



UNIVERSITAS
MERCU BUANA

**IMPLEMENTASI ETL (EXTRACT TRANSFORM LOAD) PADA DATA
WAREHOUSE TRANSAKSI MENGGUNAKAN TOOLS PENTHAHO DI
AEON MALL**

TUGAS AKHIR

Amelia Sari
41518310023

UNIVERSITAS
MERCU BUANA

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2022



**IMPLEMENTASI ETL (EXTRACT TRANSFORM LOAD) PADA DATA WAREHOUSE
TRANSAKSI MENGGUNAKAN TOOLS PENTAHO DI AEON MALL**

Tugas Akhir

Diajukan Untuk Melengkapi Salah Satu Syarat
Memperoleh Gelar Sarjana Komputer

Oleh:

Amelia Sari
41518310023

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS MERCU BUANA

JAKARTA

2022

LEMBAR PERNYATAAN ORISINALITAS

Yang bertanda tangan dibawah ini:

NIM : 41518310023

Nama : Amelia Sari

Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Menyatakan bahwa Laporan Tugas Akhir saya adalah hasil karya sendiri dan bukan plagiat. Apabila ternyata ditemukan didalam laporan Tugas Akhir saya terdapat unsur plagiat, maka saya siap untuk mendapatkan sanksi akademik yang terkait dengan hal tersebut.

Jakarta, 13 Juli 2022



Amelia Sari

UNIVERSITAS
MERCU BUANA

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Amelia Sari
NIM : 41518310023
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Noneksklusif** (*None-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul diatas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti/Noneksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya.

Selain itu, demi pengembangan ilmu pengetahuan di lingkungan Universitas Mercu Buana, saya memberikan izin kepada Peneliti di Lab Riset Fakultas Ilmu Komputer, Universitas Mercu Buana untuk menggunakan dan mengembangkan hasil riset yang ada dalam tugas akhir untuk kepentingan riset dan publikasi selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 13 Juli 2022



SURAT PERNYATAAN LUARAN TUGAS AKHIR

Sebagai mahasiswa Universitas Mercu Buana, saya yang bertanda tangan di bawah ini :

Nama Mahasiswa : Amelia Sari
NIM : 41518310023
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Menyatakan bahwa :

1. Luaran Tugas Akhir saya adalah sebagai berikut :

| No | Luaran | Jenis | Status |
|------------------------------|--|---|------------|
| 1 | Publikasi Ilmiah | Jurnal Nasional Tidak Terakreditasi | Diajukan ✓ |
| | | Jurnal Nasional Terakreditasi | |
| | | Jurnal International Tidak Bereputasi | Diterima |
| | | Jurnal International Bereputasi | |
| Disubmit/dipublikasikan di : | Nama Jurnal | : Jurnal Teknologi dan Sistem Komputer | |
| | ISSN | : e-ISSN : 2338-0403 | |
| | Link Jurnal | : https://jtsiskom.undip.ac.id/ | |
| | Link File Jurnal Jika Sudah di Publish | : | |

2. Bersedia untuk menyelesaikan seluruh proses publikasi artikel mulai dari submit, revisi artikel sampai dengan dinyatakan dapat diterbitkan pada jurnal yang dituju.
3. Diminta untuk melampirkan scan KTP dan Surat Pernyataan (Lihat Lampiran Dokumen HKI), untuk kepentingan pendaftaran HKI apabila diperlukan

Demikian pernyataan ini saya buat dengan sebenarnya.

Mengetahui
Dosen Pembimbing TA

Dwi Anindyani Rocmah, ST, MTI

Jakarta, 13 Juli 2022

Amelia Sari

LEMBAR PERSETUJUAN

Nama Mahasiswa : Amelia Sari
NIM : 41518310023
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tugas Akhir ini telah diperiksa dan disetujui

Jakarta, 13 Agustus 2022

Menyetujui,



(Dwi Anindyani Rocmah, ST, MTI)

Dosen Pembimbing

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI



NIM : 41518310023
Nama : Amelia Sari
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022

Runi
(Runi)

UNIVERSITAS

(Saruni Dwiasnati, ST, MM, MKOM)
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518310023
Nama : Amelia Sari
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 26 Agustus 2022



(Umniy Salamah, ST, MMSI)

UNIVERSITAS
MERCU BUANA

LEMBAR PERSETUJUAN PENGUJI

NIM : 41518310023
Nama : Amelia Sari
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 13 Agustus 2022



(Muhammad Rifqi, S.Kom, M.Kom)

UNIVERSITAS
MERCU BUANA

LEMBAR PENGESAHAN

NIM : 41518310023
Nama : Amelia Sari
Judul Tugas Akhir : Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tugas Akhir ini telah diperiksa dan disidangkan sebagai salah satu persyaratan untuk memperoleh gelar Sarjana pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Mercu Buana.

Jakarta, 25 Agustus 2022

Menyetujui,



(Dwi Anindyani Rocmah, ST, MTI)

Dosen Pembimbing

Mengetahui,



(Wawan Gunawan, S.Kom, MT)

Koord. Tugas Akhir Teknik Informatika



(Ir. Emil R. Kaburuan, Ph.D., IPM.)

Ka. Prodi Teknik Informatika

KATA PENGANTAR

Puji syukur kita panjatkan kepada Allah SWT yang telah memberikan rahmat hidayah dan karunia-Nya, Sehingga penulis dapat menyelesaikan penulisan Laporan Tugas Akhir ini. Judul yang penulis pilih dalam penulisan Laporan Tugas Akhir ini adalah “Implementasi ETL (Extract Transform Load) Pada Data Warehouse Transaksi Menggunakan Tools Pentaho Di AEON Mall

Tujuan dari penulisan laporan Tugas Akhir ini adalah sebagai syarat untuk melengkapi salah satu syarat memperoleh gelar Sarjana Ilmu Komputer di Universitas Mercubuana. Penulis menyadari bahwa tanpa bantuan dan bimbingan tidak mudah dalam menyelesaikan Laporan Tugas Akhir ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Allah swt.
2. Orang tua yang selalu memberikan dukungan dan doa sehingga penulis dapat menyelesaikan Penyusunan Laporan Tugas Akhir dengan lancar.
3. Ibu Dwi Anindyani Rochmah, ST, MTI Selaku Dosen pembimbing yang telah membimbing dalam Penyusunan Laporan Tugas Akhir ini.
4. Ibu Saruni Dwiasnati, ST, MM, M.KOM Selaku Dosen Penguji I,
5. Ibu Umniy Salamah, ST, MMSI Selaku Dosen Penguji II,
6. dan Bapak Muhammad Rifwi, S.KOM, M.KOM Selaku Dosen Penguji III.
7. Bapak dan Ibu Dosen/Staf pengajar Fakultas Ilmu Komputer Universitas Mercu Buana.
8. Serta pihak-pihak lain yang tidak dapat disebutkan satu-persatu yang telah memberikan dukungan semangat dalam Laporan Tugas Akhir.

Penulis menyadari bahwa dalam penulisan laporan Tugas Akhir ini masih terdapat kekurangan untuk itu penulis mengharapkan kritik dan saran. Akhir kata, penulis berharap semoga Laporan Tugas Akhir ini dapat bermanfaat bagi semua pihak demi perkembangan dan kemajuan akademik.

Jakarta, 13 Juli 2022



Amelia Sari

DAFTAR ISI

| | |
|---|-------------|
| HALAMAN SAMPUL | i |
| HALAMAN JUDUL | i |
| LEMBAR ORISINALITAS | ii |
| SURAT PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR | iii |
| SURAT PERNYATAAN LUARAN TUGAS AKHIR | iv |
| LEMBAR PERSETUJUAN | v |
| LEMBAR PERSETUJUAN PENGUJI | vi |
| LEMBAR PENGESAHAN | ix |
| ABSTRAK | x |
| ABSTRACT | xi |
| KATA PENGANTAR | xii |
| DAFTAR ISI | xiii |
| NASKAH JURNAL | 1 |
| KERTAS KERJA | 6 |
| BAB 1. LITERATUR REVIEW | 11 |
| BAB 2. ANALISIS DAN PERANCANGAN | 16 |
| BAB 3. SOURCE CODE | 24 |
| BAB 4. DATASET | 25 |
| BAB 5. TAHAPAN EKSPERIMEN | 26 |
| BAB 6. HASIL SEMUA EKSPERIMEN | 34 |
| DAFTAR PUSTAKA | 36 |
| LAMPIRAN DOKUMEN HAKI | 38 |
| LAMPIRAN KORESPONDENSI | 40 |

KERTAS KERJA

1) Data Warehouse

Menurut Paul Lane, data warehouse merupakan database relasional yang didesain lebih kepada query dan analisa dari pada proses transaksi, biasanya mengandung history data dari proses transaksi dan bisa juga data dari sumber lainnya[16]. Data warehouse merupakan suatu konsep dan kombinasi teknologi yang memfasilitasi organisasi untuk mengelola dan memelihara data historis yang diperoleh dari sistem atau aplikasi operasional.

Data warehouse adalah sebuah relational database yang di rancang untuk implementasi dari query dan analisa proses transaksi. Data warehouse berisikan data histori yang berasal dari data transaksi. Selain relational database, lingkungan dari data warehouse meliputi ekstraksi, transportasi, transformasi dan pemuatan (ETL) solusi, Online Analytical Process (OLAP), client analytic tools, dan aplikasi lain yang mengelola data dan memberikan pada pengguna bisnis.

Data warehouse didefinisikan sebagai sekumpulan data yang dimiliki enam buah sifat atau karakteristik berupa berorientasi subjek (subject oriented), terintegrasi (integrated), berorientasi pada proses (process oriented), time variant, dapat diakses dengan mudah (accessible), dan bersifat non volatile.

2) Karakteristik Data Warehouse

Karakteristik Data Warehouse Menurut Bill Inmon (1970) terdapat enam buah karakteristik pada data warehouse, yang meliputi Subject Oriented, Integrated, Process Oriented, Time Variant, Accesible dan Non Volatile.

- a) Subject Oriented adalah Data yang ditampilkan dan data yang disusun hanyalah data menurut subjek yang diperlukan untuk proses pengambilan keputusan. Data-data ini dirangkum ke dalam bentuk dimensi, yang meliputi periode waktu (time), riwayat (history), wilayah (region), dan lain-lain.
- b) Integrated adalah Data warehouse dibangun dari proses integrasi berbagai sumber data (terutamanya sumber data berupa database) yang berasal dari

berbagai aplikasi (software), menjadi satu kesatuan yang utuh. Karakteristik integrated muncul sebagai bentuk lain dari karakteristik subject oriented.

- c) Time Variant adalah Data warehouse (yang dikumpulkan dari berbagai sumber data dari berbagai aplikasi) diidentifikasi dari periode waktu penyimpanannya. Atau dengan kata lain, proses penyimpanan data-data yang dikumpulkan tersebut, berdasarkan kepada waktu. Adapun informasi yang disajikan oleh data-data pada data warehouse tersebut, dilihat sudut pandang riwayat penyimpanan (historical point of view).
- d) Non Volatile adalah Data-data dari berbagai sumber data yang dikumpulkan ke dalam data warehouse, tidak boleh mengalami manipulasi data dalam bentuk edit, update dan delete. Dengan kata lain, data warehouse mementingkan adanya historical data (data asli), guna menunjang proses analisis data yang dilakukan kelak.
- e) Process Oriented Data warehouse dipandang sebagai sebuah proses berkesinambungan di dalam pengolahan data menjadi informasi serta pengiriman informasi tersebut. Proses menjadi orientasi dari data warehouse di dalam operasionalnya.
- f) Accessible Data warehouse beserta dengan data-data di dalamnya, harus dapat diakses dengan mudah oleh pengguna. Pengguna dapat memperoleh data yang mana saja yang mereka butuhkan, baik keseluruhan maupun sebagian (parsial), sesuai dengan hak akses (privilege) yang diartikan oleh sistem atau pemilik sistem.

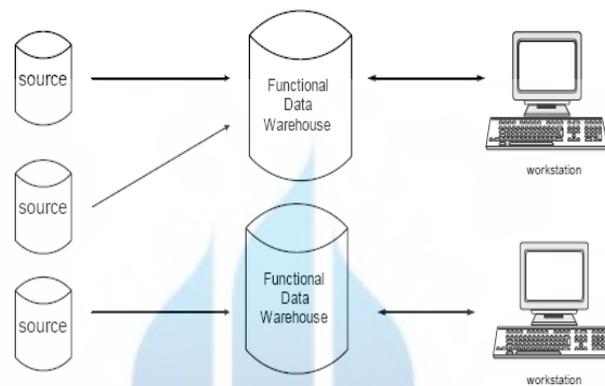
Menurut Ken Orr (1999), "Data Warehouse Architecture (DWA) atau arsitektur data warehouse merupakan sebuah cara yang dilakukan oleh arsitek data warehouse atau SDM yang berkaitan dengan perancangan dan desain sistem data warehouse di dalam sebuah organisasi, untuk mempresentasikannya ke dalam bentuk bagan/gambar/desain, yang memuat segala hal mengenai struktur data, komunikasi, pemrosesan, dan presentasi dari komputasi end to end antar komputer di dalam sistem data warehouse".

3) Arsitektur Data Warehouse

Berikut ini adalah tiga jenis Arsitektur Data Warehouse :

a. Functional Data Warehouse (Data Warehouse Fungsional)

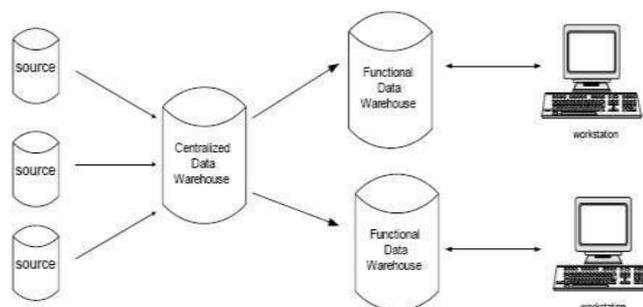
Kata operasional disini merupakan database yang diperoleh dari kegiatan sehari-hari. Data warehouse dibuat lebih dari satu dan dikelompokkan berdasar fungsi-fungsi yang ada di dalam perusahaan seperti fungsi keuangan(financial),marketing,personalia dan lain-lain.



Gambar 2.1 Bentuk data warehouse fungsional

b. Centralized data warehouse (data warehouse terpusat)

Bentuk ini terlihat seperti bentuk data warehouse fungsional, namun terlebih dahulu sumber data dikumpulkan dalam satu tempat terpusat, kemudian data disebar ke dalam fungsinya masing-masing, sesuai kebutuhan perusahaan. Data warehouse terpusat ini, biasa digunakan oleh perusahaan yang belum memiliki jaringan eksternal.



Gambar 2.2 Bentuk data warehouse data warehouse terpusat

c. Distributed data warehouse (data warehouse terdistribusi)

Pada data warehouse terdistribusi ini, digunakan gateway yang berfungsi sebagai jembatan penghubung antara data warehouse dengan workstation yang menggunakan sistem beraneka ragam. Dengan sistem terdistribusi seperti ini memungkinkan perusahaan dapat mengakses sumber data yang berada diluar lokasi perusahaan(eksternal).



Gambar 2.3 Bentuk data warehouse terdistribusi

4) Pengertian ETL (Extract, Transform dan Load)

Extract, Transform dan Load secara sederhana merupakan suatu kumpulan dari proses untuk mengambil maupun memproses data dari satu atau banyak sumber sehingga menjadi sumber baru, yaitu seperti mengolah data OLTP menjadi OLAP. Proses-prosesnya ialah sebagai berikut:

1. Extract

Proses-proses yang diperlukan untuk terhubung dengan berbagai macam sumber data, dan membuat data tersebut tersedia bagi proses-proses selanjutnya, dengan contoh mengambil data dari database, membaca file excel, membaca file dari XML, dan lain sebagainya.

2. Transform

Fungsi apa saja yang berfungsi untuk mengubah data yang masuk menjadi data yang dikehendaki. Transform adalah proses manipulasi terhadap data dari

suatu sistem sumber ke format lain pada data warehouse atau pada data mart sehingga menjadikan sebuah informasi yang bermakna .

3. Load

Semua proses yang diperlukan dalam mengisi data ke target atau dapat dikatakan proses yang akan memindahkan data yang telah ditransformasikan sebelumnya ke data warehouse. Strategi loading ke integration layer dibagi menjadi dua, yaitu strategi loading bagi tabel dimensi dan strategi loading bagi tabel fakta.

5) Pentaho

Pentaho Data Integration (PDI) atau kettle adalah software dari pentaho yang dapat digunakan untuk proses ETL (Extraction, Transformation, dan Loading). PDI dapat digunakan untuk migrasi data, membersihkan data, loading dari file ke database atau sebaliknya dalam volume besar. PDI menyediakan graphical user interface dan drag-drop komponen yang memudahkan user. Elemen utama dari PDI adalah Transformation dan job. Transformation adalah sekumpulan instruksi untuk merubah input menjadi output yang diinginkan (input-proses-output). Sedangkan job adalah kumpulan instruksi untuk menjalankan transformation. Ada tiga komponen dalam PDI yaitu spoon, pan, dan kitchen. Spoon adalah user interface untuk membuat job dan transformation. Pan adalah tools yang berfungsi membaca, merubah, dan menulis data. Sedangkan kitchen adalah program yang mengeksekusi job.(Wibisono 2014).