

ABSTRAK

Profiling memiliki peran yang penting dalam proses suatu data sehingga data yang digunakan adalah data dengan kualitas yang baik. Banyak ditemui dalam permasalahan perusahaan, bahwa data yang akan diproses merupakan data yang masih kotor. Hal ini membuat data tidak diproses atau dianalisis dengan maksimal. Jika data yang ada merupakan data kotor, maka keputusan yang akan diambil oleh perusahaan belum tentu benar.

Dalam berbagai sumber terdapat klasifikasi *profiling* dengan nama yang berbeda – beda , tetapi memiliki maksud yang sama. *Profiling* dibagi menjadi 3 bagian besar yaitu *Single Column*, *Multi Column*, dan *Dependency*. *Single column* merupakan analisis *profiling* dalam satu kolom saja. *Multi column* menggunakan analisis *profiling* pada dua atau lebih kolom yang berbeda. Sedangkan, *dependency* menggunakan analisis *profiling* pada suatu tabel yang saling bergantung.

Pada penelitian ini dibuat *package* baru berupa *Cardinalities*, *domain analysis*, dan *data type* yang termasuk kedalam *single column*. Penelitian ini menggunakan *database* dan parameter yang dapat disesuaikan dengan kebutuhan perusahaan. Hasil dari penelitian ini dapat digunakan dalam berbagai kebutuhan analisis pada perusahaan.

Kata kunci—*profiling, cardinalities, domain analysis, data type*