, Z. Wu, L. Xiong, and C. Shi, "Task offloading in hybrid-decision-based multi-cloud computing network: a cooperative multi-agent deep reinforcement learning," *Journal of Cloud Computing*, vol. 11, no. 1, p. 90, Dec. 2022, <https://doi.org/10.1186/s13677-022-00372-9>.

## BIOGRAPHIES OF AUTHORS



**Aditya Gupta** is working as a Research Associate in Computer Science and Engineering at Lincoln University College, Malaysia, and as a Senior Software Developer in a reputed organization in Lucknow, UP, India. He can be contacted at email: [gupta.aditya56@gmail.com](mailto:gupta.aditya56@gmail.com)



**Sai Kiran Oruganti** is a professor specializing in wireless power transfer and IoT security, currently affiliated with Jiangxi University of Science and Technology in China and Lincoln University College in Malaysia. His career includes research at the Ulsan National Institute of Science and Technology, developing wireless systems for Hyundai and Samsung, a faculty position at the Indian Institute of Technology, and a recognized history of innovation, including pioneering work on Zenneck Wave WPT and over 16 granted patents. He can be contacted at email: [saisharma@lincoln.edu.my](mailto:saisharma@lincoln.edu.my)