

ABSTRACT

ANALYSIS AND DESIGN OF MODULE DATA CLEANSING GENERICALLY USING OPEN SOURCE TOOLS

By

H Aidar Alvinanda SulistyO

NIM: 1202164150

In today's digital world, maintain the quality of the data is emerging as the keystone of business to obtain high-quality data to create valid and trustworthy decisions. Data collected from various sources that are dirty and will affect the quality of decision-making in an organization. Incomplete, inaccurate, or erroneous data needs to be corrected, with the data cleansing processing can significantly improve the quality of the data. Wrong format and blank data are among many problems to maintain data quality management in an organization, especially in Indonesian government agencies. Therefore this research will discuss one of the methods in data quality management that are data cleansing. We propose the design and analysis of modules in data cleansing with pattern and null method using open-source tools. The result will be implemented to Data Quality Management (DQM) dashboard to solve the problem of uniformity format in drug and food distribution permits and also the blank or null data in brand registration number permits in Indonesian government agencies. This research aims to analyze and design a data cleansing process that can be accessed flexibly using available modules and producing data according to organization needs.

Keywords: data cleansing, data quality management, data pattern, data null, open-source.