

BAB I PENDAHULUAN

Pada bab ini berisi mengenai penjelasan latar belakang, rumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penelitian yang akan digunakan sebagai dasar dan acuan dari penulisan penelitian ini.

I.1 Latar Belakang

Data merupakan salah satu komponen terpenting dalam organisasi atau perusahaan sebagai aset yang vital. Data dan informasi dapat memberikan pengetahuan mengenai pelanggan, produk dan layanan yang dimiliki (Dama International, 2017). Organisasi dimanapun memanfaatkan data dalam berinovasi dan membuat strategi bisnis yang baik dan efektif karena data dibutuhkan diseluruh lini area baik bisnis, proses dan system yang berjalan. Data yang diolah dengan baik dan sesuai dengan kebutuhan akan dipastikan memiliki kualitas yang tinggi. Data dengan kualitas yang tinggi akan memberikan dampak positif dengan menciptakan informasi dan pengetahuan yang dapat membantu mengambil keputusan dengan baik (Kusumasari & Fitria, 2017).

Data quality merupakan sebuah ukuran dari data, dimana data yang memiliki kualitas yang tinggi memiliki karakter seperti lengkap, terbaru, aman dan bisa diakses kapan saja (Dama International, 2017). Data yang memiliki kualitas yang tinggi merupakan data yang sesuai dengan tujuan dan strategi bisnis organisasi yang akan menggunakannya karena data yang berkualitas tinggi setiap organisasi berbeda-beda sesuai dengan kebutuhan masing-masing (Dama International, 2017). Jadi tantangan dalam menyiapkan, mengelola, dan memiliki data dengan kualitas yang tinggi adalah memahami kebutuhan dari pengguna data. Karena dalam kenyataannya banyak organisasi yang hanya menumpuk data tanpa adanya sistem atau cara penyimpanan dan pengelolaan yang baik membuat ukuran data menjadi besar dan sulit untuk dipahami sehingga, informasi yang terkandung didalamnya tidak dapat tersampaikan dengan baik dan membuat keputusan yang bisa diambil dari informasi data tersebut tidak memberikan dampak positif yang maksimal.

Menurut data dari survey yang dilakukan oleh *Experian Data Quality* lebih dari 90 % organisasi masih melakukan kesalahan data kontak hal ini terjadi akibat masih digunakannya cara dan sistem manual dalam pengecekan data setelah data dikumpulkan (Shelly Kramer, 2015). Akibat dari masih belum layaknya data yang dimiliki membuat tingkat produktifitas dan keberhasilan bisnis menurun. Untuk mendapatkan data dengan kualitas data yang tinggi perlu adanya pembangunan infrastruktur dan sistem pengolahan data yang terstruktur dan terintegrasi. Kualitas data yang rendah dapat mempengaruhi tata kelola data di organisasi dimana dapat mempengaruhi proses perencanaan, pengawasan dan pengendalian data dan penggunaannya (Dama International, 2017).

Untuk memiliki data yang berkualitas diperlukan pengelolaan data untuk memastikan data memiliki kualitas yang tinggi, maka diperlukan *Data Quality Management* (DQM) dapat membantu dalam menjaga kualitas data dengan memastikan tidak ada data yang hilang dan rusak dengan adanya aturan, proses, dan sistem (Dama International, 2017). Dalam DQM terdapat beberapa proses untuk memastikan data memiliki kualitas yang tinggi sesuai dengan standar karakteristik akurat, lengkap, konsisten, terintegritas, terbaru, unik dan valid (Dama International, 2017). DQM memiliki beberapa tahapan didalamnya diantaranya *data profiling*, *data cleansing*, *data quality assessment*, dan *data quality monitoring* (Effendy et al., 2020). DQM dapat digunakan untuk menyelesaikan masalah data yang secara umum disebabkan oleh perbedaan format data yang bersumber dari berbagai sumber dan memiliki ukuran yang akan semakin membesar membuat harus ada penyesuaian dan pembersihan terhadap data tersebut. Masalah tersebut dapat membuat data yang seharusnya bisa digunakan dan dapat memiliki nilai untuk membantu dalam mengambil keputusan (Laksono et al., 2020).

Untuk mendapatkan data yang berkualitas salah satu tahapannya dalam DQM adalah *data profiling*. Dalam prosesnya data profiling adalah kegiatan dimana dilakukan proses pengecekan data yang tersedia dan tersimpan untuk menentukan isi struktur dan kualitas data (Kusumasari & Fitria, 2017). *Data profiling* membantu dalam membuat data sesuai dengan kebutuhan data dari organisasi sehingga data dapat memberikan informasi secara maksimal dalam

membantu pengambilan keputusan (Laksono et al., 2020). Pada tahapan DQM selanjutnya ada proses *data cleansing* dimana pada tahapan ini dilakukan proses pembersihan data yang membuat data menjadi sesuai dengan kebutuhan dari organisasi, seperti pengecekan salah ketik, tanda baca, dan kata hubung sehingga data menjadi beraturan dan bisa dilakukan proses pengolahan data untuk menghasilkan informasi yang dapat membantu proses pengambilan keputusan (Kusumasari & Fitria, 2017).

Namun sebelum dilakukan pembersihan perlu adanya pengecekan yang dilakukan dalam proses *data profiling* seperti pengecekan persebaran data, pengecekan data yang kosong, pengecekan data yang sama dan lain sebagainya yang diwakilkan oleh modul yang akan di implementasikan pada penelitian ini yaitu *Value Distribution, Data Completeness, Clustering, Data Similarity* dan *Data Deduplication*. Penelitian ini akan diimplementasikan pada *platform open source* yang mana tujuannya agar aplikasi nantinya bisa digunakan secara gratis dan oleh siapa saja yang memerlukannya sehingga nantinya permasalahan dalam kualitas data dapat terselesaikan dengan mudah dan gratis (Colford, 2009). *Tools* yang digunakan dalam implementasi adalah Laravel, MySQL dan Pentaho dari semua *tools* tersebut semuanya dapat diakses dengan mudah dan gratis sehingga memudahkan proses penelitian ini.

Pada penelitian ini akan berfokus untuk melakukan implementasian fitur *data profiling* pada modul data quality management didalam aplikasi *website Data Governance* yang telah dirancang pada penelitian sebelumnya oleh (Dewanti, 2021) dan (Rahmani, 2021). Penelitian ini diharapkan dapat mempermudah stakeholder dalam mengelola data agar dari informasi yang didapat dari hasil pengolahan data yang dilakukan pada sistem dari website *Data Governance* dapat menghasilkan informasi yang baik dan valid.

I.2 Perumusan Masalah

Berdasarkan permasalahan yang telah dijabarkan pada latar belakang, terdapat rumusan masalah yang mendasari penelitian ini yaitu

- a. Bagaimana penyesuaian yang bisa dilakukan agar fitur *data profiling* dapat diimplementasikan pada aplikasi *Data Governace* ?
- b. Bagaimana penyesuaian yang bisa dilakukan agar fitur *data profiling* dapat dijalankan pada *platform* Linux ?

I.3 Tujuan Penelitian

Berdasarkan rumusan masalah sebelumnya, tujuan yang ingin dicapai pada penelitian ini adalah sebagai berikut :

- a. Mengimplementasikan fitur *data profiling* dalam modul *data quality management* pada aplikasi *website Data Governance*.
- b. Menyesuaikan fitur agar dapat digunakan dengan baik pada *platform* Linux

I.4 Batasan Penelitian

Berikut merupakan batasan dalam melakukan penelitian adalah sebagai berikut :

- a. Penelitian ini dilakukan didalam aplikasi *Data Governance*
- b. Melakukan penyesuaian fitur pada modul *Data Quality Management*.
- c. Melakukan implementasi fitur data profiling dalam modul *Data Quality Management*.
- d. Fitur akan diimplementasikan dalam kerangka kerja *Laravel*.
- e. Basis data yang digunakan adalah *RBDMS* dari *MySQL*.

I.5 Manfaat Penelitian

Manfaat yang diharapkan dalam penelelitian ini dibagi menjadi dua bidang yang meliputi bidang teknis dan bidang keilmuan sebagai berikut .

- a. Manfaat dalam bidang teknis adalah membantu menyelesaikan permasalahan yang ada didalam organisasi dalam mengembangkan *tools* yang dapat membantu melakukan pengelolaan data untuk memiliki data yang berkualitas.
- b. Manfaat dalam bidang keilmuan diharapkan nantinya hasil dari penelitian ini dapat membantu menyelesaikan permasalahan sejenis yang terjadi dimasa yang akan datang.

I.6 Sistematika Penulisan

Sistem penulisan pada penelitian kali ini akan dibagi menjadi beberapa bab dari pokok bahasan yang secara umum akan dibagi sebagai berikut :

BAB 1 PENDAHULUAN

Pada bab ini berisi mengenai penjelasan latar belakang, rumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penelitian yang akan digunakan sebagai dasar dan acuan dari penulisan penelitian ini.

BAB 2 TINJAUAN PUSTAKA

Pada bab tinjauan pustaka ini terdapat beberapa hal yang dijelaskan sebagai referensi dalam melakukan penelitian ini yang diantaranya adalah penjelasan dan gambaran umum mengenai *Data Governance* yang lebih spesifik proses yaitu *Data Quality Management* dan proses yang akan menjadi fokus utama dalam penelitian ini yaitu implementasi *data profiling*.

BAB 3 METODOLOGI PENELITIAN

Metedologi penelitian ini akan berisi mengenai penjelasan model konseptual penelitian dan sistematika penelitian. Model Konseptual akan berisi mengenai simulasi penyelesaian masalah yang akan diteilti. Sistematikan penelitian akan menjelaskan proses yang akan dilakukan untuk menyelesaikan penelitian.

BAB 4 ANALISIS DAN PERANCANGAN

Pada bab ini akan dibahas mengenai analisis secara menyeluruh mengenai sistem dan aplikasi eksisting yang telah dibuat dan diimplementasikan pada penelitian sebelumnya dan akan dibahas mengenai perbandingan antara aplikasi eksisting saat ini dan hasil rancangan yang akan diusulkan pada penelitian kali ini.

BAB 5 IMPLEMENTASI DAN PENGUJIAN

Pada bab ini akan dibahas mengenai implelementasi dan pengujian yang telah di usulkan dalam bab sebelumnya secara menyeluruh pada penulisan penelitian ini.

BAB 6

KESIMPULAN DAN SARAN

Pada bab ini menjelaskan kesimpulan dan saran yang berisikan tentang ringkasan hasil dari penelitian yang sudah dilakukan dan saran yang dapat dipertimbangkan untuk penelitian berikutnya.