Perancangan Data Warehouse pada Perpustakaan

STMIK AMIKOM Yogyakarta

Armadyah Amborowati
STMIK AMIKOM Yogyakarta
armagauthama@yahoo.com

Abstraksi

Data merupakan aset penting dalam sebuah organisasi yang digunakan untuk melakukan suatu strategi atau mengambil suatu keputusan. Proses pengolahan data dapat dilakukan diberbagai tempat, misalkan di database operasional, aplikasi operasional, maupun teknologi *data warehouse*.

Metode yang digunakan dalam penelitian ini adalah fase pengembangan *data warehouse*. Penelitian ini bertujuan merancang *data warehouse* di perpustakaan STMIK AMIKOM Yogyakarta. Hasil yang didapat adalah Diperlukan pengkodean khusus untuk id dosen dan id mahasiswa pada database data warehouse dan *Field* yang diperlukan pada tabel waktu sudah disesuaikan dengan kebutuhan manajemen perpustakaan.

Kata Kunci: *data warehouse*, pengambilan keputusan, *star schema*

1. Pendahuluan

1.1 Latar Belakang

Data merupakan aset penting dalam sebuah organisasi yang digunakan untuk memutuskan kebijakan, melakukan strategi, atau mengambil keputusan. Data sebelum digunakan akan diolah dulu menjadi suatu informasi yang dibutuhkan oleh organisasi. Proses pengolahan data dapat dilakukan diberbagai tempat, misalkan di database operasional, aplikasi operasional, maupun menggunakan teknologi *data warehouse*.

Data warehouse adalah suatu konsep dan kombinasi teknologi yang memfasilitasi organisasi untuk mengelola dan memelihara data historis yang

diperoleh dari sistem atau aplikasi operasional [Ferdiana, 2008]. Pemakaian teknologi *data warehouse* hampir dibutuhkan oleh semua organisasi, tidak terkecuali Perpustakaan. *Data warehouse* memungkinkan integrasi berbagai macam jenis data dari berbagai macam aplikasi atau sistem. Hal ini menjamin mekanisme akses "satu pintu bagi manajemen untuk memperoleh informasi, dan menganalisisnya untuk pengambilan keputusan.

1.2 Perumusan masalah

Berdasarkan latar belakang yang telah dikemukakan, maka permasalahan yang dapat dirumuskan adalah bagaimana merancang *star schema* untuk *data warehouse* pada Perpustakaan STMIK AMIKOM .

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk merancang *data warehouse* pada Perpustakaan STMIK AMIKOM Yogyakarta. Hasil yang diharapkan adalah sebagai berikut.

1. Rancangan *star schema* untuk data warehouse Perpustakaan STMIK AMIKOM Yogyakarta.
2. Database *data warehouse*.

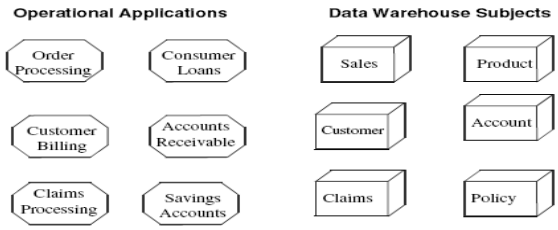
2. Landasan Teori

Data Warehouse

Data warehouse adalah suatu paradigma baru dilingkungan pengambilan keputusan strategik. *Data warehouse* bukan suatu produk tetapi suatu lingkungan dimana user dapat menemukan informasi strategik [Poniah, 2001, h.14]. *Data warehouse* adalah kumpulan data-data logik yang terpisah dengan database operasional dan merupakan suatu ringkasan. Adapun karakteristik dari *data warehouse* [Poniah, 2001,h.20-24] adalah sebagai berikut.

1. Berorientasi subyek

Data warehouse adalah tempat penyimpanan berdasakan subyek bukan berdasakan aplikasi. Subyek merupakan bagian dari suatu perusahaan. Contoh subyek pada perusahaan manufaktur adalah penjualan, konsumen, inventori, daln lain sebagainya. Gambar dibawah ini merupakan perbedaan mengenai *data warehouse* dan database operasional.



Sumber: Poniah, 2001,h.21

Gambar 1. perbedaan *data warehouse* dan database operasional

Untuk lebih jelasnya mengenai perbedaan antara database operasional dengan *data warehouse* bisa dilihat pada tabel 1 dibawah ini.

Tabel 1 Perbedaan database operasional dan *data warehouse*

	Database operasional	*Data warehouse*
Isi data	Bernilai sekarang atau up-to-date	Arsip, *history*, rangkuman
Struktur data	Dioptimasi untuk transaksi, normalisasi	Dioptimasi untuk *query* yang kompleks, Unnormalisasi
Frekuensi akses	Tinggi	Sedang-rendah
Tipe akses	Read, update, delete	Read
Penggunaan	Update secara terus menerus	Update secara periodik
Users	Banyak	Lebih sedikit

Sumber: Poniah, 2001,h.11.

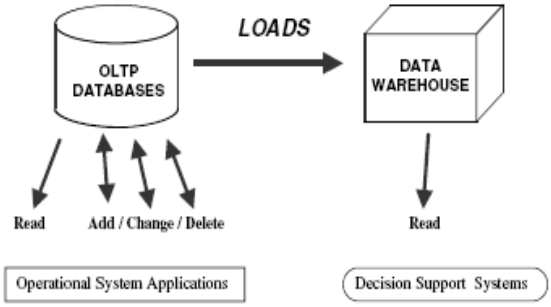
2. Data yang terintegrasi

Sumber data yang ada dalam *data warehouse* tidak hanya berasal dari database operasional (*internal source*) tetapi juga berasal dari data diluar

sistem (*external source*). Data pada sumber berbeda dapat di-*encode* dengan cara yang berbeda. Sebagai contoh, data jenis kelamin dapat di-*enkode* sebagai 0 dan 1 di satu tempat dan ”m” dan ”f” di tempat lain.

3. Nonvolatile

Data dalam database operasional akan secara berkala atau periodik dipindahkan kedalam *data warehouse* sesuai dengan jadwal yang sudah ditentukan. Misal perhari, perminggu, perbulan, dan lain sebagainya. Sekali masuk ke dalam *data warehouse*, data adalah *read-only* . Pada gambar 2 dibawah ini bisa dilihat bahwa database OLTP bisa dibaca, diupdate, dan dihapus. Tetapi pada database *data warehouse* hanya bisa dibaca.



Sumber: Poniah, 2001,h.24

Gambar 2 *Data warehouse* adalah *nonvolatile*

4. Time-Variant

Sistem operasional mengandung data yang bernilai sekarang sedangkan data dalam *data warehouse* mengandung data tidak hanya data terkini tetapi juga data *history* yang akan digunakan dalam analisis dan pengambilan keputusan. Waktu adalah dimensi penting yang harus didukung oleh semua *data warehouse*. Data untuk analisis dari berbagai sumber berisi berbagai nilai waktu, misalkan harian, mingguan, dan bulanan.

5. Ringkas

Jika diperlukan, data operasional dikumpulkan ke dalam ringkasan-ringkasan.

6. Granularity

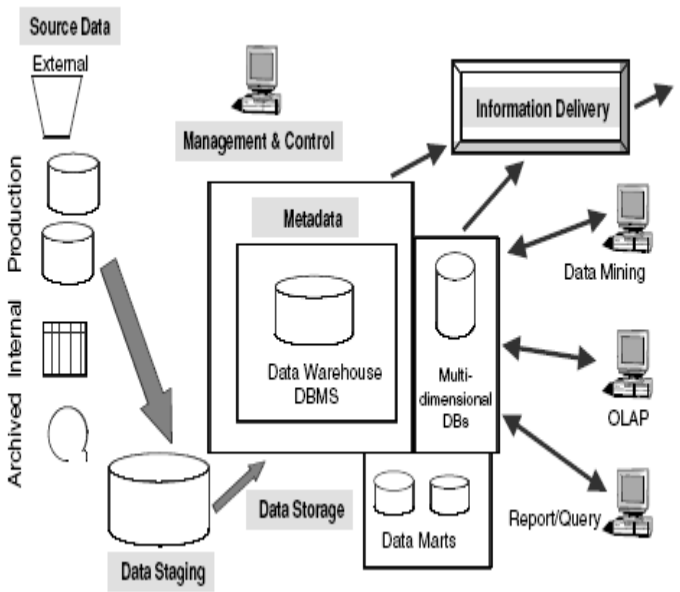
Pada sistem operasional data dibuat secara *real-time* sehingga untuk mendapatkan informasi langsung dilakukan proses *query*. Pada *data warehouse* pada menganalisis harus memperhatikan *level-of-detail* misalkan perhari, ringkasan perbulan, ringkasan per-tiga-bulan.

7. Tidak ternormalisasi

Data di dalam sebuah *data warehouse* biasanya tidak ternormalisasi dan sangat redundan.

Dasar dari suatu *data warehouse* adalah suatu data yang besar yang mengandung informasi bisnis. Data-data yang ada di dalam *data warehouse* bisa berasal dari banyak sumber, misalkan dari database operasional atau transaksional dan sumber dari luar misalkan dari web, penyedia jasa informasi, dari perusahaan lain, dan lain sebagainya. *Data warehouse* mengandung beberapa elemen penting antara lain [Mallach, 2000,h.473]:

1. Sumber data yang digunakan oleh *data warehouse*, database transaksional dan sumber data eksternal.
2. Proses ETL (*Extraction, Transformation, Loading*) dari sumber data ke database *data warehouse*.
3. Membuat suatu ringkasan atau *summary* terhadap *data warehouse* misalkan dengan menggunakan fungsi agregat.
4. Metadata.
   Metadata mengacu data tentang data. Metadata menguraikan struktur dan beberapa arti tentang data, dengan demikian mendukung penggunaan efektif atau tidak efektif dari data.
5. Database *data warehouse*.
   Database ini berisi data yang detail dan ringkasan data dari data yang ada di dalam *data warehouse*. Karena *data warehouse* tidak digunakan dalam proses transaksi individu, maka databasenya tidak perlu diorganisasikan untuk akses transaksi dan untuk pengambilan data, melainkan dioptimisasikan untuk pola akses yang berbeda di dalam analisis.
6. *Query Tools* yaitu dengan OLAP (*Online Analytical Processing*) dan data mining. Tool untuk *query* ini meliputi antarmuka pengguna akhir dalam mengajukan pertanyaan kepada database, dimana proses ini disebut sebagai *On-line Analytical Processing* (OLAP). Tool ini juga terdiri dari tool otomatis yang menemukan pola-pola di dalam data, yang sering disebut sebagai *data mining*. *Data warehouse* harus memiliki salah satu dari kedua tipe ini atau malah kedua-duanya.
7. User.
   Pengguna yang memanfaatkan *data warehouse* tersebut.

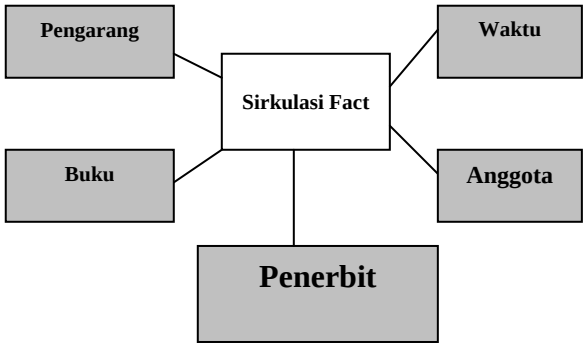


Sumber: Poniah, 2001, h. 29.

Gambar 3 Arsitektur *Data Warehouse*

Star Schema

Pada sistem OLTP (*Online Transactional Processing*) digunakan suatu teknik pemodelan data yang disebut sebagai E-R (*Entity-Relationaship*). Pada *data warehouse* digunakan teknik pemodelan data yang disebut *dimensional modelling technique.* Pemodelan dimensional adalah suatu model berbasis pemanggilan yang mendukung akses *query* volume tinggi. *Star Schema* adalah alat dimana pemodelan dimesional diterapkan dan berisi sebuah tabel fakta pusat. Tabel fakta berisi atribut deskriptif yang digunakan untuk proses *query* dan *foreign key* untuk menghubungkan ke tabel dimensi. Atribut analisis keputusan terdiri dari ukuran performa, metrik operasional, ukuran agregat, dan semua metrik yang lain yang diperlukan untuk menganalisis performa organisasi. Tabel fakta menunjukkan apa yang didukung oleh *data warehouse* untuk analisis keputusan. Tabel dimensi mengelilingi tabel fakta pusat. Tabel dimensi berisi atribut yang menguraikan data yang dimasukkan dalam tabel fakta. Tabel dimensi menunjuk bagaimana data akan dianalisis. Contoh *Star Schema* ditunjukkan pada gambar 4.



Gambar 4 Contoh *Star Schema*

Keterangan gambar.

= Tabel Fact

= Tabel Dimension

Tabel Dimensi Waktu

Tabel dimensi waktu adalah suatu tabel dimensi yang harus ada dalam sebuah *data warehouse* karena setiap *data warehouse* adalah *time series*. Waktu merupakan dimensi pertama yang harus diperhatikan dalam proses *sort order* dalam suatu database karena ketika hal itu pertama kali dilakukan, *loading* data secara berturut-turut dalam interval waktu tertentu akan masuk kedalam tempat yang sebenarnya dalam sebuah *disk*.

Data designer menyatakan secara eksplisit mengenai kebutuhan tabel dimensi waktu karena jika kunci waktu didalam tabel fakta adalah tipe data tanggal, beberapa SQL *query* mendapat kendala langsung dari kunci waktu pada tabel fakta dan menggunakan *natural* SQL untuk kendala pada bulan dan tahun.

3. Pembahasan

Penjelasan dibawah ini adalah hasil dari penelitian.

Perencanaan proyek.

Pada fase perencanaan proyek dihasilkan subyek dari *data warehouse*, yaitu sirkulasi.

Mendefinisikan kebutuhan.

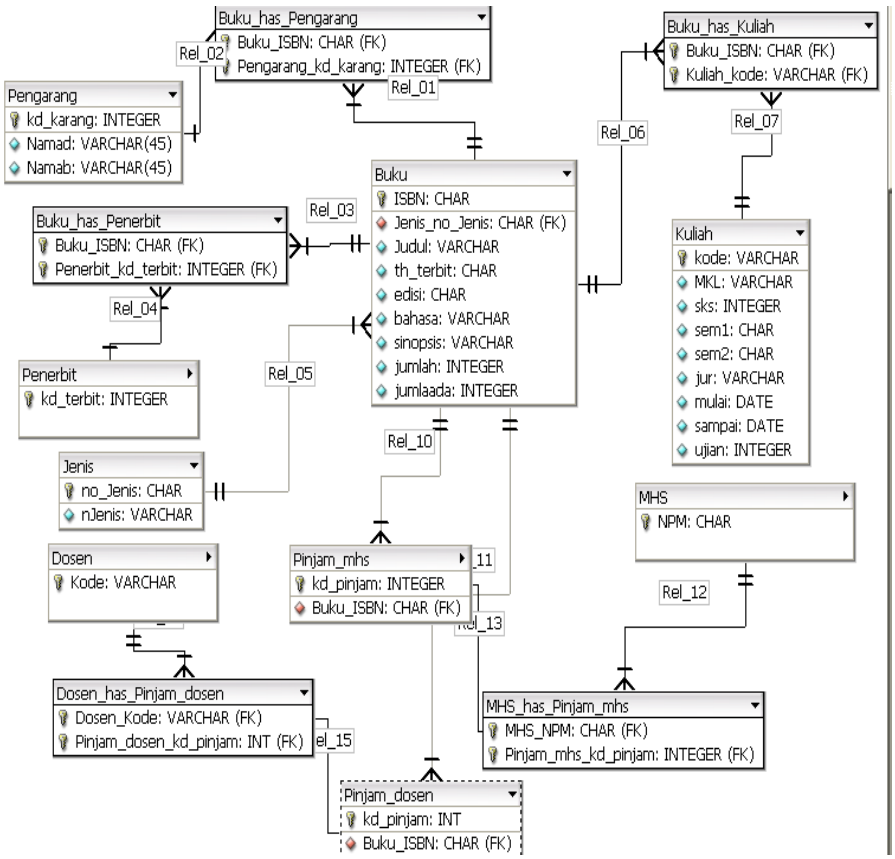
Pada fase ini mendefinisikan semua sumber data yang digunakan dalam *data warehouse* dan mendefiniskan kebutuhan informasi yang harus dihasilkan oleh *data warehouse* melalui analisis *reporting.*

Sumber data

Sumber data yang digunakan dalam *data warehouse* adalah data-data dalam database perpustakaan, yang berisi antara lain:

1. Tabel Buku
2. Tabel Jenis
3. Tabel Pengarang
4. Tabel Karang
5. Tabel IsiCD
6. Tabel Program
7. Tabel Terbit
8. Tabel Penerbit
9. Tabel Kuliah
10. Tabel Buku_mk
11. Tabel Data_buku
12. Tabel Sumber
13. Tabel Mhs
14. Tabel Dosen
15. Tabel Pinjam_Mhs
16. Tabel Pinjam_DSN
17. Tabel Sirkulasi1
18. Tabel Sirkulasi2

Adapun gambar ER-Diagram database perpustakaan yang digunakan bisa dilihat pada gambar 5.



Gambar 5. ER-Diagram Perpustakaan

Kebutuhan Informasi

Berdasarkan hasil pengamatan dan wawancara langsung kepada petugas atau manajemen perpustakaan didapat kebutuhan informasi sebagai berikut.

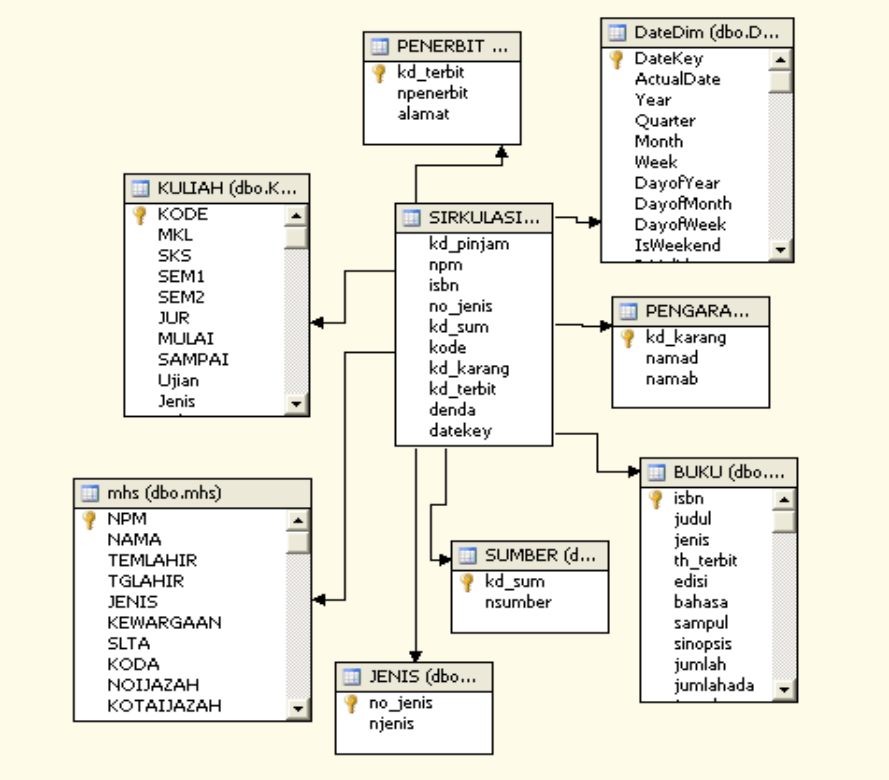
1. Tren terhadap buku-buku yang dipinjam oleh dosen dan mahasiswa.
2. Informasi mengenai jenis buku yang sering dipinjam.
3. Tren proses sirkulasi.
4. Informasi pengenai nama pengarang yang bukunya sering dipinjam.
5. Informasi mengenai nama penerbit yang bukunya sering dipinjam.

Desain

Pada fase ini aktifitas yang dilakukan adalah membuat model data dimensional yang berupa *star Schema*, mendesain proses ETL (*Extraction, Transformation, Loading*), dan menganalisis metadata yang digunakan dalam *data warehouse*.

Model data dimensional

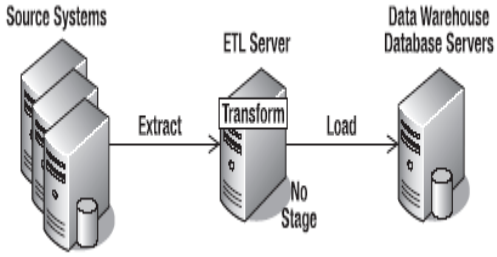
Berdasarkan ER-Diagram Perpustakaan dan kebutuhan informasi yang dibutuhkan oleh petugas dan manajemen perpustakaan maka model data dimensional yang dibuat dalam bentuk *star schema* untuk *data warehouse* perpustakaan bisa dilihat pada gambar 6.



Gambar 6. *Star Schema* untuk *data warehouse* perpustakaan

Desain proses ETL

Proses selanjutnya setelah *star schema* dibuat adalah membuat desain proses ETL. Desain proses ETL yang dipakai bisa dilihat pada gambar 7. dibawah ini. Proses ETL mengambil data dari *source systems* menggunakan *query.* ETL berkoneksi dengan source system database dan mengambil data dengan query. Setelah data hasil *query* diambil langkah selanjutnya dilakukan eksekusi proses ETL dan mengirimnya ke database *data warehouse*.



Gambar 7. Desain proses ETL

4. Kesimpulan

Adapun kesimpulan yang bisa diambil dari penelitian ini adalah sebagai berikut:

1. Diperlukan pengkodean khusus untuk id dosen dan id mahasiswa pada database data warehouse.
2. *Field* yang diperlukan pada tabel waktu sudah disesuaikan dengan kebutuhan manajemen perpustakaan.

3. Saran

Adapun saran yang diberikan untuk penelitian selanjutnya adalah sebagai berikut :

1. Pada saat fase perancangan proyek, subyek *data warehouse* bisa dikembangkan menjadi beberapa subyek sesuai dengan kebutuhan manajemen perpustakaan.
2. Diperlukan pengembangan bahwa *data warehouse* perpustakaan sebagai *data mart* dari suatu *data warehouse* di STMIK AMIKOM Yogyakarta.

DAFTAR PUSTAKA

Chaudhuri, Surajit, An Overview of Data Warehousing and OLAP Technology [Online] http://www.sigmod.org/record/issues/9703/chaudhuri.ps. tanggal akses 21 Mei 2008, pukul 11.23 WIB.

Ferdiana, Ridi ,2008, Strategi pengelolaan asset data menggunakan konsep data warehouse dalam cara pandang organisasi. [Online] http://micresearch.net/file.axd?file=Data+warehouse+dalam+definisi+organisasi+dan+bisnis.pdf tanggal akses 21 Mei 2008, pukul 10.30 WIB

Han, Jiawei; Kamber, Micheline. 2006. *Data Mining: Concepts and Techniques .* San Fransisco: Morgan Kaufmann.

Hutabarat, Bernaridho I., Data Warehouse dengan SQL Server 2005. Elex Media Komputindo: Yogyakarta.

Kalyani, 2008, Data warehouse. [Online] http://www.cse.buffalo.edu/DBGROUP/nachi/ecopres/kalyani.ppt#258,3,OLTP vs. Warehousing. Tanggal akses 4 September 2008, pukul 09.40 WIB.

Kimball, Ralph; Caserta, Joe. 2004. *The Data Warehouse ETL Toolkit*. New Delhi: Wiley Publishing:.

Mallach, Efrem G.,. 2000. *Decision Support and Data Warehouse Systems*. Singapore: Irwin McGraw Hill.

Nicholson, Scott. 2003. *The Bibliomining Process: Data Warehouse dan Data Mining for Library Decision-Making*. [Online] http://bibliomining.org. Tanggal akses 17 April 2008, pukul 09.00 WIB.

Ponniah, Paulraj. 2001. Data Warehouse Fundamentals: a Comprehensive Guide for IT Professional. New York : John Wiley & Sons.

Populate Time Dimension of AdventureWorksDW Sample Database and use it in yourDatawarehouse/cube.[Online] http://blogs.msdn.com/azazr/archive/2008/05/09/populate-time-dimension-of-adventureworksdw-sample-database-and-use-it-in-your-datawarehouse-cube.aspx. Tanggal akses 22 Juli 2008, Pukul 10.45.

Rainardi, Vincent, 2008, Building a Data Warehouse with Examples in SQL Server. Apress: New York.

Tang, ZhaoHui; MacLennan, Jamie. 2005. *Data Mining with SQL Server 2005*. Indiana Polis : Wiley Publishing.
Tanggal akses 21 April 2008, pukul 10.15 WIB.

Turban, 2005, Decision Support Systems and Intelligent Systems (Sistem Pendukung Keputusan dan Sistem Cerdas) Jilid 1. Andi Offset: Yogyakarta.

Winpec Solution, 2007, Membuat Aplikasi Reporting Service dengan SQL Server 2005. Elex Media Komputindo: Jakarta.

Yuadi, Imam, 2007, Perpustakaan Digital: Paradigma, Konsep, dan Teknologi Informasi yang digunakan. [Online] http://journal.unair.ac.id/filerPDF/PERPUSTAKAAN%20DIGITAL.pdf