

## DAFTAR PUSTAKA

- Agarwal, B. B., Tayal, S. P., & M., G. (2010). *Software Engineering and Testing*. Sudbury: Jones and Bartlett.
- Aydar, d. (2016). *Software Testing Tutorial*. [www.tutorialspoint.com: http://www.tutorialspoint.com/software\\_testing/index.htm](http://www.tutorialspoint.com/software_testing/index.htm)
- Draisbach, U., Naumann, F., Szott, S., & Wonneberg, O. (2012). Adaptive windows for duplicate detection. *Proceedings - International Conference on Data Engineering*, 1073–1083.
- Guo, L., Wang, W., Chen, F., Tang, X., & Wang, W. (2012). A Similar Duplicate Data Detection Method Based on Fuzzy Clustering for Topology Information. In *PRZEGLĄD ELEKTROTECHNICZNY (Electrical Review)*, 01b (pp. 26-31).
- Hernandez, M. A., & Stolfo, S. J. (1995). The Merge/Purge Problem for Large Database. (pp. 128-129). New York: NYS Science and Technology Foundation.
- Kroenke, D. M., Saat, S., & Nugraha, D. (2003). *Database Processing Jilid 1 Edisi 9*. Jakarta: Erlangga.
- Lee, M. L., Lu, H., Ling, T. W., & Ko, Y. T. (1999). Cleansing Data for Mining and Warehousing. *10th International Conference on Database and Expert Systems Applications*. Italy: Cleansing Data for Mining and Warehousing.
- Lemos, O. A., Franchin, I. G., & Masiero, P. C. (2009). Integration testing of Object-Oriented and Aspect-Oriented programs. *Science of Computer Programming*, 74, 861-878.
- Malhotra, J., & Bakal, J. (2015). A survey and comparative study of data deduplication techniques. 2015 International Conference on Pervasive Computing: Advance Communication Technology and Application for Society, ICPC 2015, 0–4. <https://doi.org/10.1109/PERVASIVE.2015.7087116>
- Mualifa, R. (2016). Rancang Bangun Sistem Data Cleaning untuk Master Data Konsumen di PT XYZ Dengan Menerapkan Metode Sorted Neighbourhood dan Metode N-gram.
- Murali, D., & Lavanya, V. (2016). *Duplicate Detection Using Adaptive Window Technique*. 7(11), 2–4.
- Naumann, F., & Herschel, M. (2010). An Introduction to Duplicate Detection. In *Synthesis Lectures on Data Management* (Vol. 2, Issue 1).
- Rahaman, G. M., Rahman, A., & Ripon, K. S. (2010, December 12). A Domain-Independent Data Cleaning Algorithm for Detecting Similar-Duplicates. *Journal of Computers*, Vol. 5, No. 12. Bangladesh: Academy Publisher.

- Recchia, G., & Max, L. (2013). A Comparison of String Similarity Measures for Toponym Matching. *ACM SIGSPATIAL COMP'13*. New York.
- Schacherer, C. W. (2012). SAS® Data Management Technique: Cleaning and Transforming Data for Delivery of Analytic Datasets.
- Shukla, P., & Somani, A. K. (2014). Context-Aware Duplicate Detection in Semi-structured Data Streams. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10), 1915–1926. <https://doi.org/10.1109/TPAMI.2011.272>
- Skandar, A., Rehman, M., & Anjum, M. (2015). An Efficient Duplication Record Detection Algorithm for Data Cleansing. *International Journal of Computer Applications*, 127(6), 28–37.
- Tian, Z., Lu, H., Ji, W., Zhou, A., & Tian, Z. (2001). An n-gram-based Approach for Detecting Approximately Duplicate Database Records. *Springer Verlag*.
- Tziouvara, V., Vassiliadis, P., & Simitsis, A. (2007). Deciding the physical implementation of ETL workflows. *DOLAP: Proceedings of the ACM International Workshop on Data Warehousing and OLAP*, 49–56.
- Weis, M., Naumann, F., Jehle, U., Lufter, J., & Schuster, H. (2008). Industry-scale duplicate detection. *Proceedings of the VLDB Endowment*, 1(2), 1253–1264. <https://doi.org/10.14778/1454159.1454165>