

## **ABSTRAK**

# **ANALISIS DAN PERANCANGAN *DATA CLEANSING* MENGUNAKAN *OPEN SOURCE TOOLS***

Oleh

**SITI AULIA NOOR**

**1202151364**

Seiring cepatnya perkembangan teknologi akan menghasilkan data atau informasi yang beragam sehingga dapat mempengaruhi dalam proses pengambilan keputusan. Hal ini menyebabkan banyaknya organisasi yang berlomba untuk mengolah data yang dimiliki masing-masing organisasi. Disamping itu, kualitas data menjadi faktor pendukung penting untuk mengolah suatu data agar dapat menghasilkan sebuah informasi valid yang dapat berguna bagi perusahaan. Namun, pada kenyataannya masih banyak organisasi atau perusahaan yang belum memiliki kualitas data yang baik, hal ini disebabkan oleh banyaknya kesalahan pada proses input data dan perbedaan format data untuk setiap aplikasi. Untuk menanggulangi permasalahan tersebut terdapat proses data *cleansing* yang berguna untuk menemukan kesalahan dan melakukan pembersihan data yang tidak sesuai format. Metode tersebut dapat dilakukan dengan berbagai aplikasi berbayar ataupun gratis, namun dikarenakan tingginya biaya dari aplikasi berbayar membuat banyak organisasi memanfaatkan aplikasi *open source* sehingga dilakukan perbandingan untuk menentukan aplikasi berbasis *open source* yang sesuai dengan kebutuhan organisasi. Pada penelitian ini, peneliti melakukan analisis dan perancangan arsitektur data *cleansing*. Hasil dari algoritma yang dibuat akan diimplementasikan menggunakan *open source tool* dan dilakukan komparasi dengan penelitian lain untuk melihat perbedaan hasil dari algoritma yang telah dibuat.

Kata kunci : *data cleansing, open source, data quality, algoritma, data pattern, data null*