



**ANALISIS DAN PERANCANGAN *DATA WAREHOUSE*  
AKADEMIK PERGURUAN TINGGI BINA SRIWIJAYA BERBASIS  
*PENTAHO DATA INTEGRATION (PDI) KETTLE***

**PROPOSAL PENELITIAN**

Sebagai salah satu syarat untuk mengikuti ujian proposal skripsi  
Program Studi Sistem Informasi

**DEFI SEPRIANUS  
09142278**

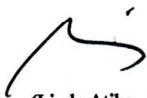
**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS BINA DARMA  
PALEMBANG  
2013**

**HALAMAN PENGESAHAN****ANALISIS DAN PERANCANGAN *DATAWAREHOUSE*  
AKADEMIK PERGURUAN TINGGI BINA SRIWIJAYA BERBASIS  
*PENTAHO DATA INTEGRATION (PDI) KETTLE*****OLEH :****DEFI SEPRIANUS  
09142278****PROPOSAL SKRIPSI**

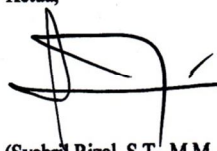
Telah diterima sebagai salah satu syarat untuk mengikuti  
Ujian Proposal

Palembang, Mei 2013  
Program Studi Teknik Informatika  
Universitas Bina Darma  
Ketua,

Pembimbing I



(Linda Atika., M.Kom)



(Syahril Rizal, S.T., M.M., M.Kom.)

Pembimbing II



(Siti Sa'uda, M.Kom)



UNIVERSITAS BINA DARMA PALEMBANG  
 FAKULTAS ILMU KOMPUTER  
 PROGRAM STUDI TEKNIK INFORMATIKA



LEMBAR KONSULTASI PROPOSAL SKRIPSI

Nama : Defi Seprianus  
 Nim : 09.142.278  
 Fakultas : Ilmu Komputer  
 Program Studi : Teknik Informatika  
 Judul : Analisis Dan Perancangan Data Warehouse Akademik  
 Perguruan Tinggi Bina Sriwijaya Berbasis Pentaho  
 Data Integration (PDI) Kettle.  
 Pembimbing I : Linda Atika, M.Kom.

No	Tanggal	Keterangan	Paraf
1	24/5/2013	revisi proposal : - Bahasa asing - format penulisan - data warehouse	
2	28/5/2013	revisi proposal dengan 1 kali ulangan.	



UNIVERSITAS BINA DARMA PALEMBANG  
 FAKULTAS ILMU KOMPUTER  
 PROGRAM STUDI TEKNIK INFORMATIKA



LEMBAR KONSULTASI PROPOSAL SKRIPSI

Nama : Defi Seprianus  
 Nim : 09.142.278  
 Fakultas : Ilmu Komputer  
 Program Studi : Teknik Informatika  
 Judul : Analisis Dan Perancangan Data Warehouse Akademik  
 Perguruan Tinggi Bina Sriwijaya Berbasis Pentaho  
 Data Integration (PDI) Kettle.  
 Pembimbing II : Siti Sa'uda, M.Kom.

No	Tanggal	Keterangan	Paraf
1.	3/2013 /5	Revisi proposal sesuai petunjuk	
2.	13/2013 /5	Revisi proposal sesuai petunjuk	
3.	15/2013 /5	Revisi proposal sesuai petunjuk	
4.	17/2013 /5	Revisi proposal sesuai petunjuk	
5.	21/2013 /5	Revisi proposal sesuai petunjuk	
6.	23/2013 /5	Acc Proposal	

**PROPOSAL PENELITIAN**  
**ANALISIS DAN PERANCANGAN *DATAWAREHOUSE***  
**AKADEMIK PERGURUAN TINGGI BINA SRIWIJAYA BERBASIS**  
***PENTAHO DATA INTEGRATION (PDI) KETTLE***

## **I. PENDAHULUAN**

### **I.1. Latar Belakang**

Informasi yang cepat dan akurat sangat dibutuhkan dalam kehidupan sehari-hari. Informasi akan menjadi suatu elemen penting dalam perkembangan masyarakat saat ini dan waktu mendatang, namun kebutuhan informasi yang tinggi kadang tidak diimbangi dengan penyajian informasi yang memadai. Sering kali informasi tersebut masih harus di gali ulang dari data yang jumlahnya sangat besar.

Pangkalan Data Perguruan Tinggi (PDPT) merupakan informasi yang di sediakan oleh Direktorat Pendidikan Tinggi mengenai kegiatan akademik seluruh perguruan tinggi di Indonesia. Untuk mendapat kan informasi PDPT, Direktorat Pendidikan Tinggi menugaskan Koordinator Perguruan Tinggi Swasta (KOPERTIS) sebagai perpanjangan tangan untuk memperoleh informasi mengenai PDPT dari Perguruan Tinggi Swasta di wilayahnya.

Bina Sriwijaya (BINAS) merupakan yayasan penyelenggara pendidikan di Sumatera Selatan. Dalam pelaksanaannya sebagai perguruan tinggi swasta di bawah naungan Kopertis Wilayah II, Binas melakukan Kegiatan rutin setiap semester yaitu pelaporan data PDPT ke Kopertis (Koordinator Perguruan Tinggi Swasta) wilayah II. Keterlambatan dalam melaporkan data PDPT dapat berpengaruh terhadap izin perguruan tinggi, kendala yang timbul adalah

bagaimana mengelola data yang terdapat di perguruan tinggi bina sriwijaya menjadi data yang dibutuhkan oleh PDPT untuk dilaporkan secara cepat dan akurat, dengan data yang banyak harus dapat di olah dalam bentuk PDPT. Dalam pelaksanaannya data dari perguruan tinggi bina sriwijaya masih dilakukan dengan cara memasukkan data ke dalam sistem PDPT yang disediakan oleh Kopertis satu-persatu dari data mahasiswa, data matakuliah, data nilai, data dosen, data kelulusan data perguruan tinggi, data aktif mahasiswa dan data dosen. Dengan begitu banyak nya data yang harus di masukkan hanya oleh satu operator maka timbul kendala yaitu sering terjadi keterlambatan dan kesalahan dalam pelaporan ke PDPT.

*Datawarehouse* merupakan penyimpanan data tetap sebagai implementasi fisik dari pendukung keputusan model data. *Data warehouse* juga biasanya dilihat sebagai arsitektur, pembangunan dan penyatuan data dari bermacam macam sumber data yang berbeda untuk mendukung struktur dan atau *query* tertentu, laporan analisis, dan pembuatan keputusan (Han, 2006).

Untuk menghasilkan *datawarehouse* dapat menggunakan *Extract, transform, dan load (ETL)* merupakan sebuah sistem yang dapat membaca data dari suatu *data store*, merubah bentuk data, dan menyimpan ke *data store* yang lain. *Data store* yang dibaca *ETL* disebut *data source*, sedangkan *data store* yang disimpan *ETL* disebut target.

Aplikasi *ETL* tersedia di dalam *Pentaho Data Integration (PDI) Kettle*. *Pentaho Data Integration (PDI) Kettle* adalah *software Open Source* dari Pentaho yang dapat digunakan untuk mengintegrasikan data. *Kettle* menyediakan fasilitas *ETL (Extraction, Transformation dan Loading)*. *Kettle* dapat digunakan untuk

membersihkan data, *loading* dari *file* ke *database* atau sebaliknya dan migrasi antar aplikasi dalam volume besar. Versi komersial dari *Kettle* adalah *Pentaho Data Integration* (PDI).

Dari uraian diatas peneliti ingin melakukan penelitian yang diambil berdasarkan latar belakang yang berjudul “**ANALISIS DAN PERANCANGAN DATAWAREHOUSE AKADEMIK PERGURUAN TINGGI BINA SRIWIJAYA BERBASIS PENTAHO DATA INTEGRATION (PDI) KETTLE**”

## **I.2. Perumusan Masalah**

Berdasarkan latar belakang yang telah di uraikan diatas, maka permasalahan yang dapat dirumuskan yaitu “Bagaimana cara Analisis Dan Perancangan *Datawarehouse* Akademik Perguruan Tinggi Bina Sriwijaya Berbasis *Pentaho Data INTEGRATION* (PDI) *Kettle* ?”

## **I.3. Batasan Masalah**

Agar penelitian lebih terarah dan memudahkan dalam pembahasan, maka perlu adanya pembatasan masalah, yaitu:

1. Merancang *data warehouse* menggunakan *tools Pentaho Data Integration* (PDI) *Kettle*.
2. Data penelitian berasal dari data Akademik untuk pelaporan PDPT
  - a. Tabel mahasiswa
  - b. Tabel mata kuliah
  - c. Tabel dosen
  - d. Tabel jadwal kuliah
  - e. Tabel transaksi kuliah mahasiswa

- f. Tabel nilai mahasiswa
- g. Tabel kelulusan mahasiswa
- h. Tabel transaksi dosen mengajar
- i. Tabel status mahasiswa
- j. Tabel profile perguruan tinggi

#### **I.4. Tujuan Penelitian dan Manfaat Penelitian**

##### **I.4.1. Tujuan Penelitian**

Menghasilkan *datawarehouse* untuk kebutuhan pelaporan PDPT bagi perguruan tinggi bina sriwijaya.

##### **I.4.2. Manfaat Penelitian**

Adapun manfaat yang diharapkan pada penelitian ini adalah:

###### 1. Bagi Penulis

Penulis dapat lebih mengetahui cara menerapkan ilmu-ilmu yang telah dipelajari selama ini dalam merancang dan menerapkan aplikasi sistem dengan teknik *data warehouse*, serta sebagai syarat dalam memperoleh gelar sarjana komputer.

###### 2. Bagi Manajemen

Model yang diperoleh diharapkan dapat digunakan untuk untuk kebutuhan pelaporan PDPT bagi perguruan tinggi bina sriwijaya.

## **II. TINJAUAN PUSTAKA**

### **2.1. Tinjauan Umum**

#### **2.1.1. Sejarah Singkat**



Bina Sriwijaya berdiri tahun 1999, yang didirikan oleh Irian Nasri dimana Bina Sriwijaya (BINAS) singkatan Bina Sriwijaya / Bimbingan Irian Nasri Pertamakali bertempat di Jln. Kebon Jahe No. 555 Palembang, dan sesuai dengan perkembangannya mulai dari Komputer 10 Unit dan sampai mencapai Ratusan Unit, serta pindah Jl. Merdeka NO. 3E Menempati Ruko 1 Unit, dengan liku-liku Perjuangan yang panjang akhirnya Berdiri Akademik Manajemen Informatika dan Komputer Bina Sriwijaya SK NO. 76/D/2008 dan sekarang Telah Terakreditasi BAN PT KEMDIKNAS 2011.

Yayasan Pendidikan Indonesia "Bina Sriwijaya" Akte Notaris Tahun 2001 dan didaftarkan di DEPHUMKAM No. 18 Tanggal 21 Juni Tahun 2010 dengan Dewan Pembina : H. Irian Nasri, M.pd Ketua Yayasan : Hj. Miaty Hermawati, M.Pd Pengawas Harian : Indra Wijaya, S.Si yang berkedudukan di Jl. A. Yani No.720, A,B,C.D 7 Ulu Palembang, Sumsel.

Yayasan Pendidikan Bina Sriwijaya (BINAS) Palembang, menyelenggarakan Pendidikan Tinggi yaitu Akademik Komputer BINA SRIWIJAYA terdiri dari 2 Program Studi yaitu Manajemen Informatika dan Teknik Komputer, dan Tahun 2012 Ini Yayasan Juga mengembangkan Sekolah dibawah naungan Diksispora Kota Palembang :

1. SD BSI (BINA SRIWIJAYA INDONESIA) Palembang
2. SMP BSI (BINA SRIWIJAYA INDONESIA) Palembang
3. SMK BSI (BINA SRIWIJAYA INDONESIA) Palembang
4. SMA BSI (BINA SRIWIJAYA INDONESIA) Palembang

### **2.1.2. Struktur Organisasi**

Tabel 2.1 Struktur Organisasi Amik Bina Sriwijaya

<b>NO</b>	<b>NAMA</b>	<b>JABATAN</b>	<b>PENDIDIKAN</b>
1.	H. Irian Nasri	Ketua Pembina	S-2
2.	Hj. Miaty Hermawati	Ketua Yayasan	S-2
3.	Indra Wijaya	Pengawas Yayasan	S-1
4.	Zet Abdullah	Pengawas Pendidikan	S-1
5.	Ibnu Aqil	Direktur	S-2
6.	Yuli Alam	Pembantu Direktur I	S-2
7.	Dahlia Novita	Pembantu Direktur Ii	S-1
8.	Novri Hardinata	Pembantu Direktur Iii	S-1
9.	M. Zulkipli Ah	Pembantu Direktur Iv	S-1
10.	Frans Habrizons	Ka. Upm	S-2
11.	Nurul Huda	Kaprodi. Manajemen Informatika	S-2
12.	Al Amin	Kaprodi. Teknik Komputer	S-1
13.	Ef Niar	Ka. Lppm	S-2
14.	Devi Andriansyah	Ka. Humas	S-1
15.	Dewi Yanti	Ka. Bsec	S-1
16.	Nanda Tri Haryati	Ka. Bstc	S-1
17.	Hormaini	Ka. Laboratorium	S-1
18.	R. A. Fitria	Ka. Bau	S-1
19.	Dahlia Novita	Ka. Baak	S-1
20.	Hendriansyah	Ka. Perpustakaan	S-1
21.	M. Nur Hidayat	Ka. Security	S-1

## 2.2. Landasan Teori

### 2.2.1. *Data Warehouse*

*Data warehouse* adalah sistem yang mengambil dan menggabungkan data secara periodik dari sistem sumber data ke penyimpanan data bentuk dimensional atau normal (Rainardi, 2008).

*Data Warehouse* adalah penyimpanan data tetap sebagai implementasi fisik dari pendukung keputusan model data. *Data warehouse* juga biasanya dilihat sebagai arsitektur, pembangunan dan penyatuan data dari bermacam-macam sumber data yang berbeda untuk mendukung struktur dan atau *query* tertentu, laporan analisis dan pembuatan keputusan (Han, 2006).

*Data warehouse* bukan merupakan *software* maupun *hardware* yang bisa kita beli untuk menghasilkan suatu keputusan strategis, melainkan suatu lingkungan dimana *user* bisa menemukan suatu informasi yang strategis atas kumpulan data-data yang dimiliki, yang bersifat *Integrated*, *Subject-Oriented*, *Nonvolatile*, dan *Time-Variant*. *Subject Oriented* berarti kita bisa mempelajari persoalan berdasarkan persoalan (*subject*) yang di inginkan saja. *Integrated* berarti data yang terdapat didalam sebuah data *warehouse* dapat berasal dari beberapa sumber terpisah yang nantinya ke semua data tersebut akan disimpan ke dalam satu bagian yang sama dengan suatu format khusus. *Nonvolatile* berarti data-data yang berada di dalam sebuah data *warehouse* tidak dapat dilakukan perubahan lagi (*edit*). *Time-Variant* berarti data-data yang tersedia dalam jumlah besar dapat dikatakan akurat atau valid sampai pada waktu-waktu tertentu saja.

Berdasarkan uraian di atas, maka penulis dapat menyimpulkan definisi dari *Data Warehouse* adalah koleksi data yang mempunyai sifat berorientasi subjek,

terintegrasi, *time-variant*, dan bersifat tetap dari koleksi data dalam mendukung proses pengambilan keputusan manajemen.

*Extract, transform, dan load (ETL)* merupakan sebuah sistem yang dapat membaca data dari suatu *data store*, merubah bentuk data, dan menyimpan ke *data store* yang lain.

Proses *ETL (Extraction, Transformation, dan Loading)* merupakan proses yang harus dilalui dalam pembentukan *data warehouse*, yaitu:

1. *Extraction*

Ekstraksi dilakukan dari sumber data yang digunakan melalui proses pemilihan data yang kemudian disimpan pada basis data sementara. Dalam penelitian ini, basis data sementara ditempatkan pada penyimpanan basis data, mesin dan *platform* yang sama dengan yang digunakan untuk *data warehouse* nantinya.

2. *Cleaning*

Proses *cleaning* adalah proses untuk membersihkan data yang rangkap, tidak konsisten, rusak, tidak benar (bernilai salah), data yang hilang (kosong), atau data dengan spasi yang berlebihan.

3. Transformasi

Proses transformasi yang dilakukan dibagi berdasarkan dua *level* yaitu *record level* dan *field level*. Pada proses ini dilakukan proses pemilihan, penggabungan dan agregasi untuk mendapatkan data ringkasan sesuai dengan dimensi yang akan dibuat. Proses transformasi ke *field* yang baru dapat dilakukan dengan (menggunakan *script*) tertentu untuk melakukannya.

#### 4. *Loading*

Proses terakhir yang perlu dilakukan adalah proses pemuatan data (*loading*). Data yang digunakan pada tahap ini merupakan data dari proses-proses yang dilakukan sebelumnya yaitu ekstraksi, pembersihan (*cleaning*) dan transformasi untuk kemudian dimasukkan ke dalam *data warehouse*. Cara pemuatan data ke dalam *data warehouse* adalah dengan menggunakan *script* yang dijalankan secara periodik.

##### 2.2.2. *MySQL*

Menurut Bunafit Nugroho (2004:1) *MySQL* adalah sebuah program *database server* yang mampu menerima dan mengirimkan datanya dengan sangat cepat, multi *user* serta menggunakan perintah standar SQL (*Structured Query Language*). *MySQL* merupakan sebuah *database server* yang *free*, artinya bebas menggunakan *database* ini untuk keperluan pribadi atau usaha tanpa harus membeli atau membayar lisensinya. *MySQL* juga berperan sebagai *client* sehingga sering disebut *database client/server*, yang *open source* dapat berjalan baik di OS (*Operating System*) manapun, dengan *Platform Windows* maupun *Linux*.

##### 2.2.3. *Pentaho Data Integration (PDI) Kettle*

*Pentaho Data Integration (PDI) Kettle* adalah *software Open Source* dari Pentaho yang dapat digunakan untuk mengintegrasikan data. *Kettle* menyediakan fasilitas *ETL (Extraction, Transformation dan Loading)*. *Kettle* dapat digunakan untuk membersihkan data, *loading* dari *file* ke *database* atau sebaliknya dan migrasi antar aplikasi dalam volume besar. Versi komersial dari *Kettle* adalah *Pentaho Data Integration (PDI)*.

### **2.3. Penelitian Sebelumnya**

Pada penelitian sebelumnya, penulis menampilkan penelitian yang berjudul “Perancangan Model *Data Warehouse* Dalam Mendukung Keputusan Jasa Pengiriman” penelitian ini dilakukan oleh Mahasiswa Universitas Bina Nusantara Fakultas Ilmu Komputer Program Studi Sistem Informasi yang bernama Tanty Oktavia. Tujuan penelitian ini adalah untuk *support* proses analisis pada perusahaan. Hasil berupa *datawarehouse* pada perusahaan jasa pengiriman.

Penelitian sebelumnya yang kedua, “Perancangan dan Pembuatan *Data Warehouse* Untuk Kebutuhan Sistem Pendukung Keputusan di Bidang Akademik Pada Jurusan Sistem Informasi, ITS, Surabaya” penelitian ini dilakukan oleh Mahasiwa Institut Teknologi Sepuluh November Fakultas Teknologi Informasi jurusan Sistem Informasi yang bernama Restia Rezalini P.S, Wiwik Anggraeni dan Radityo Prasetianto Wibowo. Tujuan penelitian ini adalah membuat *data warehouse* untuk aplikasi *dashboard* pada ITS.

## **III. METODOLOGI PENELITIAN**

### **3.1. Waktu dan Tempat Penelitian**

#### **3.1.1. Waktu Penelitian**

Waktu penelitian dilakukan pada bulan Mei 2013 sampai dengan bulan Agustus 2013.

#### **3.1.2. Tempat Penelitian**

Lokasi penelitian ini dilakukan penulis di Perguruan Tinggi Bina Sriwijaya yang beralamat di Jalan Jl. A. Yani No. 720 ( 7 Ulu) Palembang.

#### **3.1.3. Alat dan Bahan**

Dalam pembuatan analisis dan merancang *data warehouse* Perguruan Tinggi Bina Sriwijaya, alat dan bahan yang digunakan meliputi *hardware*, *software* serta bahan-bahan penunjang lainnya.

#### **3.1.4. Perangkat Keras (*Hardware*)**

Perangkat keras yang digunakan adalah laptop dengan spesifikasi berikut :

- a. *Laptop Toshiba Satellite U 400*
- b. *RAM 3GB*
- c. *Hardisk 250 GB*
- d. *Intel Pentium dual core processor T4200 @2.00Ghz*
- e. *Flash Disk 2 GB*

#### **3.1.5. Perangkat Lunak (*Software*)**

- a. *Windows 7 Professional* sebagai *operating system*
- b. *Microsoft Office 2007* untuk penulisan laporan tugas akhir ini
- c. *Software* pendukung yaitu *xampp (mysql)*, *Pentaho Kettle*.

#### **3.1.6. Metode Penelitian**

Metode pada penelitian ini menggunakan metode deskriptif yaitu penelitian tentang gejala dan keadaan yang dialami sekarang oleh subjek yang sedang diteliti, dan fenomena yang terjadi saat sekarang (ketika penelitian berlangsung) dan menyajikan apa adanya. (Subana dan Sudrajat, 2005 : 26-27).

#### **3.4. Metode Pengumpulan Data**

Metode Pengumpulan Data yang digunakan dalam penelitian ini adalah :

- a. Wawancara

Dalam metode ini penulis mengumpulkan data penelitian dengan bertanya langsung kepada pihak yang bersangkutan.

b. Kepustakaan

Mengumpulkan data dengan cara mencari dan mempelajari data-data dari buku-buku ataupun dari referensi lain yang berhubungan dengan penulisan proposal penelitian.

c. Observasi

Dalam hal ini yang dilakukan adalah melihat serta mempelajari permasalahan pembuatan *data warehouse* secara konkrit dan nyata yang ada di lapangan, juga mengumpulkan data-data pada Perguruan Tinggi Bina Sriwijaya.

### **3.5. Data Warehouse**

Metodologi perancangan *data warehouse* yang digunakan dalam penelitian ini adalah pemodelan data dimensional. Menurut Powell (2006), langkah-langkah membuat model data dimensional dimulai dari *end-user*. Karena *end-user* adalah pemakai informasi hasil *query data warehouse*. Dari perspektif tersebut, langkah-langkah dalam merancang data dimensional menurut Powell adalah sebagai berikut.

1. Proses bisnis (*Business processes*)

- a. Menentukan dan menggambarkan subyek area bisnis dari *functional requirement* yang ada.
- b. Subyek dianalisis sebagai dasar menentukan tabel fakta pada langkah selanjutnya.



## 2. *Granularity*

- a. *Granularity* adalah tingkat dari rincian kebutuhan.
- b. Pilihan yang paling aman menurut Powell adalah menyertakan semua data historis pada level yang terendah.
- c. Keuntungan menyertakan data level terendah adalah menghindari data yang hilang ketika dibutuhkan manajer eksekutif.
- d. Keuntungan lain adalah menghindari kekurangan data ketika dibutuhkan bahan analisis di masa datang meskipun untuk saat ini belum digunakan.

## 3. Identifikasi dan membentuk dimensi (*Identify and build dimensions*)

- a. Menentukan data-data yang dibutuhkan untuk mendukung subyek untuk dapat dibentuk dalam beberapa tabel dimensi.
- b. Tabel dimensi nantinya akan mendeskripsikan tabel fakta dengan menyimpan detail transaksi tabel fakta.

## 4. Membentuk fakta (*Build fact*), pembentukan fakta atas dasar subyek yang telah ditentukan.