

ABSTRAK

ANALISIS DAN PERANCANGAN MODUL *DATA CLEANSING* SECARA *GENERIC* MENGGUNAKAN *OPEN SOURCE TOOLS*

Oleh

HAIDAR ALVINANDA SULISTYO

NIM: 1202164150

Pada era digital saat ini, menjaga kualitas data merupakan batu pijakan dalam dunia bisnis dalam mendapatkan data berkualitas tinggi untuk membuat keputusan valid dan dapat dipercaya. Data dikumpulkan dari berbagai sumber, data yang kotor akan mempengaruhi kualitas pengambilan keputusan dalam suatu organisasi. Data yang tidak lengkap, tidak akurat, atau salah perlu diperbaiki. Dengan proses *data cleansing*, secara signifikan dapat meningkatkan kualitas dari data. Format yang salah dan kekosongan data merupakan sebagian di antara banyak masalah untuk menjaga kualitas data dalam suatu organisasi, terutama di lembaga pemerintah Indonesia. Oleh karena itu penelitian ini akan membahas salah satu metode dalam manajemen kualitas data yaitu *data cleansing*. Kami mengusulkan sebuah desain dan analisis dari modul *data cleansing* menggunakan metode *data pattern* dan metode *data null* menggunakan *open-source tools*. Hasilnya akan diimplementasikan ke dalam *dashboard Data Quality Management (DQM)* untuk mengatasi permasalahan tidak seragamnya format Nomor Izin Edar (NIE) obat dan makanan serta data kosong atau *blank* dalam Nomor Izin Edar Merk di lembaga pemerintah Indonesia. Tugas akhir ini bertujuan untuk menganalisis dan merancang proses pembersihan data yang dapat diakses secara fleksibel menggunakan modul yang tersedia dan menghasilkan data sesuai dengan kebutuhan organisasi.

Kata kunci: *data cleansing, data quality management, data pattern, data null, open-source*.