

DAFTAR PUSTAKA

- [1] E. R. E. Sirait, "IMPLEMENTASI TEKNOLOGI BIG DATA DI LEMBAGA PEMERINTAHAN INDONESIA," *Jurnal Penelitian Pos dan informatika*, vol. 6, no. 2, p. 113, Dec. 2016.
- [2] H. Sh. Abdallah, M. H. Khafagy, and F. A. Omara, "Case Study: Spark GPU-Enabled Framework to Control COVID-19 Spread Using Cell-Phone Spatio-Temporal Data," *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1303–1320, 2020.
- [3] J. Wei, M. Chen, L. Wang, P. Ren, Y. Lei, Y. Qu, Q. Jiang, X. Dong, W. Wu, Q. Wang, K. Zhang, and X. Zhang, "Status, challenges and trends of data-intensive supercomputing," *CCF Transactions on High Performance Computing*, vol. 4, no. 2, pp. 211–230, Jun. 2022.
- [4] Imam Machdi, Alfatihah Reno Mnspm, Arie Wahyu Wijayanto, Amanda Pratama Putra, and Setia Pramana, *Implementasi Big Data dan Kecerdasan Artifisial untuk Statistik Ofisial*. Penerbit BRIN, Mar. 2023, pp. 47–53.
- [5] A. H. A. AL-Jumaili, R. C. Muniyandi, M. K. Hasan, J. K. S. Paw, and M. J. Singh, "Big Data Analytics Using Cloud Computing Based Frameworks for Power Management Systems: Status, Constraints, and Future Recommendations," *Sensors*, vol. 23, no. 6, p. 2952, Mar. 2023.
- [6] W. McKinney, "Python for Data Analysis," *O'Reilly Media*, vol. 2, Oct. 2017.
- [7] Y. K. Gupta and S. Kumari, "A Study of Big Data Analytics using Apache Spark with Python and Scala," in *2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*. IEEE, pp. 471–478. [Online]. Available: <https://ieeexplore.ieee.org/document/9315863/>
- [8] Harshal S. Kudale, Mihir V. Phadnis, Pooja J. Chittar, Kalpesh P. Zarkar, and Prof. Balaji K. Bodhke, "A REVIEW OF DATA ANALYSIS AND VISUALIZATION OF OLYMPICS USING PYSPARK AND DASH-PLOTLY," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 04, no. 06, pp. 2093–2097, Jun. 2022.
- [9] T. R. Rao, P. Mitra, R. Bhatt, and A. Goswami, "The big data system, components, tools, and technologies: A survey," vol. 60, pp. 1165–1245, 2019.
- [10] K. Kambatla, G. Kollias, V. Kumar, and A. Grama, "Trends in big data analytics," *Journal of Parallel and Distributed Computing*, vol. 74, no. 7, pp. 2561–2573, 2014.
- [11] A. Oussous, F.-Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, "Big Data technologies: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 431–448, Oct. 2018.
- [12] P. Kumar, P. Kumar, N. Zaidi, and V. S. Rathore, "Analysis and Comparative Exploration of Elastic Search, MongoDB and Hadoop Big Data Processing," in *Soft Computing: Theories and Applications*, M. Pant, K. Ray, T. K. Sharma, S. Rawat,

and A. Bandyopadhyay, Eds. Singapore: Springer Singapore, 2018, vol. 584, pp. 605–615.

- [13] S. Salloum, R. Dautov, X. Chen, P. X. Peng, and J. Z. Huang, “Big data analytics on Apache Spark,” *International Journal of Data Science and Analytics*, vol. 1, no. 3-4, pp. 145–164, Nov. 2016.
- [14] M. T. Gabdullin, Y. Suinullayev, Y. Kabi, J. W. Kang, and A. Mukasheva, “Comparative Analysis of Hadoop and Spark Performance for Real-time Big Data Smart Platforms Utilizing IoT Technology in Electrical Facilities,” *Journal of Electrical Engineering & Technology*, Jun. 2024.
- [15] Sifat Ibtisum, Ehsan Bazgir, S M Atikur Rahman, and S. M. Saokat Hossain, “A comparative analysis of big data processing paradigms: Mapreduce vs. apache spark,” *World Journal of Advanced Research and Reviews*, vol. 20, no. 1, pp. 1089–1098, Oct. 2023.
- [16] J. Russell, *Getting Started with Impala: Interactive SQL for Apache Hadoop*. "O'Reilly Media, Inc.", Sep. 2014.
- [17] T. Aydoğan, M. İlkuçar, and M. A. Akca, “An Analysis on the Comparison of the Performance and Configuration Features of Big Data Tools Solr and Elasticsearch,” *International Journal of Intelligent Systems and Applications in Engineering*, vol. 4, no. Special Issue-1, pp. 8–12, Dec. 2016.
- [18] M. A. Khan, M. R. Karim, and Y. Kim, “A Two-Stage Big Data Analytics Framework with Real World Applications Using Spark Machine Learning and Long Short-Term Memory Network,” *Symmetry*, vol. 10, no. 10, p. 485, Oct. 2018.
- [19] “ZooKeeper: Because Coordinating Distributed Systems is a Zoo,” <https://zookeeper.apache.org/doc/current/zookeeperOver.html>.
- [20] Avro. Avro. Apache Avro. [Online]. Available: <https://avro.apache.org/docs/>
- [21] Oozie. Oozie - Apache Oozie Workflow Scheduler for Hadoop. [Online]. Available: <https://oozie.apache.org/index.html>
- [22] Ambari. Introduction to Apache Ambari. 2024-02-01. [Online]. Available: <https://ambari.apache.org/>
- [23] Y. Mao, Y. Fu, S. Gu, S. Vhaduri, L. Cheng, and Q. Liu, “Resource Management Schemes for Cloud-Native Platforms with Computing Containers of Docker and Kubernetes,” Oct. 2020.
- [24] A. Amirkhanyan, “Methoden und Frameworks für geo-raumzeitliche Datenanalysen Methods and frameworks for GeoSpatioTemporal data analytics,” Ph.D. dissertation, Universität Potsdam, 2020.
- [25] S. Moosavi, M. H. Samavatian, A. Nandi, S. Parthasarathy, and R. Ramnath, “Short and Long-term Pattern Discovery Over Large-Scale Geo-Spatiotemporal Data,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. Anchorage AK USA: ACM, Jul. 2019, pp. 2905–2913.

- [26] Y. Xu, S.-L. Shaw, Z. Zhao, L. Yin, Z. Fang, and Q. Li, "Understanding aggregate human mobility patterns using passive mobile phone location data: A home-based approach," *Transportation*, vol. 42, no. 4, pp. 625–646, Jul. 2015.
- [27] J. Reif and D. Schmücker, "Exploring new ways of visitor tracking using big data sources: Opportunities and limits of passive mobile data for tourism," *Journal of Destination Marketing & Management*, vol. 18, p. 100481, Dec. 2020.
- [28] J. Yu and M. Sarwat, "Geospatial Data Management in Apache Spark: A Tutorial," in *2019 IEEE 35th International Conference on Data Engineering (ICDE)*. Macao, Macao: IEEE, Apr. 2019, pp. 2060–2063.
- [29] J. Raun, R. Ahas, and M. Tiru, "Measuring tourism destinations using mobile tracking data," *Tourism Management*, vol. 57, pp. 202–212, Dec. 2016.
- [30] S. Shekhar, Z. Jiang, R. Y. Ali, E. Eftelioglu, X. Tang, V. M. V. Gunturi, and X. Zhou, "Spatiotemporal Data Mining: A Computational Perspective," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2306–2338, Dec. 2015.
- [31] M. Chen, S. Mao, Y. Zhang, and V. C. Leung, *Big Data: Related Technologies, Challenges and Future Prospects*, ser. SpringerBriefs in Computer Science. Cham: Springer International Publishing, 2014.
- [32] H. B. Abdalla, "A brief survey on big data: Technologies, terminologies and data-intensive applications," *Journal of Big Data*, vol. 9, no. 1, p. 107, Nov. 2022.
- [33] M. Tiru, "Overview of the sources and challenges of mobile positioning data for statistics," in *Proceedings of the International Conference on Big Data for Official Statistics, Beijing, China, 2014*, pp. 28–30.
- [34] R. Ahas, S. Silm, O. Järv, E. Saluveer, and M. Tiru, "Using Mobile Positioning Data to Model Locations Meaningful to Users of Mobile Phones," *Journal of Urban Technology*, vol. 17, no. 1, pp. 3–27, Apr. 2010.
- [35] N. I. Hasanah, S. Somantri, and A. Sujjada, "IMPLEMENTATION OF THE HAVERSINE FORMULA IN THE WEBGIS OF PRIVATE UNIVERSITIES IN WEST JAVA AND BANTEN," *IDEALIS : InDonEsiA journal Information System*, vol. 7, no. 2, pp. 178–188, Jun. 2024.
- [36] V. Kurama, "Python Programming in 15 min Part 1," <https://towardsdatascience.com/python-programming-in-15-min-part-1-3ad2d773834c>, Jul. 2017.
- [37] Z. N. Rashid, S. R. M. Zebari, K. H. Sharif, and K. Jacksi, "Distributed Cloud Computing and Distributed Parallel Computing: A Review," in *2018 International Conference on Advanced Science and Engineering (ICOASE)*, Oct. 2018, pp. 167–172.
- [38] H.-C. Lu, F. J. Hwang, and Y.-H. Huang, "Parallel and distributed architecture of genetic algorithm on Apache Hadoop and Spark," *Applied Soft Computing*, vol. 95, p. 106497, Oct. 2020.

- [39] C. Carrión, “Kubernetes Scheduling: Taxonomy, Ongoing Issues and Challenges,” *ACM Computing Surveys*, vol. 55, no. 7, pp. 1–37, Jul. 2023.
- [40] S. Guan, C. Zhang, Y. Wang, and W. Liu, “Hadoop-based secure storage solution for big data in cloud computing environment,” *Digital Communications and Networks*, vol. 10, no. 1, pp. 227–236, Feb. 2024.
- [41] Hadoop, “Hadoop – Apache Hadoop 3.3.6,” <https://hadoop.apache.org/docs/stable/>.
- [42] W. G. Choi and S. Park, “A write-friendly approach to manage namespace of Hadoop distributed file system by utilizing nonvolatile memory,” *The Journal of Supercomputing*, vol. 75, no. 10, pp. 6632–6662, Oct. 2019.
- [43] D. García-Gil, S. Ramírez-Gallego, S. García, and F. Herrera, “A comparison on scalability for batch big data processing on Apache Spark and Apache Flink,” *Big Data Analytics*, vol. 2, no. 1, p. 1, Dec. 2017.
- [44] C. Zhu, B. Han, and Y. Zhao, “A comparative performance study of spark on kubernetes,” *The Journal of Supercomputing*, vol. 78, no. 11, pp. 13 298–13 322, Jul. 2022.
- [45] Ilum. What is Ilum? | Ilum Documentation - Managed Spark Cluster. [Online]. Available: <https://illum.cloud/docs/>
- [46] ILUM. Architecture | Ilum Documentation - Managed Spark Cluster. [Online]. Available: <https://illum.cloud/docs/architecture/>
- [47] M. Holopainen, “Monitoring Container Environment with Prometheus and Grafana.”
- [48] J. Pivotto and B. Brazil, *Prometheus: Up & Running*. "O'Reilly Media, Inc."
- [49] M. Großmann and C. Schenk, “A Comparison of Monitoring Approaches for Virtualized Services at the Network Edge,” in *2018 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, pp. 85–90. [Online]. Available: <https://ieeexplore.ieee.org/document/8695277/?arnumber=8695277>
- [50] B. T. Hasan and D. B. Abdullah, “Real-Time Resource Monitoring Framework in a Heterogeneous Kubernetes Cluster,” in *2022 Muthanna International Conference on Engineering Science and Technology (MICEST)*, pp. 184–189. [Online]. Available: <https://ieeexplore.ieee.org/document/9790264/?arnumber=9790264>
- [51] Grafana. Grafana Cloud | Observability platform overview. Grafana Labs. [Online]. Available: <https://grafana.com/products/cloud/>
- [52] A. Makris, I. Kontopoulos, E. Psomakelis, S. N. Xyalis, T. Theodoropoulos, and K. Tserpes, “Performance Analysis of Storage Systems in Edge Computing Infrastructures,” vol. 12, no. 17, p. 8923. [Online]. Available: <https://www.mdpi.com/2076-3417/12/17/8923>