

## **BAB II**

### **TINJAUAN PUSTAKA DAN DASAR TEORI**

#### **2.1 Tinjauan Pustaka**

Lusidai (2017) "*Perancangan data warehouse pada perpustakaan Universitas Nusantara PGRI Kediri*" dimana penelitian tersebut menjelaskan bahwa sulitnya melakukan pengolahan data dan pemeliharaan data-data operasional perpustakaan. Maka penulis melakukan perancangan data warehouse perpustakaan beserta aplikasinya yang akan dijadikan pangkalan data sebagai dasar pengambilan keputusan mampu menyajikan informasi yang multidimensi secara tepat, akurat dan dalam waktu yang cepat.

Andri dkk. (2015) "*Analisis Dan Perancangan data warehouse pada perpustakaan Universitas Binadharma Palembang*" menggunakan perangkat lunak Pentaho untuk memanfaatkan teknologi informasi terutama penyimpanan dalam sebuah pengembangan aplikasi yang ada dalam perpustakaan yang dapat mempermudah proses pelaporan dan analisis.

Paramita. A (2016) "*Perancangan Data Warehouse Perpustakaan Yayasan Lentera Insan*" merancang data warehouse untuk memproses data perpustakaan yang belum terintegrasi. Metode penelitian yang digunakan oleh penulis dalam penelitian ini adalah metode grounded

(*grounded research*) yaitu suatu metode penelitian berdasarkan pada fakta dan menggunakan analisis perbandingan.

Supriyatna. A dan Wahyudi. M (2012) penelitian yang dilakukan adalah ”*Perancangan Data Warehouse Pada Perpustakaan Bina Sarana Informatika*” penelitian ini menggunakan metode OLAP membuat rancangan data warehouse beserta aplikasinya yang dijadikan sebagai pangkalan data yang berisi mengenai informasi perpustakaan yang akan dijadikan sebagai pengambilan keputusan dan menyajikan informasi secara multidimensi yang tepat, akurat, dan cepat.

Ambarwati (2017) melakukan Perancangan dan Pembuatan Data Warehouse pada Perpustakaan STMIK AMIKOM Yogyakarta dalam penelitian ini penulis menggunakan metode OLAP yang membahas perancangan skema, database data warehouse proses loading per periode hingga ringkasan data warehouse yang diolah menjadi *reporting*.

Dahlan dkk. (2013) “*Perancangan Data Warehouse Pada Perpustakaan Perguruan Tinggi XYZ*” penelitian ini menggunakan *Snowflake Schema* yang membahas proses ETL pada database perpustakaan dan database data warehouse dan menghasilkan data histori proses loading sesuai yang akan diinginkan.

Wulandari (2021) Perancangan Data Warehouse Perpustakaan SD N Brengosan 1. Penelitian ini membuat perancangan data warehouse pada

perpustakaan, menggunakan *Star Schema* yang akan menjelaskan hubungan database perpustakaan dan database data warehouse, menjelaskan proses ETL yang akan menghasilkan laporan yang akan dijadikan data histori untuk pelaporan dan pengambilan keputusan.

Tabel 2.1 merupakan perbandingan penelitian yang diusulkan peneliti sebelumnya.

**Tabel 2.1 Tabel Perbandingan Penelitian**

| Penulis                             | Judul   | Metode                          | Keterangan   |
|-------------------------------------|---|---------------------------------|--|
| Lusidai (2017)                      | <i>Perancangan data warehouse pada perpustakaan Universitas Nusantara PGRI Kediri</i>             | OLAP                            | Menyajikan informasi yang multidimensi secara cepat akurat daam waktu yang tepat.  |
| Andri dkk. (2015)                   | <i>Analisis Dan Perancangan data warehouse pada perpustakaan Universitas Binadharna Palembang</i> | OLAP                            | Bertujuan menghasilkan informasi sebagai pengambilan keputusan.  |
| Supriyatna. A dan Wahyudi. M (2012) | <i>Perancangan Data Warehouse Pada Perpustakaan Bina Sarana Informatika</i>                       | OLAP<br>Nine-Step Methodholog y | Membuat rancangan data yang dijadikan sebagai pangkalan data sebagai pengambilan keputusan dan menyajikan informasi secara multidimensi yang tepat, akurat, dan cepat. |
| Ambarwati (2017)                    | <i>Perancangan dan pembuatan Data Warehouse pada Perpustakaan STMIK AMIKOM Yogyakarta</i>         | OLAP                            | Membahas perancangan skema, database data warehouse proses loading per periode hingga ringkasan data warehouse yang diolah menjadi <i>reporting</i> .                  |
| Wulandari (2021)                    | <i>Analisis Data Warehouse Perpustakaan SD N Brengosan 1</i>                                      | OLAP                            | Mennjadikan data historis yang dapat dilihat pada periode tertentu dan digunakan sebagi pengambil keputusan  |

## 2.2 Dasar Teori

Pada bagian ini penulis akan memaparkan beberapa tinjauan sebagai dasar teori dalam kesuksesan penerapan perancangan data warehouse pada SD N Brengosan 1.

### 2.2.1 Data Warehouse

Menurut Inmon dalam Imhoff, Gallemmo dan Geiger (2003), data warehouse merupakan sekumpulan data yang berorientasi pada subjek, terintegrasi, memiliki rentang waktu dan tidak mudah berubah yang digunakan dalam mendukung pembuatan keputusan strategis. Menurut Silvers (2008), terdapat tujuh karakteristik pada sebuah data warehouse, yaitu :

1. *Subject-oriented* (Berorientasi Subjek), data warehouse didesain untuk menganalisa data berdasarkan *subject-oriented* tertentu dalam organisasi, bukan pada proses atau fungsi aplikasi tertentu.
2. *Integrated* (Terintegrasi), data warehouse dapat menyimpan data-data yang berasal dari sumber-sumber yang terpisah ke dalam suatu format yang konsisten dan saling terintegrasi satu dengan lainnya. Syarat integrasi sumber data dipenuhi dengan berbagai cara, seperti konsisten dalam penambahan variabel. Konsisten dalam ukuran variabel, konsisten dalam struktur pengkodean dan konsisten dalam atribut fisik dari data.

3. *Time Variant* (Rentang Waktu), seluruh data pada data warehouse dapat dikatakan akurat atau valid pada rentang waktu tertentu. Cara-cara yang digunakan untuk melihat interval waktu yang digunakan dalam mengukur keakuratan suatu data warehouse: a. Cara yang paling sederhana adalah menyajikan data warehouse pada rentang waktu tertentu. b. Menggunakan variasi/perbedaan waktu yang disajikan dalam data warehouse baik *implicit* maupun *explicit*, secara *explicit* dengan unsur waktu dalam hari, minggu, bulan dan waktu tertentu.
4. Variasi yang disajikan data warehouse melalui serangkaian *snapshot* yang panjang.
5. *Nonvolatile* maksudnya data pada data warehouse tidak di-*update* secara *real time* tetapi direfresh dari sistem operasional secara reguler. Data yang baru selalu ditambahkan bagi basis data itu sebagai sebuah perubahan. Basis data tersebut secara terus menerus menyerap data baru, kemudian disatukan dengan data sebelumnya. Hal ini berbeda dengan basis data operasional yang dapat melakukan *update*, *insert* dan *delete* terhadap data, sedangkan pada data warehouse hanya ada dua kegiatan manipulasi data yaitu loading data (menggambil data) dan akses data.

### 2.2.2 OLAP (*On Line Analytical Processing*)

Salah satu *tools* yang digunakan untuk mengakses informasi dalam data warehouse secara efektif untuk proses *online analysis*, memberikan respon yang cepat terhadap *analytical queries* yang kompleks (William, C. Amo.2000). Multidimensional data model dan teknik agregasi data yang dimiliki oleh OLAP dapat mengatur dan membuat kesimpulan dari data dalam jumlah besar, sehingga dapat dievaluasi secara cepat dengan menggunakan *online analysis* dan *graphical tool*. Sistem OLAP menyediakan kecepatan dan fleksibilitas untuk melakukan *support analisis* secara *real time*. OLAP Merupakan suatu metode khusus untuk melakukan analisis terhadap data yang terdapat di dalam media penyimpanan data (database) dan kemudian membuat laporannya sesuai dengan permintaan user.

Untuk tujuan tersebut data yang berupa informasi dibuat dalam format khusus. Multidimensional merupakan Suatu sistem yang menghasilkan konseptual dari data secara multidimensional, yang meliputi dukungan penuh untuk hierarki dan multiple hierarki yang merupakan cara *logic* untuk menganalisis bisnis dan organisasi. Menurut Silvers (2008), terdapat dua konsep model data pada data warehouse, yaitu:

1. Fakta (*Fact*), fakta dikenal sebagai suatu kejadian atau transaksi.

Fakta adalah sesuatu yang terjadi. Sebuah tabel fakta menggabungkan entitas yang diidentifikasi dalam *logical* data model. Beberapa tabel

fakta yang dihubungkan dengan tabel-tabel dimensi yang sama dan dipandang sebagai kumpulan dari *star schema* di dalam tabel fakta mempunyai dua tipe kolom, yaitu kolom yang menyimpan nilai-nilai *numeric* atau yang biasa disebut dengan *measure* dan kolom yang menyimpan *foreign key* yang mengacu ke tabel lainnya (tabel dimensi yang berisi fakta *numeric*). Tabel yang umumnya mengandung sesuatu yang dapat diukur (*measure*), seperti harga, jumlah barang, dan sebagainya. *Fact table* juga merupakan kumpulan *foreign key* dari *primary key* yang terdapat pada masing-masing dimension tabel. *Fact table* juga mengandung data yang historis.

2. Dimensi (*Dimensions*), dimensi adalah data yang menggambarkan kualifikasi entitas perusahaan yang terlibat dalam fakta. Hubungan antara tabel fakta dan dimensi dapat digambarkan dengan skema dimensional, skema data warehouse yang paling umum digunakan adalah skema bintang (*Star Schema*) yang terdiri dari sebuah tabel fakta yang dikelilingi oleh tabel dimensi. Dimensi tabel berelasi dengan *fact table*, tabel yang berisi data detail yang menjelaskan *foreign key* yang terdapat pada *fact table*.

### 2.2.3 OLTP

Menurut Stair dan Reynolds (2010), OLTP adalah suatu bentuk pengolahan data dimana setiap transaksi diproses dengan segera, tanpa penundaan mengumpulkan transaksi ke dalam *batch*. Memiliki karakteristik dengan jumlah data yang besar namun transaksi yang

dilakukan cukup sederhana seperti *insert*, *update*, dan *delete*. Hal utama yang menjadi perhatian dari sistem yang dilakukan OLTP adalah melakukan *query* secara cepat dan mudah untuk diperbaiki dan dapat diakses.

#### **2.2.4 Star Schema**

*Star Schema* adalah sarana untuk mengorganisir data mart–data mart atau gudang-gudang data dengan menggunakan basis data relasional. Menggunakan tabel-tabel dimensi untuk mendeskripsikan data-data yang terdapat di dalam tabel fakta.

Setiap perusahaan pada umumnya menjual produk, pengetahuan, maupun jasa. Sehingga sistem penjualan adalah sebuah sistem yang terdapat di sebagian besar perusahaan. Karakteristik utama dari skema bintang adalah bahwa tabel dimensinya tidak dinormalisasi., tabel fakta berisi data-data yang diekstrak dari database operasional.

#### **2.2.5 Pentaho Data Integration**

Pentaho adalah sebuah perusahaan yang menyediakan produk dan solusi (layanan) dalam hal *business intelligence*. Pentaho sendiri memiliki beberapa produk (aplikasi) yang memberikan kemudahan dalam mengelola data, *analisis data*, *data mining*, dan *reporting*. Salah satu aplikasi dari Pentaho yang dapat digunakan untuk mengelola data (dalam hal ini migrasi dan pemetaan basis data) dan mampu untuk melakukan proses *extract*, *transform*, dan *load* (ETL) data adalah Pentaho data integration. Aplikasi Pentaho data integration dapat

digunakan untuk melakukan migrasi data, membersihkan data, memuat data dari file ke basis data atau sebaliknya hingga volume yang besar (Wibisono, 2014). Pentaho data integration dapat melakukan proses *transformation* dan *job* dengan cara yang sangat sederhana dan intuitif, sehingga memudahkan pengguna dalam merancang skema dan memeliharanya. Adapun maksud dari *transformation* adalah sekumpulan instruksi untuk mengubah *input* data menjadi sebuah *output* yang telah disesuaikan dengan kebutuhan. Sedangkan *job* merupakan kumpulan instruksi yang diberikan untuk menjalankan proses *transformation*. Fitur utama yang dimiliki Pentaho data integration yaitu *spoon*, *pan*, dan *kitchen*. *Spoon* merupakan bagian antar muka aplikasi yang digunakan sebagai tempat untuk membuat rancangan atau instruksi untuk menjalankan proses *job* dan *transformation*. *Pan* yaitu utilitas untuk membaca, mengubah, dan menulis data. *Kitchen* yaitu utilitas yang digunakan untuk menjalankan atau mengeksekusi *job*. Adapun utilitas-utilitas yang digunakan di dalam *job* disebut dengan *step*.

### **2.2.6 Requirement Traceability Matrix**

*Requirement Traceability Matrix* (RTM) ialah tabel yang berisi daftar *requirement*, atribut yang bervariasi untuk setiap *requirement*, dan status dari *requirement* untuk memastikan semua *requirement* telah terpenuhi. Dengan RTM kita bisa melakukan *trace* terhadap setiap *requirement* yang ada, mulai dari *business requirement* sampai pada

*testing*. Ada banyak sekali template RTM yang beredar di dunia maya. Namun RTM tersebut harus disesuaikan dengan kebutuhan kita. Tujuan utama dari RTM adalah agar semua anggota *project* mengetahui *requirement* yang ada.