

Implementasi Pengembangan Sistem Model Water Fall Untuk Data Warehouse Akademik

¹Arik Sofan Tohir, ²Kusrini, ³Sudarmawan

^{1,2,3} Universitas Amikom

^{1,2,3} Sleman, Yogyakarta

E-mail: ¹arik.sofan.tohir@gmail.com, ²kusrini@amikom.ac.id,

³sudarmawan@amikom.ac.id

Abstract— Data warehouse merupakan suatu konsep dan suatu teknologi untuk menyimpan data transaksional dari beberapa sumber yang sudah melalui proses penyaringan dan pemilihan data. Dengan memanfaatkan proses Ectract, Transform and Load (ETL) pada data warehouse, data OLTP diproses untuk menghasilkan data yang baik dan siap digunakan untuk proses analisis. Untuk perancangan data warehouse ini akan dibangun dengan menggunakan Nine-Step Methodolgy dari Kimbal, sehingga data warhouse yang dihasilkan bisa sesuai dengan yang diharapkan. Sedangkan untuk pengembangan sistem menggunakan pendekatan System Developement Life Circle (SDLC) dengan model water fall. Dengan menggunakan model wate fall akan dibangun sebuah prototype untuk mengimplementasikan hasil perancangan data warehouse.

Keywords— ETL, SDLC, Water Fall, Data Warehouse

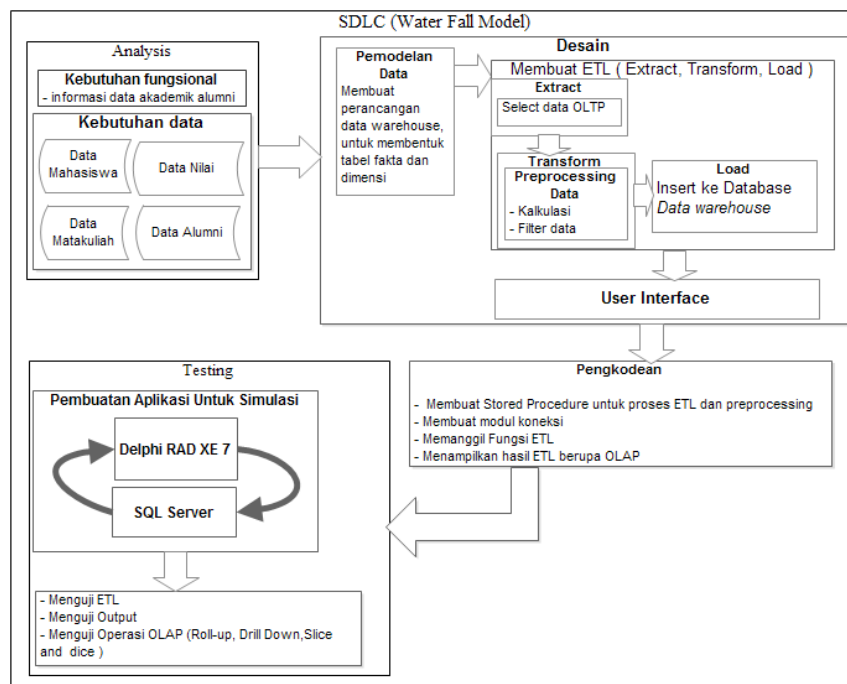
I. PENDAHULUAN

Data warehouse adalah suatu konsep dan kombinasi teknologi yang memfasilitasi organisasi untuk mengelola dan memelihara data historis yang diperoleh dari sistem atau aplikasi operasional [1]. Konsep data warehouse menjadi sangat terkenal dan banyak dipakai sekitar awal tahun 2003 [2]. Tujuan utama pembuatan data warehouse adalah untuk menyatukan data yang beragam ke dalam sebuah tempat penyimpanan, sehingga pengguna dapat dengan mudah menjalankan query (pencarian data), menghasilkan laporan, dan melakukan analisis [3]. Data history akademik mahasiswa yang selalu bertambah setiap periode, tentu nya akan sangat memerlukan waktu yang lama ketika ingin menampilkan suatu informasi terkait kondisi akademik mahasiswa dan memerlukan perintah query yang kompleks juga memerlukan perhitungan yang banyak dan digunakan untuk mendukung pengambilan keputusan. Data warehouse adalah tempat penyimpanan data terpusat, dimana data dari basis data operasional dan sumber data lainnya diintegrasikan, dibersihkan dan diarsipkan untuk mendukung pengambilan keputusan [4]. *System Development Life Circle* adalah proses mengembangkan atau mengubah suatu sistem perangkat lunak dengan menggunakan model-model

dan metodologi yang digunakan orang untuk mengembangkan sistem perangkat lunak sebelum nya (berdasarkan best practice atau cara-cara yang sudah teruji baik). Dalam hal ini penulis akan menggunakan pendekatan dengan model waterfall, *model waterfall* (air terjun) sering juga disebut model sekuensial linier (*sequential linier*) atau alur hidup klasik (*classic life circle*) [5].

II. METODE PENELITIAN

Metode yang akan digunakan untuk pengembangan sistem menggunakan pendekatan *System Development Life Circle (SDLC)* dengan model *water fall*. Sedangkan untuk perancangan *data warehouse* digunakan pendekatan metode kimbil dengan menggunakan *nine step* (sembilan langkah). Untuk implementasi pengembangan *water fall* ini akan fokus pada penerapan model untuk membangun sebuah *prototype* sistsem *data warehouse* alumni. Gambar 1 menunjukkan alur penelitian yang akan digunakan untuk mengimplementasi model *water fall* untuk *data warehouse* alumni.



Gambar 1. ALUR PENELITIAN

A. Data warehouse

Data warehouse adalah suatu koleksi data yang digunakan untuk pengambilan keputusan manajemen, yang berorientasi subjek (*Subject-oriented*), terintegrasi (*Integrated*), dimensi waktu (*time variant*) dan tidak mudah berubah (*nonvolatile*) [6].

1. Berorientasi subjek (Subject-Oriented), sebuah data warehouse terorganisasi pada subject utama, seperti pembeli, pemasok, produk dan penjualan.

2. Terintegrasi (Integrated), dibangun dengan mengintegrasikan berbagai macam sumber, seperti database relational, file dan catatan transaksi On-Line. Pembersihan data (Cleaning), dan teknik integrasi data diterapkan untuk memastikan konsistensi dalam penamaan, struktur pengkodean, pemberian atribut dan sebagainya.
3. Dimensi waktu (Time variant), data disimpan untuk menyediakan informasi dari sebuah perspektif riwayat (misal : 5-10 tahun terakhir).
4. Tidak mudah berubah (Nonvolatile).

B. *Nine Step Methodology*

1. *Choosing The Process* adalah melakukan pemilihan proses yang mengacu pada materi subjek yang dibutuhkan oleh *data mart*. Pada tahapan ini ditentukan pada proses bisnis apa *data warehouse* akan digunakan.
2. *Choosing The Grain* adalah memutuskan secara pasti apa yang diwakili atau direpresentasikan oleh sebuah tabel fakta. Pada tahap ini akan ditentukan tingkat *detail* data yang bisa didapatkan dari model dimensional.
3. *Identifying And Conforming The Dimension*, adalah membuat set dimensi yang dibutuhkan untuk menjawab seluruh pertanyaan yang diajukan pada tabel fakta
4. Pemilihan Fakta (*Choosing The Fact*) adalah pemilihan fakta dimaksudkan sebagai pemilihan tabel fakta yang dapat mengimplementasikan semua *grain* yang digunakan pada *data mart*.
5. Menyimpan Pre-Kalkulasi Di Tabel Fakta (*Storing Pre-Calculation In The Fact Table*), Setelah tabel fakta terpilih, setiap tabel fakta tersebut harus diperiksa ulang untuk menentukan apakah terdapat fakta-fakta yang dapat di terapkan pre kalkulasi dan kemudian dilakukan penyimpanan pada tabel fakta
6. Melengkapi Tabel Dimensi (*Rounding Out The Dimension Table*), pada tahap ini dilakukan pemeriksaan ulang pada tabel dimensi dan menambahkan deskripsi teks terhadap dimensi, serta menentukan hirarki atribut dimensi untuk mempermudah proses analisis.
7. Pemilihan Durasi Database (*Choosing The Duration Of The Database*) durasi berguna untuk mengukur seberapa lama tabel fakta yang dapat disimpan
8. Menelusuri Perubahan Dimensi Secara Perlahan (*Tracking Slowly Changing Dimension*), Dimensi dapat berubah secara perlahan seiring berjalannya waktu dan kebutuhan. Terdapat tiga perubahan dimensi secara perlahan, yaitu :
 - a. Atribut dimensi yang berubah dan di tulis ulang
 - b. Atribut dimensi yang berubah dan menimbulkan data dimensi baru
 - c. Atribut dimensi yang telah berubah menimbulkan alternatif sehingga nilai atribut lama dan yang baru dapat di akses secara bersamaan pada dimensi yang sama
9. Menentukan Prioritas Dan Mode Query (*Deciding The Query Priorities And The Query Modes*), Pada tahap ini masalah perancangan fisik (*physical design*) dipertimbangkan, seperti masalah

keamanan data, masalah pengarsipan data, masalah *backup data*. Masalah perancangan fisik yang paling kritis, yang dapat mempengaruhi pandangan pengguna akhir terhadap *data mart* adalah masalah pembagian fisik tabel fakta ke dalam *disk*. Extract Transform Load (ETL) merupakan proses yang penting dalam *data warehouse*, dengan ETL inilah data dari operasional dapat dimasukkan ke dalam *data warehouse*. ETL juga dapat digunakan untuk mengintegrasikan data dengan sistem sistem yang sudah ada sebelumnya. Tujuan ETL adalah mengumpulkan, menyaring, megolah dan menggabungkan data-data yang relevan dari berbagai sumber untuk disimpan ke dalam *data warehouse*. Hasil dari proses ETL adalah dihasilkannya data yang memenuhi kriteria dari *data warehouse* seperti data yang historis, terpadu, terangkum, statis, dan memiliki struktur yang dirancang untuk proses analisis. Extract merupakan langkah pertama pada proses ETL adalah mengestrak data dari sumber-sumber data.

C. *System Development Life Circle*

1. Analisis, pada tahap analisis ini dibagi menjadi dua tahapan yaitu analisis kebutuhan fungsional terkait dengan apa saja yang ingin dihasilkan dalam pembuatan *software* untuk menampilkan informasi dari *data warehouse* yang akan dibangun. Berdasarkan analisis kebutuhan fungsional maka langkah berikutnya adalah tahap menganalisis kebutuhan data yang terkait dengan pembuatan *data warehouse* untuk kebutuhan penyajian data alumni.
2. Desain, pada tahap ini merupakan tahap multi langkah yang fokus pada desain pembuatan program perangkat lunak termasuk struktur data, arsitektur perangkat lunak, representasi antar muka dan prosedur pengkodean.
3. Pengkodean dimana pada proses ini adalah mentransformasikan desain yang sudah dirancang kedalam sebuah bahasa pemrograman.
4. Testing, pada bagian ini juga akan dibuatkan simulasi untuk menguji data dan model yang sudah dibuat. Dalam pengujian ini menggunakan software Delphi RAD Studio XE 7 yang merupakan salah satu bahasa pemrograman dengan basic dari bahasa pemrograman pasca dan untuk data digunakan *database engine SQL Server*.

III. HASIL DAN PEMBAHASAN

a. Analisis

Untuk tahapan analisis dibedakan menjadi dua bagian yaitu :

1. Kebutuhan fungsional

Bertujuan untuk menggali kebutuhan fungsi yang diinginkan dari sistem yang akan dibangun dan bisa menjawab pertanyaan-pertanyaan yang kemungkinan muncul dari suatu laporan yang ada. Dalam penelitian ini kebutuhan fungsional yang di perlukan bisa menjawab pertanyaan terkait,

jumlah alumni yang lulus, periode kelulusan, IPK, asal sekolah, masa kuliah, asal sekolah alumni.

2. Kebutuhan data

Bertujuan untuk menganalisa keperluan data yang diperlukan dalam pembuatan sebuah sistem.

b. Desain

Pada tahapan desain ini yaitu membuat keperluan *data warehouse* alumni dengan menggunakan *Nine Step Methodology*.

1. *Choosing The Process*

Tabel 1. CHOOSING THE PROCESS

Proses Bisnis	Deskripsi	Fungsi yang terlibat
Penyajian data akademik	Mendata kegiatan akademik, menginputkan seluruh informasi yang terkait dengan akademik.	Bagian akademik

2. *Choosing Grain*

Tabel 2. CHOOSING GRAIN

Grain	Deskripsi	Proses bisnis yang terlibat
Informasi data akademik	Meyajikan data alumni yang bisa dilihat dari berbagai sudut yang terkait dengan informasi yang ada keterkaitannya dengan alumni meliputi jumlah kelulusan dalam periode tertentu, ipk, masa tunggu alumni untuk mendapatkan pekerjaan, alumni yang melaksanakan studi lanjut, asal sekolah	Penyajian data akademik

3. *Identifying And Conforming The Dimension*

Tabel 3. IDENTIFYING AND THE CONFORMING THE DIMENSION

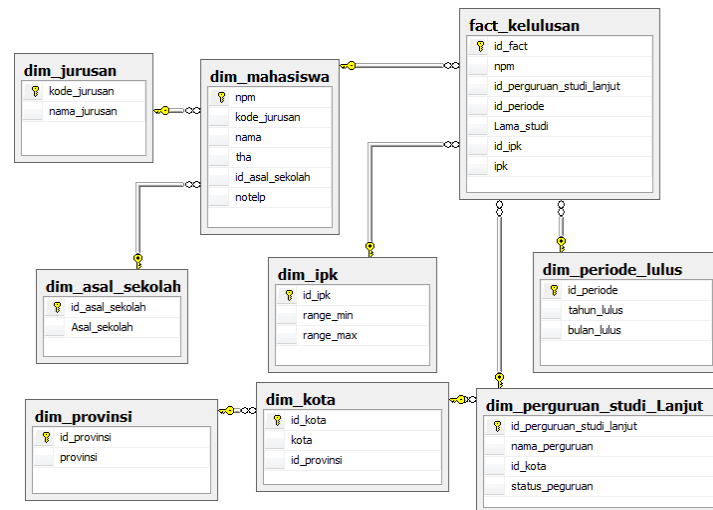
Dimensi	Deskripsi	Grain
Dim_jurusan	Dim_jurusan diidentifikasi kode_jurusan, nama_jurusan	Informasi data akademik
Dim_mahasiswa	Dim_mahasiswa diidentifikasi menjadi kode_jurusan, npm, nama, tha, notelp, id_asal_sekolah	Informasi data akademik
Dim_periode_lulus	Dim_periode_lulus diidentifikasi id_periode, tahun_lulus, bulan_lulus	Informasi data akademik
Dim_asal_sekolah	Dim_asal_sekolah diidentifikasi menjadi id_asal_sekolah, nama_asal_sekolah	Informasi data akademik
Dim_perguruan_stu di_lanjut	Dim_perguruan_studi_lanjut diidentifikasi menjadi id_perguruan_studi_lanjut,	Informasi data akademik

Dim_ipk	Dim_ipk diidentifikasi menjadi id_ipk, range_min, range_max	Informasi data akademik
---------	---	-------------------------

4. *Choosing The Fact*

Tabel 4. CHOOSING THE FACT

Fakta	Deskripsi	Dimensi
Fact_kelulusan	Fact_kelulusan merupakan tabel fakta yang menyimpan informasi terkait dengan kelulusan mahasiswa.	Dim_jurusan Dim_mahasiswa Dim_periode Dim_perguruan_studi_lanjut Dim_asal_sekolah Dim_provinsi Dim_kota Dim_ipk



Gambar 2. SNOW FLAKE DATA WAREHOUSE AKADEMIK

5. *Storing Pre-Calculation In The Fact Table*

Pada tahap ini ditentukan apakah pada tabel fakta memerlukan pre kalkulasi dan kemudian dilakukan penyimpanan pada tabel fakta. Berikut prekalkulasi yang akan ada.

- a. $ipk = \text{getIPK}(\text{dim_mahasiswa.npm})$
- b. $\text{lama_masa_studi} = \text{getMasaStudi}(\text{dim_mahasiswa.npm})$

6. *Rounding Out The Dimension Table*

Tabel 5. DIM_JURUSAN

Atribut	Data Type	Length
Kode_jurusan	Char	2
Nama_jurusan	Varchar	50

Tabel 6. DIM_PERIODE_LULUS

Atribut	Data Type	Length
id_periode	Int	11
tahun_lulus	Char	4
bulan_lulus	Char	2

Tabel 7. DIM_MAHASISWA

Atribut	Data Type	Length
Npm	Char	10
kode_jurusan	Char	10
Nama	Varchar	50
Tha	Char	4
Notelp	Varchar	15
Id_asal_sekolah	Int	11

Tabel 8. DIM_PERGURUAN_STUDI_LANJUT

Atribut	Data Type	Length
id_perguruan_studi_lanjut	int	11
nama_perguruan	Varchar	50
Provinsi	Varchar	50
Kota	Varchar	50
status_peguruan	Varchar	20

Tabel 9. DIM_ASAL_SEKOLAH

Atribut	Data Type	Length
id_asal_sekolah	int	11
Asal_sekolah	Varchar	50

Tabel 10. DIM_PROVINSI

Atribut	Data Type	Length
id_provinsi	int	11
Provinsi	Varchar	50

Tabel 11. DIM_KOTA

Atribut	Data Type	Length
id_kota	int	11
id_provinsi	Int	11
Kota	Varchar	50

Tabel 12. DIM_IPK

Atribut	Data Type	Length
id_ipk	int	11
Range_min	Float	-
Range_max	Float	-

7. *Choosing The Duration Of The Database*

Diasumsikan bahwa untuk pembuatan *data warehouse* merupakan data 5 tahun terakhir yang terbaru. Data diambil dari proses OLTP data akademik mahasiswa dan data alumni yang sudah bekerja maupun belum bekerja.

8. Tracking Slowly Changing Dimension

Tabel 13. TRACKING SLOWLY CHANGING DIMENSION

<i>Dimensi</i>	Atribut	Tipe SCD
Dim_jurusan	kode_jurusan	1
	nama_jurusan	1
Dim_asal_sekolah	Id_asal_sekolah	-
	Asal_sekolah	1
Dim_periode_lulus	id_periode	-
	tahun_lulus	1
	bulan_lulus	1
Dim_mahasiswa	npm	1
	kode_jurusan	1
	nama	1
	tha	1
	notelp	1
Dim_perguruan_studi_lanjut	id_perguruan_studi_lanjut	-
	nama_perguruan	1
	Provinsi	1
	Kota	1
	status_perguruan	1
Dim_provinsi	id_provinsi	1
	provinsi	1
Dim_kota	id_kota	-
	id_provinsi	1
Dim_ipk	Id_ipk	-
	Range_min	1
	Range_max	1

9. Deciding The Query Priorities And The Query Modes

Pada langkah ini lebih pada permasalahan dengan *physical* data warehouse terkait dengan pengurutan index, untuk data warehouse ini sudah diurutkan berdasarkan primary key pada masing-masing tabel. Selain itu terkait dengan pengarsipan data dan backup data.

a. Pengkodean

Tahap selanjut adalah melakukan pengkodean dimana pada proses ini adalah mentransformasikan desain yang sudah dirancang kedalam sebuah bahasa pemrograman

b. Testing

Untuk menguji data dan model yang sudah dibuat. Dalam pengujian ini menggunakan software Delphi RAD Studio XE 7 yang merupakan salah satu bahasa pemrograman dengan basic dari bahasa pemrograman pasca dan untuk data digunakan *database engine* SQL Server

IV. KESIMPULAN DAN SARAN

Dari hasil Implementasi Pengembangan Sistem Model *Water Fall* Untuk *Data Warehouse* Akademik dapat ditarik kesimpulan bahwa model *water fall* dapat digunakan untuk merancang prototype *data warehouse* dalam bidang akademik. Data-data transaksional akademik (OLTP) dapat dilakukan proses *Extract, Transform and Load* kedalam *data warehouse*, sehingga akan didapatkan data yang bisa digunakan untuk analisis data. Untuk pengembangan berikutnya penelitian ini bisa dilanjutkan dengan menambahkan fitur-fitur output yang sesuai dengan borang akreditasi, sehingga hasil output yang dihasilkan dapat digunakan untuk proses akreditasi

DAFTAR PUSTAKA

- [1] J. Han and M. Kamber, *Data Mining: Concepts and Techniques Second Edition*. Oxford: Morgan Kaufman Publisher, 2006.
- [2] A. . Rosa and M. Shalahuddin, *Rekayasa Perangkat Lunak*. Bandung: Modula, 2011.
- [3] Parsiyono, Kusri, and A. Sunyoto, "Perancangan Data Warehouse Akademik Di Sekolah Tinggi Agama Budha," *J. Inf.*, vol. 1, 2015.
- [4] G. Karya and A. Sandi, "Penerapan Business Intelligence untuk Analisis Data Profil Mahasiswa di Perguruan Tinggi," in *SNASTIKOM*, 2012.
- [5] Taufik, "Model Executive Information System Dengan Menggunakan Online Analytical Processing Dan Data Warehouse Bidang Akademik," *Scan*, vol. IX, no. 2, 2014.
- [6] A. Supriyatna, "Sistem Analisis Data Mahasiswa Menggunakan Aplikasi Online Analytical Processing (OLAP) Data Warehouse," *J. Pilar Nusa Mandiri*, vol. XII, no. 1, 2016.