REFERENCES

[1] P. M. Mell and T. Grance, *The NIST definition of cloud computing*. National Institute of Standards and Technology, 2011.

[2] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *Journal of internet services and applications*, vol. 1, pp. 7–18, 2010.

[3] A. Jayanetti, S. Halgamuge, and R. Buyya, "Multi-agent deep reinforcement learning framework for renewable energy-aware workflow scheduling on distributed cloud data centers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 35, no. 4, pp. 604–615, 2024.

[4] R. J. Sethi and Y. Gil, "Scientific workflows in data analysis: Bridging expertise across multiple domains," *Future Generation Computer Systems*, vol. 75, pp. 256–270, 2017.

[5] S. Cheikh and J. Walker, "Solving task scheduling problem in the cloud using a hybrid particle swarm optimization approach," *International Journal of Applied Metaheuristic Computing*, vol. 13, pp. 1–25, 2022.

[6] H. Topcuoglu, S. Hariri, and M.-Y. Wu, "Performance-effective and low-complexity task scheduling for heterogeneous computing," *IEEE transactions on parallel and distributed systems*, vol. 13, no. 3, pp. 260–274, 2002.

[7] J. Malmodin, N. Lövehagen, P. Bergmark, and D. Lundén, "Ict sector electricity consumption and greenhouse gas emissions - 2020 outcome," *Telecommunications Policy*, vol. 48, no. 3, p. 102701, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0308596123002124

[8] G. Zhou, W. Tian, R. Buyya, R. Xue, and L. Song, "Deep reinforcement learning-based methods for resource scheduling in cloud computing: a review and future directions," *Artificial Intelligence Review*, vol. 57, no. 5, 2024.

[9] C. F. Hayes, R. Rădulescu, E. Bargiacchi, J. Källström, M. Macfarlane, M. Reymond, T. Verstraeten, L. M. Zintgraf, R. Dazeley, F. Heintz *et al.*, "A practical guide to multi-objective reinforcement learning and planning," *Autonomous Agents and Multi-Agent Systems*, vol. 36, no. 1, p. 26, 2022.

[10] McKinsey & Company, "How data centers and the energy sector can sate ai's hunger for power," 2024. [Online]. Available: https://www.mckinsey.com/industries/private-capital/our-insights/how-data-centers-and-the-energy-sector-can-sate-ais-hunger-for-power

[11] IEA, "Electricity 2024 - analysis and forecast to 2026," https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf, 2024, accessed: 2025-05-17.

[12] S. Di, D. Kondo, and F. Cappello, "Characterizing and modeling cloud applications/jobs on a google data center," *The Journal of Supercomputing*, vol. 69, no. 1, pp. 139–160, 2014.

[13] L. Huang, J. Jia, B. Yu, B.-G. Chun, P. Maniatis, and M. Naik, "Predicting execution time of computer programs using sparse polynomial regression," *Advances in neural information processing systems*, vol. 23, 2010.

[14] E. H. Houssein, A. G. Gad, Y. M. Wazery, and P. N. Suganthan, "Task scheduling in cloud computing based on meta-heuristics: Review, taxonomy, open challenges, and future trends," *Swarm and Evolutionary Computation*, vol. 62, p. 100841, 2021.

[15] Z. Tong, H. Chen, X. Deng, K. Li, and K. Li, "A scheduling scheme in the cloud computing environment using deep q-learning," *Information Sciences*, vol. 512, pp. 1170–1191, 2020.

[16] M. Sudheer, K. Reddy, M. KUMAR, O. Khalaf, C. Romero, and G. Sahib, "DRLBTSA: Deep reinforcement learning based task-scheduling algorithm in cloud computing," *Multimedia Tools and Applications*, vol. 83, pp. 1–29, 2023.

[17] "Apache airflow documentation," 2024. [Online]. Available: https://airflow.apache.org/

[18] "Argo workflows documentation," 2024. [Online]. Available: https://argoproj.github.io/workflows/

[19] R. N. Calheiros, R. Ranjan, A. Beloglazov, C. A. F. De Rose, and R. Buyya, "Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms," *Software: Practice and Experience*, vol. 41, no. 1, pp. 23–50, 2011.

[20] J. J. Durillo, H. M. Fard, and R. Prodan, "Moheft: A multi-objective list-based method for workflow scheduling," in *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, 2012, pp. 185–192.

[21] X. Zhou, G. Zhang, J. Sun, J. Zhou, T. Wei, and S. Hu, "Minimizing cost and makespan for workflow scheduling in cloud using fuzzy dominance sort based heft," *Future Generation Computer Systems*, vol. 93, pp. 278–289, 2019.

[22] L. Xing, J. Li, Z. Cai, and F. Hou, “Evolutionary optimization of energy consumption and makespan of workflow execution in clouds,” *Mathematics*, vol. 11, no. 9, p. 2126, 2023.

[23] H. Mao, M. Alizadeh, I. Menache, and S. Kandula, “Resource management with deep reinforcement learning,” in *Proceedings of the 15th ACM workshop on hot topics in networks*, 2016, pp. 50–56.

[24] Y. Yang, C. He, B. Yin, Z. Wei, and B. Hong, “Cloud task scheduling based on proximal policy optimization algorithm for lowering energy consumption of data center,” *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 16, no. 6, pp. 1877–1891, 2022.

[25] K. Lei, P. Guo, Y. Wang, J. Xiong, and W. Zhao, “An end-to-end hierarchical reinforcement learning framework for large-scale dynamic flexible job-shop scheduling problem,” in *2022 International Joint Conference on Neural Networks (IJCNN)*, Padua,Italy, 2022, pp. 1–8.

[26] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, “How powerful are graph neural networks?” *arXiv preprint arXiv:1810.00826*, 2018.

[27] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, “Graph attention networks,” *arXiv preprint arXiv:1710.10903*, 2017.

[28] J. Wilkes, “Google cluster-usage traces v3,” https://github.com/google/cluster-data/blob/master/ClusterData2019.md, 2020, accessed: 2025-05-17.

[29] Alibaba, “Alibaba cluster trace program - cluster data v2018,” https://github.com/alibaba/clusterdata, 2018, accessed: 2025-05-17.

[30] C. Audet, J. Bigeon, D. Cartier, S. Le Digabel, and L. Salomon, “Performance indicators in multiobjective optimization,” *European journal of operational research*, vol. 292, no. 2, pp. 397–422, 2021.

[31] E. Zitzler, K. Deb, and L. Thiele, “Comparison of multiobjective evolutionary algorithms: Empirical results,” *Evolutionary computation*, vol. 8, no. 2, pp. 173–195, 2000.

[32] W. Liu, Y. Yan, Y. Sun, H. Mao, M. Cheng, P. Wang, and Z. Ding, “Online job scheduling scheme for low-carbon data center operation: An information and energy nexus perspective,” *Applied Energy*, vol. 338, p. 120918, 2023.

[33] K. Lei, P. Guo, Y. Wang, J. Zhang, X. Meng, and L. Qian, “Large-scale dynamic scheduling for flexible job-shop with random arrivals of new jobs by hierarchical

reinforcement learning," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 1, pp. 1007–1018, 2024, conference Name: IEEE Transactions on Industrial Informatics. [Online]. Available: https://ieeexplore.ieee.org/document/10114974

[34] J. Zhou, T. Wang, P. Cong, P. Lu, T. Wei, and M. Chen, "Cost and makespan-aware workflow scheduling in hybrid clouds," *Journal of Systems Architecture*, vol. 100, p. 101631, 2019.

[35] K. M. U. Ahmed, M. H. J. Bollen, and M. Alvarez, "A review of data centers energy consumption and reliability modeling," *IEEE Access*, vol. 9, pp. 152 536–152 563, 2021.

[36] E. Deelman, K. Vahi, G. Juve, M. Rynge, S. Callaghan, P. J. Maechling, R. Mayani, W. Chen, R. Ferreira da Silva, M. Livny, and K. Wenger, "Pegasus, a workflow management system for science automation," *Future Generation Computer Systems*, vol. 46, pp. 17–35, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X14002015

[37] F. Cao and M. M. Zhu, "Energy-aware workflow job scheduling for green clouds," in *2013 IEEE international conference on green computing and communications and IEEE internet of things and IEEE cyber, physical and social computing*. IEEE, 2013, pp. 232–239.

[38] J. J. Durillo, V. Nae, and R. Prodan, "Multi-objective energy-efficient workflow scheduling using list-based heuristics," *Future Generation Computer Systems*, vol. 36, pp. 221–236, 2014.

[39] R. Shaw, E. Howley, and E. Barrett, "An advanced reinforcement learning approach for energy-aware virtual machine consolidation in cloud data centers," in *2017 12th International Conference for Internet Technology and Secured Transactions (ICITST)*. IEEE, 2017, pp. 61–66.

[40] S. S. Mangalampalli, G. R. Karri, P. R. Ch, K. S. Pokkuluri, P. Chakrabarti, and T. Chakrabarti, "An energy and temperature aware deep reinforcement learning workflow scheduler in cloud computing," *IEEE Access*, vol. 12, pp. 163 424–163 443, 2024.

[41] M. Reymond and A. Nowe, "Pareto-dqn: Approximating the pareto front in complex multi-objective decision problems," in *Proceedings of the Adaptive and Learning Agents Workshop 2019 (ALA-19) at AAMAS*, 2019, pp. 1–6.

[42] W. Zhang and H. Ou, "Reinforcement learning based multi objective task scheduling for energy efficient and cost effective cloud edge computing," *Scientific Reports*, vol. 15, no. 1, 2025.

[43] A. Bartlett, C. Liem, and A. Panichella, "The pursuit of diversity: Multi-objective testing of deep reinforcement learning agents," *arXiv preprint arXiv:2510.14727*, 2025.

[44] Y.-H. Hung, "Investigating how the cloud computing transforms the development of industries," *IEEE Access*, vol. 7, pp. 181 505–181 517, 2019.

[45] M. Ghobakhloo, "The future of manufacturing industry: a strategic roadmap toward industry 4.0," *Journal of manufacturing technology management*, vol. 29, no. 6, pp. 910–936, 2018.

[46] MarketsandMarkets, "Cloud computing market size, size, growth & latest trends," https://www.marketsandmarkets.com/Market-Reports/cloud-computing-market-234.html, 2025, accessed: 2025-05-17.

[47] SPEC, "SPECpower_ssj 2008 results," https://www.spec.org/power_ssj2008/results/power_ssj2008/, 2008, accessed: 2025-05-17.

[48] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing," *Future generation computer systems*, vol. 28, no. 5, pp. 755–768, 2012.

[49] X. Li, W. Yu, R. Ruiz, and J. Zhu, "Energy-aware cloud workflow applications scheduling with geo-distributed data," *IEEE Transactions on Services Computing*, vol. 15, no. 2, pp. 891–903, 2020.

[50] M. Alouane and H. El Bakkali, "Virtualization in cloud computing: Existing solutions and new approach," in *2016 2nd International Conference on Cloud Computing Technologies and Applications (CloudTech)*, 2016, pp. 116–123.

[51] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauly, M. J. Franklin, S. Shenker, and I. Stoica, "Resilient distributed datasets: A {Fault-Tolerant} abstraction for {In-Memory} cluster computing," in *9th USENIX symposium on networked systems design and implementation (NSDI 12)*, 2012, pp. 15–28.

[52] A. Jayanetti, S. Halgamuge, and R. Buyya, "Deep reinforcement learning for energy and time optimized scheduling of precedence-constrained tasks in edge–cloud computing environments," *Future Generation Computer Systems*, vol. 137, pp. 14–30, 2022.

[53] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 2012, pp. 13–16.

[54] D. Kebbal, E. Talbi, and J. Geib, "Building and scheduling parallel adaptive applications in heterogeneous environments," in *ICWC 99. IEEE Computer Society International Workshop on Cluster Computing*. IEEE, 1999, pp. 195–201.

[55] D. Meedeniya, *Deep Learning: A Beginners' Guide*. CRC Press LLC, 2023. [Online]. Available: www.routledge.com/9781032473246

[56] K. Lei, P. Guo, W. Zhao, Y. Wang, L. Qian, X. Meng, and L. Tang, "A multi-action deep reinforcement learning framework for flexible job-shop scheduling problem," *Expert Systems with Applications*, vol. 205, p. 117796, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417422010624

[57] Y.-H. Chang, C.-H. Liu, and S. D. You, "Scheduling for the flexible job-shop problem with a dynamic number of machines using deep reinforcement learning," *Information*, vol. 15, no. 2, p. 82, 2024, publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2078-2489/15/2/82

[58] G. Andreadis, L. Versluis, F. Mastenbroek, and A. Iosup, "A reference architecture for datacenter scheduling: Extended technical report," *arXiv preprint arXiv:1808.04224*, 2018.

[59] S. Euting, C. Janiesch, R. Fischer, S. Tai, and I. Weber, "Scalable business process execution in the cloud," in *2014 IEEE International Conference on Cloud Engineering*. IEEE, 2014, pp. 175–184.

[60] L. Thai, B. Varghese, and A. Barker, "A survey and taxonomy of resource optimisation for executing bag-of-task applications on public clouds," *Future Generation Computer Systems*, vol. 82, pp. 1–11, 2018.

[61] M. Challa and D. Sudha, "An efficient approach for minimization of energy and makespan in cloud computing," *Annals of the Romanian Society for Cell Biology*, vol. 25, no. 6, pp. 7422–7430, 2021.

[62] M. Hussain, L.-F. Wei, A. Lakhan, S. Wali, S. Ali, and A. Hussain, "Energy and performance-efficient task scheduling in heterogeneous virtualized cloud computing," *Sustainable Computing: Informatics and Systems*, vol. 30, p. 100517, 2021.

[63] T. A. Genez, L. F. Bittencourt, and E. R. Madeira, "Workflow scheduling for saas/paas cloud providers considering two sla levels," in *2012 IEEE Network Operations and Management Symposium*, 2012, pp. 906–912.

[64] S. Kardani-Moghaddam, R. Buyya, and K. Ramamohanarao, “Adrl: A hybrid anomaly-aware deep reinforcement learning-based resource scaling in clouds,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 3, pp. 514–526, 2020.

[65] S. Abrishami, M. Naghibzadeh, and D. H. Epema, “Deadline-constrained workflow scheduling algorithms for infrastructure as a service clouds,” *Future generation computer systems*, vol. 29, no. 1, pp. 158–169, 2013.

[66] E. Barrett, E. Howley, and J. Duggan, “A learning architecture for scheduling workflow applications in the cloud,” in *2011 IEEE ninth European conference on web services*, 2011, pp. 83–90.

[67] W.-N. Chen and J. Zhang, “A set-based discrete pso for cloud workflow scheduling with user-defined qos constraints,” in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2012, pp. 773–778.

[68] G. Yao, Y. Ding, and K. Hao, “Using imbalance characteristic for fault-tolerant workflow scheduling in cloud systems,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 12, pp. 3671–3683, 2017.

[69] X. Zhu, J. Wang, H. Guo, D. Zhu, L. T. Yang, and L. Liu, “Fault-tolerant scheduling for real-time scientific workflows with elastic resource provisioning in virtualized clouds,” *IEEE transactions on parallel and distributed systems: a publication of the IEEE Computer Society*, vol. 27, no. 12, pp. 3501–3517, 2016.

[70] X. Ye, J. Li, S. Liu, J. Liang, and Y. Jin, “A hybrid instance-intensive workflow scheduling method in private cloud environment,” *Natural Computing*, vol. 18, pp. 735–746, 2019.

[71] M. Masdari, S. ValiKardan, Z. Shahi, and S. I. Azar, “Towards workflow scheduling in cloud computing: a comprehensive analysis,” *Journal of Network and Computer Applications*, vol. 66, pp. 64–82, 2016.

[72] A. Verma and S. Kaushal, “Cost minimized pso based workflow scheduling plan for cloud computing,” *International Journal of Information Technology and Computer Science*, vol. 7, no. 8, pp. 37–43, 2015.

[73] S. Mohammadi, L. PourKarimi, and H. Pedram, “Integer linear programming-based multi-objective scheduling for scientific workflows in multi-cloud environments,” *The Journal of Supercomputing*, vol. 75, no. 10, pp. 6683–6709, 2019.

[74] A. C. Zhou, W. Xue, Y. Xiao, B. He, S. Ibrahim, and R. Cheng, “Taming system dynamics on resource optimization for data processing workflows: A probabilistic approach,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 1, pp. 231–248, 2021.

[75] T. Thanavanich and P. Uthayopas, “Efficient energy aware task scheduling for parallel workflow tasks on hybrids cloud environment,” in *2013 International Computer Science and Engineering Conference (ICSEC)*. IEEE, 2013, pp. 37–42.

[76] F. Juarez, J. Ejarque, and R. M. Badia, “Dynamic energy-aware scheduling for parallel task-based application in cloud computing,” *Future Generation Computer Systems*, vol. 78, pp. 257–271, 2018.

[77] M. Safari and R. Khorsand, “Energy-aware scheduling algorithm for time-constrained workflow tasks in dvfs-enabled cloud environment,” *Simulation Modelling Practice and Theory*, vol. 87, pp. 311–326, 2018.

[78] L. Zhang, L. Wang, Z. Wen, M. Xiao, and J. Man, “Minimizing energy consumption scheduling algorithm of workflows with cost budget constraint on heterogeneous cloud computing systems,” *IEEE Access*, vol. 8, pp. 205 099–205 110, 2020.

[79] R. Medara, R. S. Singh, and M. Sompalli, “Energy and cost aware workflow scheduling in clouds with deadline constraint,” *Concurrency and Computation: Practice and Experience*, vol. 34, no. 13, p. e6922, 2022.

[80] J. Wang, C. Huang, K. He, X. Wang, X. Chen, and K. Qin, “An energy-aware resource allocation heuristics for vm scheduling in cloud,” in *2013 IEEE 10th International Conference on High Performance Computing and Communications & 2013 IEEE International Conference on Embedded and Ubiquitous Computing*. IEEE, 2013, pp. 587–594.

[81] W. Zheng and S. Huang, “An adaptive deadline constrained energy-efficient scheduling heuristic for workflows in clouds,” *Concurrency and Computation: Practice and Experience*, vol. 27, no. 18, pp. 5590–5605, 2015.

[82] M. Khaleel and M. M. Zhu, “Energy-efficient task scheduling and consolidation algorithm for workflow jobs in cloud,” *International Journal of Computational Science and Engineering*, vol. 13, no. 3, pp. 268–284, 2016.

[83] H. Li, J. Li, W. Yao, S. Nazarian, X. Lin, and Y. Wang, “Fast and energy-aware resource provisioning and task scheduling for cloud systems,” in *2017*

18th international symposium on quality electronic design (ISQED), 2017, pp. 174–179.

[84] L. Ye, Y. Xia, S. Tao, C. Yan, R. Gao, and Y. Zhan, "Reliability-aware and energy-efficient workflow scheduling in iaas clouds," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 3, pp. 2156–2169, 2022.

[85] N. Mansouri and R. Ghafari, "Cost-efficient task scheduling algorithm for reducing energy consumption and makespan of cloud computing," *Computer and Knowledge Engineering*, vol. 5, no. 1, pp. 1–12, 2022.

[86] S. Pandey, L. Wu, S. M. Guru, and R. Buyya, "A particle swarm optimization-based heuristic for scheduling workflow applications in cloud computing environments," in *2010 24th IEEE international conference on advanced information networking and applications*, 2010, pp. 400–407.

[87] S. Yassa, R. Chelouah, H. Kadima, and B. Granado, "Multi-objective approach for energy-aware workflow scheduling in cloud computing environments," *The Scientific World Journal*, vol. 2013, no. 1, p. 350934, 2013.

[88] K. Sellami, P. F. Tiako, L. Sellami, and R. Kassa, "Energy efficient workflow scheduling of cloud services using chaotic particle swarm optimization," in *2020 IEEE Green Technologies Conference (GreenTech)*. IEEE, 2020, pp. 74–79.

[89] A. A. H. Al-Mahruqi, G. Morison, B. G. Stewart, and V. Athinarayanan, "Hybrid heuristic algorithm for better energy optimization and resource utilization in cloud computing," *Wireless Personal Communications*, vol. 118, no. 1, pp. 43–73, 2021.

[90] N. Khaledian, K. Khamforoosh, R. Akraminejad, L. Abualigah, and D. Javaheri, "An energy-efficient and deadline-aware workflow scheduling algorithm in the fog and cloud environment," *Computing*, vol. 106, no. 1, pp. 109–137, 2024.

[91] B. Xiang, B. Zhang, and L. Zhang, "Greedy-ant: ant colony system-inspired workflow scheduling for heterogeneous computing," *IEEE Access*, vol. 5, pp. 11 404–11 412, 2017.

[92] M. Shojafar, M. Kardgar, A. A. R. Hosseinabadi, S. Shamshirband, and A. Abraham, "Tets: a genetic-based scheduler in cloud computing to decrease energy and makespan," in *Hybrid Intelligent Systems: 15th International Conference HIS 2015 on Hybrid Intelligent Systems, Seoul, South Korea, November 16-18, 2015 15*, 2016, pp. 103–115.

[93] Y. Gu and C. Budati, "Energy-aware workflow scheduling and optimization in clouds using bat algorithm," *Future Generation Computer Systems*, vol. 113, pp. 106–112, 2020.

[94] A. Mohammadzadeh, M. Masdari, and F. S. Gharehchopogh, "Energy and cost-aware workflow scheduling in cloud computing data centers using a multi-objective optimization algorithm," *Journal of Network and Systems Management*, vol. 29, no. 3, p. 31, 2021.

[95] A. Mohammadzadeh, M. Masdari, F. S. Gharehchopogh, and A. Jafarian, "A hybrid multi-objective metaheuristic optimization algorithm for scientific workflow scheduling," *Cluster Computing*, vol. 24, no. 2, pp. 1479–1503, 2021.

[96] A. Mohammadzadeh, M. Akbari Zarkesh, P. Haji Shahmohamd, J. Akhavan, and A. Chhabra, "Energy-aware workflow scheduling in fog computing using a hybrid chaotic algorithm," *The Journal of Supercomputing*, vol. 79, no. 16, pp. 18 569–18 604, 2023.

[97] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," *IEEE transactions on evolutionary computation*, vol. 6, no. 2, pp. 182–197, 2002.

[98] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: Solving problems with box constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577–601, 2014.

[99] Q. Zhang and H. Li, "Moea/d: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on evolutionary computation*, vol. 11, no. 6, pp. 712–731, 2007.

[100] W. Guo, W. Tian, Y. Ye, L. Xu, and K. Wu, "Cloud resource scheduling with deep reinforcement learning and imitation learning," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3576–3586, 2021.

[101] N. Liu, Z. Li, J. Xu, Z. Xu, S. Lin, Q. Qiu, J. Tang, and Y. Wang, "A hierarchical framework of cloud resource allocation and power management using deep reinforcement learning," in *2017 IEEE 37th international conference on distributed computing systems (ICDCS)*, 2017, pp. 372–382.

[102] S. Mangalampalli, S. S. Hashmi, A. Gupta, G. R. Karri, K. V. Rajkumar, T. Chakrabarti, P. Chakrabarti, and M. Margala, "Multi objective prioritized workflow scheduling using deep reinforcement based learning in cloud computing," *IEEE access*, vol. 12, pp. 5373–5392, 2024.

[103] H. Lu, C. Gu, F. Luo, W. Ding, and X. Liu, "Optimization of lightweight task offloading strategy for mobile edge computing based on deep reinforcement learning," *Future Generation Computer Systems*, vol. 102, pp. 847–861, 2020.

[104] J. L. Berral, R. Gavalda, and J. Torres, "Adaptive scheduling on power-aware managed data-centers using machine learning," in *2011 IEEE/ACM 12th International Conference on Grid Computing*. IEEE, 2011, pp. 66–73.

[105] M. Cheng, J. Li, and S. Nazarian, "Drl-cloud: Deep reinforcement learning-based resource provisioning and task scheduling for cloud service providers," in *2018 23rd Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2018, pp. 129–134.

[106] A. R. Malipatil, M. Paramasivam, D. Gulyamova, A. Saravanan, J. V. N. Ramesh, E. Muniyandy, and R. Ghodhbani, "Energy-efficient cloud computing through reinforcement learning-based workload scheduling." *International Journal of Advanced Computer Science & Applications*, vol. 16, no. 4, 2025.

[107] J. Lou, Z. Tang, and W. Jia, "Energy-efficient joint task assignment and migration in data centers: A deep reinforcement learning approach," *IEEE Transactions on Network and Service Management*, vol. 20, no. 2, pp. 961–973, 2022.

[108] N. Shan, X. Cui, and Z. Gao, ""drl+ fl": An intelligent resource allocation model based on deep reinforcement learning for mobile edge computing," *Computer Communications*, vol. 160, pp. 14–24, 2020.

[109] Z. Zhu, G. Zhang, M. Li, and X. Liu, "Evolutionary multi-objective workflow scheduling in cloud," *IEEE Transactions on parallel and distributed Systems*, vol. 27, no. 5, pp. 1344–1357, 2015.

[110] Y. Zhou and X. Jiao, "Knowledge-driven multi-objective evolutionary scheduling algorithm for cloud workflows," *IEEE Access*, vol. 10, pp. 2952–2962, 2021.

[111] Azure, "Azure public dataset v2," https://github.com/Azure/AzurePublicDataset, 2019, accessed: 2025-05-17.

[112] R. N. Calheiros and R. Buyya, "Meeting deadlines of scientific workflows in public clouds with tasks replication," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1787–1796, 2013.

[113] W. Hu, X. Li, and X. Li, "Hybrid cloud workflow scheduling method with privacy data," *IEEE Access*, vol. 8, pp. 211 540–211 552, 2020.

[114] A. C. Zhou, B. He, and C. Liu, “Monetary cost optimizations for hosting workflow-as-a-service in iaas clouds,” *IEEE transactions on cloud computing*, vol. 4, no. 1, pp. 34–48, 2015.

[115] M. A. Rodriguez and R. Buyya, “Deadline based resource provisioning and scheduling algorithm for scientific workflows on clouds,” *IEEE transactions on cloud computing*, vol. 2, no. 2, pp. 222–235, 2014.

[116] Y. Wang and X. Zuo, “An effective cloud workflow scheduling approach combining pso and idle time slot-aware rules,” *IEEE/CAA journal of automatica sinica*, vol. 8, no. 5, pp. 1079–1094, 2021.

[117] H. Gu, X. Li, and Z. Lu, “Scheduling spark tasks with data skew and deadline constraints,” *IEEE Access*, vol. 9, pp. 2793–2804, 2020.

[118] O. H. Ahmed, J. Lu, A. M. Ahmed, A. M. Rahmani, M. Hosseinzadeh, and M. Masdari, “Scheduling of scientific workflows in multi-fog environments using markov models and a hybrid salp swarm algorithm,” *IEEE Access*, vol. 8, pp. 189 404–189 422, 2020.

[119] M. C. Calzarossa, M. L. Della Vedova, L. Massari, G. Nebbione, and D. Tessera, “Multi-objective optimization of deadline and budget-aware workflow scheduling in uncertain clouds,” *IEEE Access*, vol. 9, pp. 89 891–89 905, 2021.

[120] N. Zhou, F. Li, K. Xu, and D. Qi, “Concurrent workflow budget-and deadline-constrained scheduling in heterogeneous distributed environments,” *Soft Computing*, vol. 22, pp. 7705–7718, 2018.

[121] W. Chen and E. Deelman, “Workflowsim: A toolkit for simulating scientific workflows in distributed environments,” in *2012 IEEE 8th international conference on E-science*, 2012, pp. 1–8.

[122] J. J. Durillo and A. J. Nebro, “jmetal: A java framework for multi-objective optimization,” *Advances in engineering software*, vol. 42, no. 10, pp. 760–771, 2011.

[123] H. Gupta, A. Vahid Dastjerdi, S. K. Ghosh, and R. Buyya, “ifogsim: A toolkit for modeling and simulation of resource management techniques in the internet of things, edge and fog computing environments,” *Software: Practice and Experience*, vol. 47, no. 9, pp. 1275–1296, 2017.

[124] M. C. Silva Filho, R. L. Oliveira, C. C. Monteiro, P. R. Inácio, and M. M. Freire, “Cloudsim plus: a cloud computing simulation framework pursuing software engineering principles for improved modularity, extensibility and correctness,”

in *2017 IFIP/IEEE symposium on integrated network and service management (IM)*, 2017, pp. 400–406.

[125] H. Casanova, "Simgrid: A toolkit for the simulation of application scheduling," in *Proceedings First IEEE/ACM International Symposium on Cluster Computing and the Grid*, 2001, pp. 430–437.

[126] R. Grandl, G. Ananthanarayanan, S. Kandula, S. Rao, and A. Akella, "Multi-resource packing for cluster schedulers," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4, pp. 455–466, 2014.

[127] "Prefect documentation," 2024. [Online]. Available: https://docs.prefect.io/

[128] "Kubernetes documentation," 2024. [Online]. Available: https://kubernetes.io/

[129] J. Rothman and J. Chamanara, " An RL-Based Model for Optimized Kubernetes Scheduling ," in *2023 IEEE 31st International Conference on Network Protocols (ICNP)*. Los Alamitos, CA, USA: IEEE Computer Society, Oct. 2023, pp. 1–6. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/ICNP59255.2023.10355623

[130] N. D. Cicco, G. F. Pittalà, G. Davoli, D. Borsatti, W. Cerroni, C. Raffaelli, and M. Tornatore, "Drl-forch: A scalable deep reinforcement learning-based fog computing orchestrator," in *2023 IEEE 9th International Conference on Network Softwarization (NetSoft)*, 2023, pp. 125–133.

[131] R. S. Sutton, "Reinforcement learning: An introduction," *A Bradford Book*, 2018.

[132] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *ArXiv*, vol. abs/1707.06347, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:28695052

[133] D. Cordeiro, G. Mounié, S. Perarnau, D. Trystram, J.-M. Vincent, and F. Wagner, "Random graph generation for scheduling simulations," in *3rd International ICST Conference on Simulation Tools and Techniques*, ser. SIMUTools '10. Brussels, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2010.

[134] S. Huang, R. F. J. Dossa, C. Ye, J. Braga, D. Chakraborty, K. Mehta, and J. G. Araújo, "Cleanrl: High-quality single-file implementations of deep reinforcement learning algorithms," *Journal of Machine Learning Research*, vol. 23, no. 274, pp. 1–18, 2022.

[135] M. Kumar, S. Sharma, A. Goel, and S. Singh, “A comprehensive survey for scheduling techniques in cloud computing,” *Journal of Network and Computer Applications*, vol. 143, pp. 1–33, 2019.

[136] P. ERDdS and A. R&wi, “On random graphs i,” *Publ. math. debrecen*, vol. 6, no. 290-297, p. 18, 1959.

[137] W. M. van Der Aalst, A. H. Ter Hofstede, B. Kiepuszewski, and A. P. Barros, “Workflow patterns,” *Distributed and parallel databases*, vol. 14, pp. 5–51, 2003.

[138] S. Fu, “Failure-aware resource management for high-availability computing clusters with distributed virtual machines,” *Journal of Parallel and Distributed Computing*, vol. 70, no. 4, pp. 384–393, 2010.

[139] “Kind documentation,” 2026. [Online]. Available: https://kind.sigs.k8s.io/

[140] M. Li and X. Yao, “Quality evaluation of solution sets in multiobjective optimisation: A survey,” *ACM Computing Surveys*, vol. 52, pp. 1–38, 03 2019.

[141] S. Chandrasiri and D. Meedeniya, “Energy-efficient dynamic workflow scheduling in cloud environments using deep learning,” *Sensors*, vol. 25, no. 5, p. 1428, 2025.

[142] ——, “Multi-objective cloud workflow scheduling with two-stage deep reinforcement learning and graph neural networks,” *IEEE Access*, 2026.