

PERANCANGAN DATA WAREHOUSE IPK MAHASISWA STUDI KASUS: PRODI PTIK UNIVERSITAS NEGERI MANADO

Irene R.H.T. Tangkawarow, ST., MISD. CIBIA

¹⁾ Pendidikan Teknologi Informasi dan Komunikasi Universitas Negeri Manado
Jl. Kampus Fatek Unima Tonsaru Tondano Sulawesi Utara Indonesia
email : irene.tangkawarow@gmail.com

ABSTRACT

Support data warehouse expected to solve the evaluation problems of teaching and learning outcomes as well as the information relevance received as a support in decision-making by leadership (executive) level. The data warehouse is important to designed utilize the existing information resources. IPK (Indeks Prestasi Kumulatif) data warehouse can help the process of evaluation, decision making even further planning of PTIK study program. Diversity of data source in PTIK study program made decision-making and evaluation not easy. Pentaho Data Integration used for integrate data in PTIK easily. IPK data warehouse design with multidimensional database modeling approach by using dimension table and fact table.

Key words

Data warehouse, Pentaho Data Integration, multidimensional, dimension table, fact table.

1. Pendahuluan

Penggunaan teknologi informasi yang terintegrasi dengan proses pekerjaan pada suatu institusi atau perusahaan sudah menjadi kebutuhan yang mutlak saat ini. Hal ini disebabkan oleh adanya kebutuhan dari institusi untuk mengoptimalkan kemampuan untuk menganalisa masalah-masalah yang dihadapi yang nantinya akan berpengaruh dalam proses pengambilan keputusan. Ketersediaan data yang lengkap dan tepat merupakan tolak ukur dari kelangsungan hidup suatu institusi.

Salah satu indikasi keberhasilan suatu program studi adalah dengan adanya data IPK mahasiswa yang dapat diakses dengan cepat, tepat dan akurat. Masalah yang terjadi saat ini pada umumnya terletak pada proses input data-data beragam sehingga menyebabkan keterlambatan pemrosesan data. Penelitian ini diharapkan dapat menanggulangnya dengan adanya menggunakan

integrasi data dengan menggunakan *Pentaho Data Integration- Kettle*. Aplikasi ini dapat mengintegrasikan data yang siap diolah di dalam *data warehouse* sehingga nantinya dapat disajikan dengan akurat dan tepat waktu. Proses integrasi ini dikenal dalam *Business Intelligence* sebagai proses *Extract Transformation Loading (ETL)*. Proses ETL akan mengubah data *On-Line Transactional Processing (OLTP)* menjadi data *On-Line Analytical Processing (OLAP)*.

Dalam penelitian ini akan dibangun *datawarehouse* untuk menampilkan IPK mahasiswa khususnya pada Prodi PTIK. Universitas Negeri Manado berdiri sejak tahun 1955 yang awalnya merupakan IKIP Manado. Prodi PTIK sendiri berdiri sejak tahun 2010.

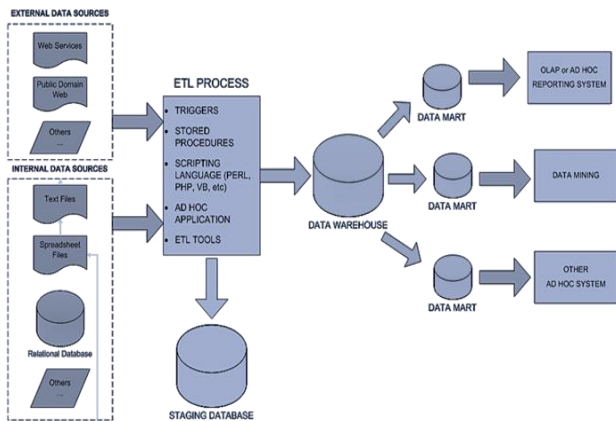
Data yang akan digunakan adalah data mahasiswa dan data IPK mahasiswa. Diharapkan penelitian ini dapat memberikan data yang cepat, tepat dan akurat terlebih untuk kebutuhan evaluasi, pengambilan keputusan dan perencanaan pengembangan prodi kedepan.

2. Perancangan Data Warehouse

2.1 Data Warehouse

Data Warehouse adalah suatu database khusus yang digunakan sebagai “gudang data” atau data yang telah terkonsolidasi dari sumber-sumber data berbagai sistem informasi yang ada pada suatu organisasi/perusahaan. Menurut Kimball, ada beberapa requirement untuk *data warehouse*, beberapa diantaranya adalah:

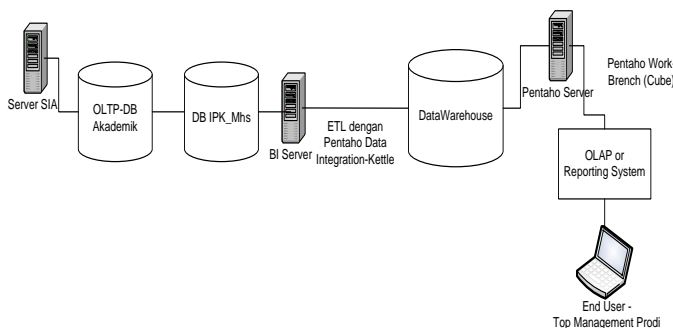
- Data warehouse harus membuat informasi dari suatu perusahaan/institusi/organisasi mudah untuk diakses.
- Data warehouse harus menampilkan informasi perusahaan/institusi secara konstan.
- Data warehouse harus menyajikan data yang nantinya akan dijadikan dasar atau pedoman untuk pengambilan keputusan.



Gambar 1. Arsitektur Data Warehouse (Kimball,2002)

2.2 Perancangan Arsitektur Data Warehouse PTIK.

Data *source* yang akan diolah diambil dari database akademik Prodi PTIK. Data tersebut berisi semua data akademik dari mahasiswa yang ada di prodi PTIK yang biasanya akan di *upload* per semester untuk ditampilkan pada Sistem Informasi Akademik Universitas Negeri Manado. Pemilahan data dilakukan untuk memilah data apa saja yang akan digunakan tanpa mengganggu data operasional yang sementara digunakan. Proses *staging* ini dilakukan juga untuk memudahkan proses ETL nantinya, karena data yang diambil benar-benar hanyalah data yang dibutuhkan untuk Data Warehouse IPK mahasiswa. Berikut adalah desain arsitektur data warehouse prodi PTIK:



Gambar 2. Arsitektur Fisik Data Warehouse Prodi PTIK

Untuk sumber data (*data source*) menggunakan Microsoft Access dan Ms, Excel sedangkan untuk datawarehouse setelah selesai proses ETL akan menggunakan MySQL.

2.3 Pemodelan Data Dimensional

Pemodelan multidimensional database akan terdiri dari **tabel fakta** (*fact tables*) dan **tabel dimensi** (*dimension tables*) yang saling terkait. Suatu tabel fakta berisi

berbagai nilai agregasi yang menjadi dasar pengukuran (*measure*) serta beberapa key yang terkait ke tabel dimensi yang akan menjadi sudut pandang dari *measure* tersebut.

Struktur dari fact table dan dimension table ini memiliki skema perancangan sehingga dapat meningkatkan performa dan kemudahan dalam penerjemahan ke sistem OLAP. Schema inilah yang menjadi dasar untuk melakukan data warehousing. Dua schema yang paling umum digunakan oleh berbagai OLAP *engine* adalah skema bintang (*Star Schema*) dan skema butir salju (*Snowflake Schema*). Dalam penelitian ini penggambaran data multidimensional menggunakan *Star Schema*.

Tabel dimensi merupakan tabel yang berisi data yang menunjukkan hasil tinjauan dari berbagai sudut pandang. Tabel dimensi nantinya akan menyusun *cube*. Table dimensi yang ada yaitu:

Tabel dim_mhs

Tabel dimensi ini berisi data mahasiswa yang hanya dibatasi pada nim, nama tempat dan tanggal lahir, jenis kelamin dan tahun masuk.

dim_mhs	
PK	sk_mhs
	nim
	nama_mhs
	tempat_lahir
	tanggal_lahir
	jk
	thn_masuk
	kd_prodi

Tabel dim_dosen

Tabel dimensi ini berisi data dosen yang hanya dibatasi pada nama dosen, dan kode dosen. Untuk kode dosen merupakan nidn.

dim_dosen	
PK	sk_dosen
	kd_dosen
	nip
	nama_dosen
	Gelar_dosen
	TempatLahir
	TanggalLahir
	JK
	No_KTP
	Kd_Prodi
	Kd_MK
	Nama_MK

Tabel dim_mk

Tabel dimensi ini berisi data mata kuliah, bobot sks dari matakuliah tersebut dan di semester berapa mk tersebut ada.

dim_mk	
PK	sk_mk
	kd_mk nama_mk sks semester GanjilGenap Kd_Dosen

Tabel dim_waktu

Tabel dimensi ini berisi data waktu berupa hari, kuartal, semester, bulan, tahun dan tanggal.

dim_waktu	
PK	sk_waktu
	hari kuartal semester bulan tahun tanggal

Tabel dim_ipk

Tabel dimensi ini berisi data index prestasi kumulatif (IPK) mahasiswa pada setiap tahun akademik dan semester ganjil maupun genap.

dim_ipk	
PK	sk_ipk
	nim tahun_akademik ganjilgenap ip_semester sks_semester ip_kumulatif sks_kumulatif

Tabel dim_nilai

Tabel dimensi ini berisi data nilai mahasiswa pada setiap matakuliah setiap tahun akademik dan semester ganjil maupun semester genap.

dim_nilai	
PK	sk_nilai
	nim tahun_akademik ganjilgenap kd_mk kelas nilai_akhir (0-4) grade(A-E)

Tabel *fact* merupakan tabel yang berisi fakta-fakta bisnis, umumnya merupakan tabel rincian transaksi yang telah terjadi (Mulyana, 2015). *Fact Table* yang dirancang dalam data warehouse ini diambil berhubungan dengan data evaluasi dari Prodi PTIK. Satu fact table *fact_ipk* saja sebenarnya apa sudah cukup menampilkan data ipk dari tahun ke tahun. Namun dianggap perlu juga adanya informasi tentang nilai dari tiap mahasiswa. Untuk itulah, ditentukan dua *fact table* yaitu *fact_ipk* dan *fact_nilai*.

Tabel fact_nilai

Tabel ini berisi semua data nilai mahasiswa PTIK pada tiap matakuliah yang sudah di kontrak.

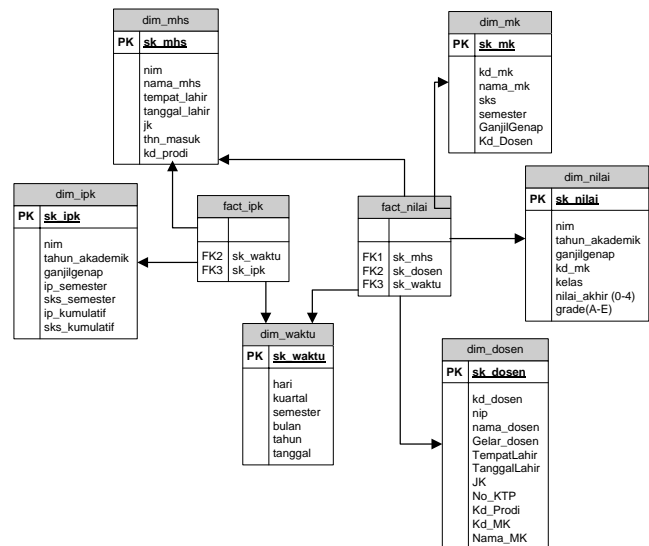
fact_nilai	
FK1	sk_mhs
FK2	sk_dosen
FK3	sk_waktu

Tabel fact_ipk

Tabel ini berisi semua data ipk mahasiswa PTIK disetiap tahun akademik berdasarkan semester ganjil dan genap.

fact_ipk	
FK2	sk_waktu
FK3	sk_ipk

Berikut ini gambaran star schema dari ipk prodi dan nilai mahasiswa.



Gambar 3. Star Schema Data Warehouse IPK Prodi PTIK

3. Hasil Percobaan

3.1 Pentaho Data Integration

Pentaho Data Integration (PDI) atau **Kettle** adalah utilitas ETL *open source* di bawah Pentaho Corp. Amerika. Proyek ini awalnya merupakan inisiatif dari **Matt Casters**, seorang programmer dan konsultan Business Intelligence (BI) dari Belgia yang telah menangani berbagai proyek BI

untuk perusahaan besar. Saat ini Kettle merupakan utilitas ETL yang sangat populer dan salah satu yang terbaik di pasaran. Beberapa kelebihanannya adalah sebagai berikut:

- Memiliki koleksi modul pengolahan data yang cukup banyak. Lebih dari 100 modul atau *step*
- Memiliki modul yang memudahkan perancangan model *data warehouse* seperti *Slowly Changing Dimension* dan *Junk Dimensions*.
- Performa dan kemampuan skalabilitas yang baik dan sudah terbukti.
- Dapat dikembangkan dengan berbagai *plugin* tambahan

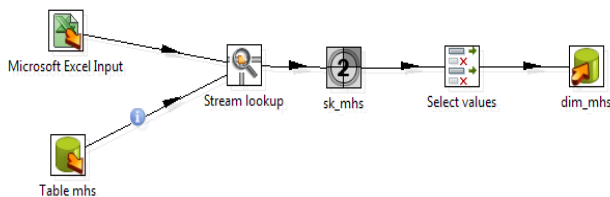
Utility Kettle yang akan digunakan dalam integrasi data ini menggunakan *Spoon*.

3.2. Integrasi Data untuk Table Dimensi

Dengan menggunakan PDI- Kettle, keberagaman data yang ada pada prodi PTIK dapat diintegrasikan menjadi database dengan *platform database* yang digunakan adalah MySQL. Tabel-tabel dimensi inilah yang akan membentuk *Data Warehouse* IPK untuk Prodi PTIK.

Tabel Dimensi Mahasiswa (Dim_Mhs)

Data source dari table *dim_mhs* ini berasal dari daftar nama mahasiswa yang terdaftar sebagai mahasiswa pada tahun ajaran 2014 semester ganjil (1). Dimana total mahasiswa yang terdaftar adalah 1293. Namun data ini hanyalah berupa nama dan NIM saja yang disimpan dalam file Excel. Sedangkan dalam data *warehouse* membutuhkan data mahasiswa yang lengkap, antara lain tempat lahir, tanggal lahir, jenis kelamin, tahun masuk, dsb. Oleh karena itu, kolaborasi data dengan data induk yang digunakan pada PDPT (Pangkalan Data PT) perlu dilakukan. Data PDPT tersebut merupakan data Ms.Access. Penggabungan data source dari file Excel dan data pada Ms.Access dapat dilakukan dengan menggunakan PDI ini. Berikut ini gambaran *transformation design* untuk menggunakan PDI-Kettle.



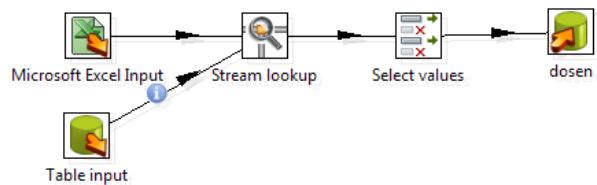
Gambar 4. Transformasi untuk tabel dim_mhs

Setelah transformasi dilakukan dihasilkan tabel dimensi mahasiswa sebagai berikut:

pk_mhs	NIM	Nama	TempatLahir	TanggalLahir	TahunMasuk	JK
127414215175		NIHEL KOLONIAM	TONJANO	19971004	2014	L
127514215176		DIYANTI EKAPUTRI KUNIA	PAPAYONGAN	19941207	2014	P
127614215177		KUNTHA LINDA	KAYUNDA	19960117	2014	P
127714215179		IRFAN MOHAMAD ZEIN BINA	TASTIMALAYA	19960407	2014	L
127814215180		ANDRE K M RUMENGAN	JAKARTA	19960526	2014	L
127914215181		KARITHA WIDYAYANTI TANJALO	GORONTALO	19951115	2014	P
128014215182		DEBALDO MOLORE WIDURWANAN	PITALING	19960323	2014	L
128114215183		STEFANUS ROBOG	KUNDA	19960509	2014	L
128214215184		KEVIN FRANCISCO KARANGAN	MARADO	19930922	2014	L
128314215185		RAYONO PONDAM	MARADO	19931107	2014	L
128414215186		RIZKY FERY SOMENGO	TONJONO	19970205	2014	L
128514215187		RAHMALUD STEFAN FERR	TONJANO	19960923	2014	L
128614215188		GRANIFIE G.E MORINGGA	TONJANO	19960628	2014	L
128714215189		JOHANNY PRADIRYO LEMPAS	KARANGSIANG	19960517	2014	L
128814215190		LEONARDI RONGGOWO	TIKOP	19960405	2014	L
128914215191		HEMCH VAN DE MORIS BUDONI	SEKI	19960217	2014	L
129014215192		INDRA DAVID RESIK	MOTOLING SATU	19950722	2014	L
129114215193		RESON TERBAK	BEKKA	19930205	2014	L
129214215194		TOMAS EFERE	OTOMBA	19971006	2014	L
129314215195		JAMUS SEARO	PECELANA	19971128	2014	L

Tabel Dimensi Dosen (dim_dosen)

Data dosen yang ada pada prodi (*Data Source*) adalah data yang tersimpan dalam file excel. Penggabungan data dilakukan untuk memperoleh data dosen yang lengkap.



Gambar 5. Transformasi untuk tabel dim_dosen

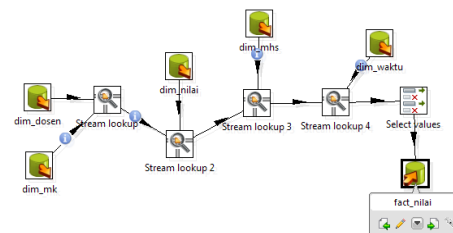
Setelah transformasi dilakukan dihasilkan tabel dimensi dosen sebagai berikut:

pk_dosen	nama_dosen	nama_dosen	detail_dosen	tempatLahir	TanggalLahir	
1	200005650	200005650	KEPTAMUN LIMBONG TAMBARA	M.SI	RAJEBRAB	19830803
2	3000410301	132320080	OTTO C DALIDE	ST., Mh.	PADEUSABAN	19840606
3	4000705306	130900982	OLIVIA ERNIKE SELVIE LIANDU	S.T.	MARADO	19931006
4	5000927309	132308448	HERY ERDIE POLLI	Dev.	MIRAMALA	19830507
5	6000907904	132206171	FRISKY YUSRIKA TOBAS	(DUL)	MARADO	19750709
6	7000910401	-	BENCI DOLY LOW	S.T.	MIRAMALA	19700709
7	8001114212	131789679	GLADLY CAREY BORNHARDY	ST. NISD	MARADO	19841009
8	9001204204	131414890	MORTHE YULFIRE PANGEMAN	Dev.	MIRAMALA	19621111
9	10013017907	132324806	NEKETEPE SAMPALTA	Dev.	TERAKTE	19200612
10	11013094211	131597016	JANNE DEVI TICOM	(DUL)	PADEUSABAN	19750513
11	12014149406	130513536	RIO H ANAST	Dev.	TANJURA	19620913
12	13014110105	130613536	CIRES TIONAN TONAS	(DUL)	MIRAMALA	19840104
13	14010125303	131448314	ROBERT LITA	M.SI	SOGONG	19511014
14	15010104611	132321741	RAM DOLFER BOGASARA	M.PS	KATAPAN	19931218
15	160140448304	(DUL)	DOOR ANDRIEF	S.Ps	PAJO	19840516
16	17010145007	131789686	YVVI P. RANTUNG	ST. NISD	(DUL)	19930416
17	180101046205	131597016	CHRISTINE T H MAKUPU	Dev.	MARADO	19650118
18	19010117109	131113200811001	BENNER EL MORGUNTA	M.SI	SOGONG	19200919
19	20020038701	(DUL)	REVI HARLANI WIRNO PARADISI	ST. M.Eng	Yung Pandang	19711119
20	21020078912	131599110	MARIO T. PARINSI	S.Kom, MTI	(DUL)	19700320
21	220202046111	131697907	LUCKIE BOONE	M.Ps	MIRAMALA	19890720
22	23024074604	132099012	RIKKA W. MORGUNGA	Dev.	MIRAMARAD	19610622
23	24020308609	131448313	JULIEN BIRIHMAN	M.SI	MARADO	19660724
24			JAMES JEFFREY SUNAWO	Dev.	MIRAMALA	19600325

Gambar 6. Tabel data dari dim_dosen

Tabel Fact Nilai (fact_nilai)

Fact tabel dibuat dalam transformasi dengan menggabungkan tabel dimensi yaitu *dim_dosen*, *dim_mk*, *dim_mhs*, *dim_nilai*, *dim_waktu* yang sudah didesain sebelumnya.



Gambar 7. Transformasi ETL untuk tabel fact_nilai

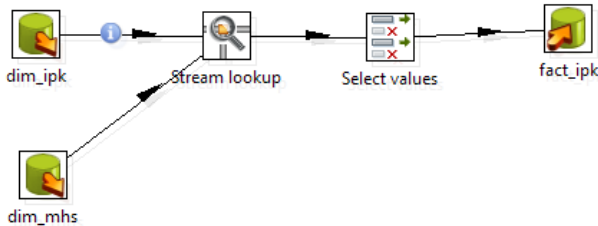
Tabel fact_nilai setelah transformasi dijalankan adalah sebagai berikut:

sk_mhs	sk_mk	sk_dosen	sk_waktu	sk_nilai
21	65	18	20160102	1
21	67	22	20160102	2
21	66	32	20160102	3
21	(NULL)	(NULL)	20160102	4
21	(NULL)	(NULL)	20160102	5
21	(NULL)	(NULL)	20160102	6
(NULL)	66	32	20160102	7
(NULL)	(NULL)	(NULL)	20160102	8
(NULL)	67	22	20160102	9
(NULL)	66	32	20160102	10
203	44	29	20160102	11
203	78	7	20160102	12
203	(NULL)	(NULL)	20160102	13
203	54	17	20160102	14
203	64	4	20160102	15
203	56	3	20160102	16
203	80	20	20160102	17
203	(NULL)	(NULL)	20160102	18
203	81	19	20160102	19
203	66	32	20160102	20
203	43	24	20160102	21

Gambar 8. Table data fact_nilai

Tabel Fact Nilai (fact_ipk)

Fact tabel dibuat dalam transformasi dengan menggabungkan tabel-tabel dimensi (dim_ipk dan dim_mhs) yang sudah didesign sebelumnya.



Gambar 9. Transformasi ETL untuk tabel fact_ipk

Tabel fact_ipk setelah transformasi dijalankan adalah sebagai berikut:

Nama	JK	TahunMasuk	SKSeme...	IPSemester	SKSKumulatif	IPKumulatif
FADLY FENDRY GUNDONG	L	2011	18	3	133	3.01
FRISILIA FRELINA IRENE ENGKOL	P	2011	(NULL)	(NULL)	(NULL)	(NULL)
ANVELIKA CLARA LOIS SASUWUK	P	2011	18	3	132	2.94
GRACE CILLYA LEGESANG	P	2011	20	2.35	136	2.98
JEANE ANITA TAROREH	P	2011	12	2.85	134	3.28
JESI MARCELLA SEMBUNG	P	2011	(NULL)	(NULL)	(NULL)	(NULL)
DENNIS IMANUEL HASEGER	L	2011	(NULL)	(NULL)	(NULL)	(NULL)
JOSTA JOEY BARLEDO MAKAENAS	L	2011	13	1.38	122	2.41
FRICKY JORDY NAJOAN	L	2011	18	3	124	2.35
BRANDO JESRI SEMUEL MENTANG	L	2011	18	3	124	2.79
RIBCA M. MENGGO	P	2011	13	1.62	124	2.7
BILLY HANURIAN SUMILAT	L	2011	13	2.15	123	2.88
RICHARD BAERHEL MANOI	L	2011	13	2.92	134	3.03
MARINI BERNADET MANARISIP	P	2011	14	3.5	136	3.32
KRISTOVEL FERDINANTY TAMUNTUAN	L	2011	18	3	136	2.59
SARTIKA KANNA	P	2011	(NULL)	(NULL)	(NULL)	(NULL)
FRANS HALIM TELAUMBANUA	L	2011	18	3	128	3
JEANE BASTIANA MONTOH	P	2011	4	1.5	116	2.79
PITER LOKON	L	2011	18	3	128	3
MICHAEL HAIGLEN SENDUK	L	2011	(NULL)	(NULL)	(NULL)	(NULL)
DIESLY FREYSILIA VIOLITA REPP	P	2011	13	2.15	128	2.51
MARSELLA NIKITA LANTANG	P	2011	19	2.89	135	3.03
DELANO MARTIN ULARAN	L	2011	18	3	132	2.44

Gambar 10. Tabel data fact_ipk

4. Kesimpulan

Keberagaman data yang menjadi sumber data yang ada pada prodi PTIK dapat diatasi dengan menggunakan Pentaho Data Integration (PDI)-Kettle. Hasil perancangan IPK Data warehouse yang dihasilkan menjadikan data dapat dilihat dalam bentuk tabel yang lebih tertata sehingga mudah diolah. Data inilah yang siap diolah menjadi Pentaho schema workbench dan selanjutnya dashboard dapat disajikan dengan menggunakan Pentaho Business Intelligence (BI) Server.

REFERENSI

- [1] "CIBIA Courseware", 2015, Multimatics, Jakarta.
- [2] JRP, Mulyana., 2014, "Pentaho: Solusi Open Source untuk Membangun Data Warehouse", ANDI, Yogyakarta.
- [3] Kimball, Ralph., Ross, Margy., 2002, "The Data Warehouse Toolkit Second Edition", John Wiley & Sons, Inc, Canada.
- [4] Kimball, Ralph., Caserta, Joe., 2004, "The Data Warehouse ETL Toolkit", John Wiley & Sons, Inc, Canada.[online].
- [5] Thia Feris, "Pentaho Knowledge Based", [online], (<http://pentaho-en.phi-integration.com>)

Irene R.H.T. Tangkawarow, memperoleh gelar **ST** pada Universitas Katolik De La Salle Manado tahun 2006 dan pada tahun 2012 memperoleh gelar **MISD** dari Hoogheschool van Arnhem en Nijmegen Netherland. Baru saja diawal tahun 2015 tertarik dengan Business Intelligence dengan menggunakan aplikasi Pentaho sehingga mengikuti TOT pada Multimatics Jakarta sehingga bisa disertifikasi dengan gelar **CIBIA** (Certified International Business Intelligence Associate). Saat ini sebagai Staf Pengajar program studi Pendidikan Teknologi Informasi dan Komunikasi (PTIK) Universitas Negeri Manado.