# DAFTAR PUSTAKA


Acronis, Inc. (2009). How Deduplication Benefits Companies of All Sizes. 4.

Adom, D., Hussein, E. K., & Agyem, J. A. (2018). THEORETICAL AND CONCEPTUAL FRAMEWORK: MANDATORY INGREDIENTS OF A QUALITY RESEARCH. 439.

Aggarwal, C. C., & Reddy, C. K. (2013). *DATA CLUSTERING Algorithms and Applications.* New York: CRC Press.

Becker, J., Niehaves, B., & Janiesch, C. (2010). Socio-Technical Perspectives on Design Science in IS Research. *Information Systems and eBusiness Management, Vol.9, issue 1*, 109-131.

Boselli, R., Cesarini, M., Mercorio, F., & Mezzanzanica, M. (2014). Planning Meets Data Cleansing. 1.

Chezian, R. (2016). Methods and techniques to evaluate the performance of Data. 777.

Desai, B. C. (2014). The State of Data. 77.

Devi, S., & Kalia, D. (2015). Study of Data Cleaning & Comparison of Data Cleaning Tools. 362-364.

Do, H. H., & Rahm, E. (2000). Data Cleaning: Problems and Current Approaches. 3-5.

Dwiandriani, F. (2017). ANALISIS DAN PERANCANGAN ARSITEKTUR APLIKASI DATA PROFILING BERBASIS OPEN SOURCE MENGGUNAKAN METODE REVERSE ENGINEERING UNTUK MENINGKATKAN DATA QUALITY MANAGEMENT PADA STUDI KASUS PT XYZ. Bandung.

Experian Information Solution, Inc. (2013). The State of Data Quality. 5-8.

Gunatama, M. (2018). ANALISIS MULTI COLUMN MENGGUNAKAN OPENSOURCE PLATFORM UNTUK PROFILING DATA: OUTLIERS DAN DEDUPLICATION. Bandung.

Gunatama, M., Kusumasari, T. F., & Hasibuan, M. A. (2018). Deduplication for Data Profiling using Open Source Platform.

Hellerstein, J. M. (2008). Quantitative Data Cleaning for Large Databases. 1-3.

Hevner, A. R., Ram, S., March, S. T., & Park, J. (2004). Design Science in Information Systems Research. *MIS Quarterly Vol. 28 No. 1*, 75-105.

Hevner, A., & Chatterjee, S. (2010). *Design Research in Information System : Theory and Practice.* New York: Springer.

Hildenbrand, T., Rothlauf, R., Geisser, M., Heinzl, A., & Kude, T. (2008). Approach to Collaborative Software Development. *International Conference on Complex, Intelligent and Software Intensive Systems* (pp. 523-528). Barcelona: IEEE.

Hu, W., Zaveri, A., Qiu, H., & Dumontier, M. (2017). Cleaning by clustering: methodology for addressing data quality issues in biomedical metadata.

Khoirunisa, A. N., Kusumasari, T. F., & Hasibuan, M. A. (2017). ANALYSIS AND DESIGN OF APPLICATION ARCHITECTURE DATA.

Kumar, R., & Chandrasekaran, D. R. (2011). ATTRIBUTE CORRECTION-DATA CLEANING USING ASSOCIATION RULE AND CLUSTERING METHODS. 22.

Kumar, V., & Khosla, C. (2018). A thorough analysis and survey on Unstructured data. *Data Cleaning*.

Kusumasari, T. F., & Fitria. (2016). Data Profiling for Data Quality Improvement with Openrefine.

Leszek Ziora, A. C. (2015). The role of big data solutions in the management of organizations.

Medhat, D., Hassan, A., & Salama, C. (2015). A Hybrid Cross-Language Name Matching. 204-205.

Meister, D. (2013). Advanced Data Deduplication Techniques and Their Applications.

Noor, S. A. (2019). *ANALISIS DAN PERANCANGAN DATA CLEANSING MENGGUNAKAN OPEN SOURCE TOOLS.* Bandung.

Oni, S., Chen, Z., Susan, H., & Jademi, O. (2019). A Comparative Study of Data Cleaning Tools.

Pratiktio, R. P. (2020). ANALISIS DAN PERANCANGAN PEDOMAN, PROSES DAN TEKNIK PENILAIAN MANAJEMEN KUALITAS DATA MENGGUNAKAN DAMA-DMBOK DAN PROCESS ASSESSMENT MODEL. Bandung.

Pulla, V. V., Varol, C., & Al, M. (2016). Open Source Data Quality Tools: Revisited. 7-8.

Sun, G.-Z., Dong, Y., Chen, D.-W., & Wei, J. (2010). Data backup and recovery based on data de-duplication. 379.

Udechukwu, A., Ezeife, C., & Barker, K. (2003). Independent De-Duplication In Data Cleaning. 56.

Vandenbelt, D. (2017, November 15). *Filter Rows*. Retrieved May 12, 2020, from https://wiki.pentaho.com/display/EAI/Filter+Rows

Vandenbelt, D. (2017, November 15). *Insert - Update*. Retrieved May 12, 2020, from https://wiki.pentaho.com/display/EAI/Insert+-+Update

Vandenbelt, D. (2017, November 15). *Sort Rows*. Retrieved May 12, 2020, from https://wiki.pentaho.com/display/EAI/Sort+rows

Wangikar, V. C., & Desmukh, R. R. (2015). Current Approaches and Issues. *Data Cleaning*.

Wilkerson, P. (2020, February 4). *Select Values*. Retrieved May 12, 2020, from https://wiki.pentaho.com/display/EAI/Select+Values

Wilkerson, P. (2020, February 3). *Table Input*. Retrieved May 12, 2020, from https://wiki.pentaho.com/display/EAI/Table+Input

Wilkerson, P. (2020, February 3). *Unique Rows (HashSet)*. Retrieved May 12, 2020, from https://wiki.pentaho.com/pages/viewpage.action?pageId=25044474

Ye, N. (2003). *The Hand Book of Data Mining.* New Jersey: Lawrence Erlbaum Associates.