

**LAPORAN AKHIR
HIBAH PENELITIAN DOSEN PEMULA (PDP)**



**DESAIN ARSITEKTUR DATABASE SECARA REAL-TIME
ANTAR DATABASE HETEROGEN UNTUK MENGELOLA
INTEGRASI DATA ANTAR DATABASE EPIDEMIOLOGI**

Tahuk ke 1 dari rencana 1 tahun

Oleh :

- 1. Muslih, M.Kom, NIDN-0604057501 (Ketua)**
- 2. Elkaf Rahmawan Pramudya, M.Kom, NIDN-0612067502 (anggota)**

**LEMBAGA PENELITIAN DAN PENGABDIAN PADA MASYARAKAT
UNIVERSITAS DIAN NUSWANTORO SEMARANG
DESEMBER 2014**

HALAMAN PENGESAHAN

Judul Kegiatan : DESAIN ARSITEKTUR DATABASE SECARA REAL-TIME
ANTAR DATABASE HETEROGEN UNTUK MENGELOLA
INTEGRASI DATA ANTAR DATABASE EPIDEMIOLOGI

Peneliti / Pelaksana

Nama Lengkap : MUSLIH M.Kom
NIDN : 0604057501
Jabatan Fungsional :
Program Studi : Desain Komunikasi Visual
Nomor HP : 08122565708
Surel (e-mail) : muslih@dsn.dtnus.ac.id

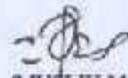
Anggota Peneliti (1)

Nama Lengkap : ELKAF RAHMAWAN PRAMUDYA M.Kom
NIDN : 0612067502
Perguruan Tinggi : Universitas Dian Nuswantoro
Institusi Mitra (jika ada) :
Nama Institusi Mitra :
Alamat :
Pesanggung Jawab :
Tahun Pelaksanaan : Tahun ke 1 dari rencana 1 tahun
Biaya Tahun Berjalan : Rp. 12.500.000,00
Biaya Keseluruhan : Rp. 15.000.000,00

Mengetahui
Dekan Fakultas Ilmu Komputer


(Dr. Abdul Syukur)
NIP/NIK 0686.11.1992.017

Semarang, 3 - 11 - 2014,
Ketua Peneliti,


(MUSLIH M.Kom)
NIP/NIK0686.11.1996.082

Menyetujui,
Kepala LP2M Udinus


(Dr. Y. Trias Catur Pramudi, S.Si.M.Kom)
NIP/NIK 0686.11.1984.046



RINGKASAN

Pengaruh globalisasi terhadap perusahaan yang semakin luas area bisnisnya, maka organisasi menuntut inovasi teknologi informasi yang dapat mengelola peningkatan jumlah data dan skalabilitas jarak transaksi (antara sistem aplikasi dengan database maupun antar database itu sendiri). Perusahaan yang memiliki cabang diberbagai lokasi yang semakin jauh dan meluas, maka sistem aplikasi komputer memerlukan pilihan arsitektur basis data yang optimal dalam mengimbangi perkembangan bisnis tersebut termasuk yang berkaitan dengan model distribusi dan integrasi datanya. Seperti pada pengelolaan data epidemiologi kesehatan, dimana sumber data tersebar pada database yang ada pada berbagai lokasi rumah sakit dan poliklinik pada suatu wilayah kabupaten atau suatu kota tertentu. Permasalahannya database sumber (source) bersifat heterogen sehingga mengalami potensi konflik (kesulitan) dalam melakukan integrasi menuju pada pusat data epidemiologi (target) pada dinas kesehatan. Potensi konflik yang terjadi adalah ketidakseragaman skema relasi (konflik skema), ketidakakuratan isi (konflik data). Untuk itu dalam integrasi memerlukan analisis database sumber yang bersifat heterogen dengan melakukan strukturisasi dan sinkronisasi sebagai persiapan integrasi data. Dengan permasalahan integrasi antar database distribusi tersebut maka dalam penelitian ini bertujuan mendesain arsitektur database terdistribusi dengan metode replikasi yang akan diimplementasikan pada integrasi database epidemiologi sehingga akan didapatkan sebuah arsitektur database terdistribusi yang bisa mengatasi ketersediaan data pada sistem surveilans terpadu (SST). Untuk mengatasi masalah tersebut maka diperlukan suatu pengembangan arsitektur basis data terdistribusi untuk pola sinkronisasi dan integrasi dengan metode replikasi mapping schema. Metode replikasi mapping schema generator merupakan kemajuan dari rekayasa konsep teknologi DDBMS (Distributed Database Management System) yang mampu melakukan sinkronisasi (captures, routes, transforms) dan integrasi data yang bersifat heterogen secara real-time. Proses utama dalam metode replikasi mapping schema mencakup proses mendeteksi konflik schema, representasi schema matching, replikasi Integration common dan schema

mapping. DDBMS mengelola data yang tersebar di site-site dalam sebuah jaringan atau node-node dari sebuah sistem multiprocessor. Prinsip basis data terdistribusi adalah suatu basis data dengan skema global (global schema) yang berada dibawah kendali sistem manajemen basis data (DBMS) terpusat dengan piranti penyimpanan data yang terpisah-pisah dalam skema lokal (local schema) dalam suatu jaringan komputer. Tujuan jangka panjang dalam penelitian ini adalah merancang teknologi dan aplikasi dalam mengembangkan teknik integrasi data dari berbagai ragam aplikasi dan database tanpa harus menyeragamkan aplikasi dan database yang sudah ada. Dengan demikian schema local dapat dipertahankan dalam mendapatkan schema global melalui teori rekayasa sinkronisasi dan integrasi basis data. Sedangkan target khusus yang akan dicapai adalah memperoleh model arsitektur database tersebar untuk integrasi data yang dapat diterapkan dalam mengelola dan mengembangkan sistem informasi epidemiologi terintegrasi pada dinas kesehatan dengan metode replikasi mapping schema . Metode penelitian yang digunakan adalah dengan studi literatur dan studi lapangan. Setelah melakukan studi awal kegiatan penelitian dilanjutkan dengan observasi dan studi pustaka, analisa permasalahan dalam pernacangan arsitektur database. Tahapan berikutnya adalah melakukan desain pola integrasi database dan dilakukan uji integrasi dan replikasi untuk mendapatkan kesimpulan integrasi antar database heterogen. Hasil dalam kasus penelitian ini adalah integrasi antar dua relasi yang terjadi konflik (surveila_rs_A dan data_center_SST) menggunakan mapping schema (relasi_map_ICD_X).

Kata kunci: DDBMS, schema global. Schema local, sinkronisasi

PRAKATA

Alhamdulillah, puji syukur peneliti panjatkan kehadirat Allah SWT yang telah memberikan rahmat dan hidayahNya sehingga peneliti dapat menyusun dan menyelesaikan laporan akhir penelitian ini. Penelitian yang berjudul DESAIN ARSITEKTUR DATABASE SECARA REAL TIME ANTAR DATABASE HETEROGEN UNTUK MENGELOLA INTEGRASI DATA ANTAR DATABASE EPIDEMIOLOGI, merupakan penelitian dosen pemula yang dibiayai DIKTI selama satu tahun.

Selama melakukan penelitian dan selesainya laporan kemajuan penelitian ini tidak terlepas dari bantuan dan dorongan dari berbagai pihak, baik secara moril dan materiiil. Oleh karena itu peneliti mengucapkan terimakasih kepada :

1. Direktur Jendral Pendidikan Tinggi (Dirjen Dikti) Kemendikbud
2. Direktorat Penelitian dan Pengabdian kepada Masyarakat (Ditlitabmas)
3. Kopertis wilayah VI Jawa Tengah
4. Rektor UDINUS Semarang
5. Ketua Lembaga Penelitian UDINUS Semarang
6. Dosen dan mahasiswa Udinus Semarang
7. Istri, anak dan kerabat atau keluarga besar

Meskipun sudah memperhatikan berbagai aspek yang berhubungan dengan dengan penulisan laporan akhir dalam penelitian ini, peneliti menyadari masih banyak terdapat kekurangan dan kelemahan dalam penelitian ini. Saran dan kritik yang bersifat membangun merupakan masukan yang peneliti harapkan. Semoga penelitian ini bermanfaat dan dapat dikembangkan oleh para peneliti lainnya.

Semarang, 24 Oktober 2014

Muslih, M.Kom
NIDN. 0604057501

DAFTAR ISI

Halaman Pengesahan	i
Ringkasan	ii
Prakata	v
Daftar Isi	vi
Daftar Tabel	vii
Daftar Gambar	viii
Daftar Lampiran	ix
Bab 1 Pendahuluan	1
Bab 2 Tinjauan Pustaka	
2.1. Sinkronisasi.....	3
2.1.1. Pengertian Sinkronisasi.....	3
2.1.2. Protokol Sinkronisasi.....	3
2.2. Replikasi.....	3
2.2.1. Pengertian Replikasi.....	3
2.2.2. Jenis Replikasi.....	3
2.2.3. Teknik Replikasi.....	4
2.3. Basis data Distribusi.....	4
2.3.1. Definisi Basis Data Distribusi.....	4
2.3.2. Desain Basis data Distribusi.....	5
2.3.3. Fragmentasi Sistem Multi Database Distribusi.....	7
2.3.4. Masalah Integrasi Pada Multi Database.....	8
Bab 3 Tujuan Dan Manfaat Penelitian.....	11
Bab 4 Metodologi Penelitian	
Bab 5 Hasil Dan Pembahasana	
5.1. Mendeteksi Konflik Antar Skema Unit Surveilans.....	16
5.2. Representasi Konflik.....	18
5.3. Mapping Relasi... ..	18
Bab 6 Kesimpulan dan Saran.....	26
Daftar Pustaka	
Lanpiran	

DAFTAR TABEL

Tabel 1. Tahapan Penelitian yang sudah dilakukan dan akan dilanjutkan.....	29
--	----

DAFTAR GAMBAR

Gambar 1. Framework Sistem Tersebar	5
Gambar 2. Perancangan Secara Top Down	6
Gambar 3. Perancangan Secara Button Up.....	6
Gambar4.Bagan Proses Restrukturisasi dan Integrasi Basis Data.....	10
Gambar 5. Tahapan Penelitian.....	13
Gambar 6. Arsitektur Integrasi Epidemiologi	16
Gambar 7. Deteksi Konflik Skema level instance	17
Gambar 8. Relasi Data Surveilans sumber dan target	18
Gambar 9. Mapping Relation map_ICD_X.....	19
Gambar 10. Nama dan struktur tabel yang berbeda untuk transformasi konflik integrasi	20
Gambar 11. Skema Multidatabase.....	22
Gambar 12. Rancangan Pola DML Multidatabase.....	23
Gambar 13. Transpormasi Relasi.....	25

DAFTAR LAMPIRAN

Lampiran 1. Personalia tenaga peneliti beserta kualifikasinya

Lampiran 2. Publikasi

BAB 1 PENDAHULUAN

Pembangunan dibidang kesehatan merupakan bagian integral dan terpenting dalam pembangunan nasional. Tujuan diselenggarakan pembangunan kesehatan adalah untuk meningkatkan kesadaran, kemauan dan kemampuan hidup sehat bagi setiap orang agar terwujudnya derajat kesehatan masyarakat yang optimal. Hal ini sesuai dengan amanat Undang-Undang Dasar 1945 pasal 28 H ayat (1) bahwa setiap orang berhak hidup sejahtera lahir dan batin, bertempat tinggal dan mendapatkan lingkungan hidup baik dan sehat serta berhak mendapatkan pelayanan kesehatan. Seiring dengan sistem otonomi daerah maka penjabaran tujuan pembangunan nasional yang sesuai dengan amanat UUD 1945 tersebut menjadi tugas pokok pemerintah daerah, dimana parameter keberhasilan pembangunan daerah dapat dilihat dari pencapaian Indeks Pembangunan Manusia (*IPM*). Komponen penting yang mempengaruhi *IPM* adalah indikator kesehatan per kapita. Dengan indikator pencapaian derajat kesehatan perkapita menjadi upaya peningkatan kualitas sumberdaya manusia, sehingga secara tidak langsung akan mendukung percepatan pembangunan nasional.

Indikator kesehatan masyarakat sangat erat kaitanya dengan epidemiologi suatu kasus pada suatu daerah tertentu. Epidemiologi adalah wabah penyakit terutama yang menular secara cepat dan tak terduga pada suatu wilayah tertentu. Agar wabah tidak meluas eskalasinya maka diperlukan sistem monitoring untuk mengembangkan suatu metode dalam menganalisis secara sistematis keadaan dan keberadaan suatu penyakit dalam upaya untuk mengatasi dan menaggulangi secara cepat dan terintegrasi. Untuk itu Departemen Kesehatan telah mengeluarkan keputusan menteri **No. 1479/MENKES/SK/X/2003** tentang : Pedoman Penyelenggaraan Sistem Surveilans Epidemiologi Penyakit Menular Dan Penyakit Tidak Menular Terpadu. Dalam pedoman surveilans tersebut menegaskan diperlukannya suatu Sistem Surveilans Terpadu (*SST*) dengan dukungan basis data yang setandar dimana sistem pengawasan utama epidemiologi meliputi semua unit pelayanan kesehatan (Puskesmas, Laboratorium, Rumah Sakit) di semua Pemerintah Daerah Propinsi dan

Pemerintah Daerah Kabupaten/Kota dengan model : Sistem Pencatatan Pelaporan Puskesmas Terpadu (*SP2PT*) dan Sistem Pelaporan Rumah Sakit (*SPRS*). Di tingkat pemerintah daerah pelaksanaan operasional *SST* tersebut sepenuhnya diserahkan kepada dinas kesehatan daerah untuk bisa menjadi sistem informasi epidemiologi dalam rangka mendukung pemberantasan penyakit menular dan tidak menular secara nasional. Direktorat Jendral Pemberantasan Penyakit Menular dan Penyehatan Lingkungan (Dirjen PPM & PL Departemen Kesehatan) sebagai lembaga pemerintah pusat yang mendapat tugas dan bertanggung jawab dalam bidang pengendalian maupun pemberantasan penyakit secara nasional.

Namun dalam pelaksanaan dan penyelenggaraan sistem surveilans terpadu (*SST*) tersebut ditingkat kabupaten/kota menghadapi suatu kendala dalam melakukan pengiriman data kesehatan ke dinas kesehatan kabupaten/kota sebagai penanggung jawab kesehatan di tingkat pemerintah daerah. Kendala ini disebabkan karena sumber data unit surveilans (puskesmas, laboratorium, rumahsakit) terletak secara geografis tersebar dan diperoleh dari beragam aplikasi dan database manajemen sistem (*DBMS*) yang beragam (heterogen). Sinkronisasi dan integrasi data antara unit surveilans (puskesmas, laboratorium, rumahsakit) dengan dinas kesehatan kabupaten/kota menjadi masalah utamanya. Hal ini disebabkan karakter masing-masing unit surveilans memiliki ketidak seragaman platform aplikasi dan database (*heterogen*). Untuk itu dibutuhkan suatu pola kaidah sinkronisasi dan integrasi data antar unit surveilans (skema local) dengan dinas kesehatan sebagai “*Data Center*” (skema global) dalam model aljabar relasi dan notasi skema relasi (*relation schema*) serta skema basis data (*database schema*). Sehingga bisa menjadi modal ilmiah dalam membangun basis data terdistribusi untuk mendukung sistem informasi terpadu epidemiologi sebagai pelaksanaan sistem surveilans terpadu (*SST*) tingkat kabupaten/kota.

BAB 2 TINJAUAN PUSTAKA

2.1. Sinkronisasi

2.1.1. Pengertian Sinkronisasi

Sinkronisasi adalah proses penyesuaian data terhadap skala waktu dari proses osilasi yang terjadi antara proses osilasi tersebut (*Balanov, 2010*). Pada dasarnya sinkronisasi terdiri dari dua jenis yaitu *one-way file synchronizatiion* (sinkronisasi satu arah) dimana file-file yang telah mengalami perubahan pada bagian pusat (*source*) akan dibuat salinannya dan dipindah ke lokasi targetnya. Pada *one-way file synchronizatiion* ini, tidak ada file dari target yang akan menuju ke bagian *source*. Sedangkan pada jenis yang kedua, *two-way file synchronizatiion* (sinkronisasi 2 arah) proses pembuatan salinan dan pemindahannya dapat berjalan 2 arah baik dari source ke target maupun sebaliknya.

2.1.2. Protokol Sinkronisasi

Dalam teknik sinkronisasi dibutuhkan beberapa protokol yang digunakan mendukung komunikasi dan replikasi. Beberapa teknik sinkronisasi adalah HotSyn, Intellisync, SyncML, CPISync.

2.2. Replikasi

2.2.1. Pengertian Replikasi

Replikasi adalah suatu teknik untuk melakukan copy dan pendistribusian data dan objek-objek database dari satu database ke database lain dan melaksanakan sinkronisasi antara database sehingga konsistensi data dapat terjamin (*Dollimore, 2012*). Dengan menggunakan teknik replikasi ini, data dapat didistribusikan ke lokasi yang berbeda melalui koneksi jaringan lokal maupun internet.

2.2.2. Jenis Replikasi

Terdapat beberapa jenis replikasi diantaranya adalah :

1. Snapshot Replication : Mendistribusikan data yang dapat dilihat pada saat tertentu tanpa melakukan update.
2. Transactional Replication : Jenis replikasi ini lebih mementingkan dan memelihara kekonsistenan transaksi yang terjadi.
3. Merge Replication : Memungkinkan pengguna bekerja dan merubah data sesuai dengan wewenangnya. Pada saat server tidak dikoneksikan keseluruh lokasi dalam topologi, replikasi merubah data ke nilai yang sama.

2.2.3. Teknik Replikasi

Cara replikasi dalam DBMS terdistribusi dapat dikelompok dalam 2 teknik replikasi www.learning.unl.ac.uk/csp003n/lectures/w021-ddb-arch.pdf :

1. Teknik Single Master Replicated : Dengan metode ini, salah satu komputer berfungsi sebagai master dan yang lainnya berfungsi sebagai slave. Pada prosesnya, komputer digunakan sebagai server akan dapat read dan write kedalam database. Sedangkan komputer yang berfungsi sebagai slave, hanya akan read saja kedalam basis data tersebut. Apabila kita melakukan perubahan data pada master, maka otomatis data pada slave akan berubah. Tetapi jika kita melakukan perubahan data pada slave , basi data pada master tidak akan berubah.
2. Teknik Multi Master Replicated Dengan metode ini, salah satu komputer berfungsi sebagai master server dan yang lainnya berfungsi sebagai master server juga. Pada prosesnya, setiap komputer akan dapat write dan read data dalam database. Apabila kita melakukan perubahan data pada master server 1, maka otomatis data pada master server 2 akan berubah. Begitu juga jika kita melakukan perubahan data pada master server 2, maka basis data pada basis data pada master server 1 akan berubah.

2.3. Basis Data Distribusi

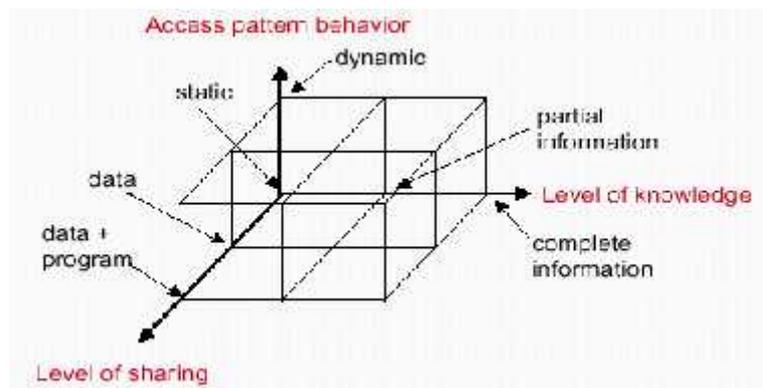
2.3.1. Definisi Basis Data Distribusi

Basis data terdistribusi adalah basis data yang disebarakan pada sejumlah lokasi (A.Silberschatz, 2005). Setiap lokasi tersebut memiliki kewenangan sendiri dalam

mengelola basis data. Masing-masing lokasi bisa melakukan transaksi lokal dan transaksi global. Basis data terdistribusi memiliki beberapa keuntungan seperti : meningkatkan performance, reliabilitas, ketersediaan data, memudahkan perluasan, dan meningkatkan Otonomi.

2.3.2. Desain Sistem Basis Data Distribusi

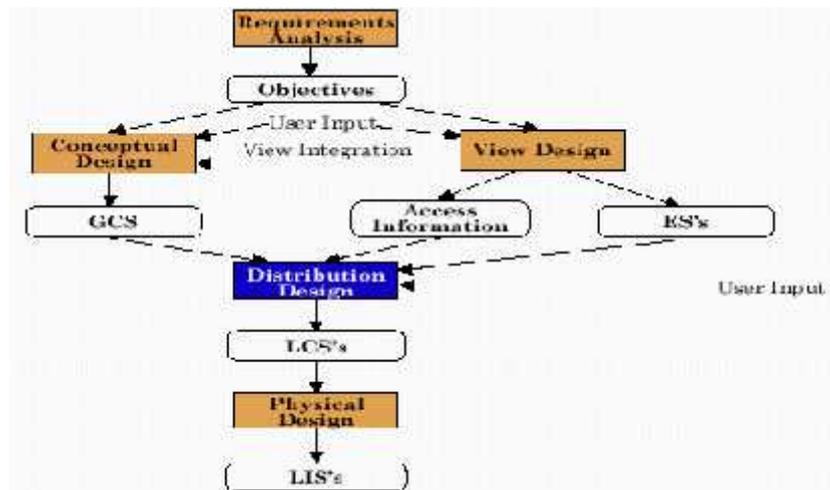
Desain sistem basis data terdistribusi meliputi bagaimana letak data dan program dalam suatu jaringan komputer. Dalam kasus sistem database tersebar, distribusi aplikasi meliputi 2 hal, yaitu distribusi software DBMS dan distribusi program aplikasi program aplikasi. Organisasi dari sitem tersebar dapat meliputi (Date C.J,2005) : Level Sharing, Pola Akses, Level Pengetahuan Pola Akses.



Gambar 1. : Framework Sistem Tersebar

Dalam strategi perancangan system database terdistribusi terdapat 2 pendekatan yaitu :

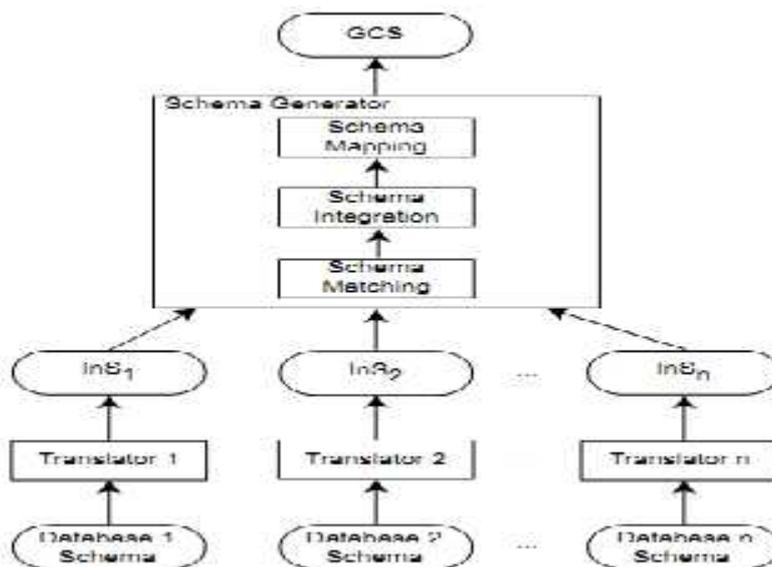
- a. Proses Perancangan Top Down dengan ciri – ciri pendekatan : Merancang system database dari awal, DBMS bersifat seragam (homogen)



Gambar 2. : Perancangan Secara Top Down

b. Perancangan Button Up

Titik awal dari dari perancangan secara button up adalah individual local conceptual schema. Proses perancangan berupa proses integrasi local schema menjadi global conceptual schema. Ciri-ciri pendekatan ini adalah : database sudah ada dalam suatu situs, mengintegrasikan beberapa database heterogen menjadi satu database utama.



Gambar 3. : Perancangan Secara Button Up

2.3.4. Fragmentasi Sistem Multi Basis Data Distribusi

Fragmentasi data merupakan proses dimana basis data akan dipecah-pecah kedalam unit-unit logic yang disebut fragment yang kemudian akan disimpan dalam site yang berbeda (A. Tannenbaum, 2008). Suatu sistem basis data terdistribusi adakalanya dibentuk dari beberapa basis data yang berlainan. Sistem seperti ini disebut *multidatabase*, yang pembentukannya dilakukan dengan integrasi basis data. Dalam melakukan integrasi ini, boleh jadi ada ketidakseragaman antara basis data-basis data yang membentuknya dan mengakibatkan konflik, baik *konflik skema* (akibat ketidakseragaman skema relasi) maupun *konflik data* (akibat ketidakakuratan isi, misalnya presisi, besaran dan satuan, juga data yang tidak tepat atau sudah tidak berlaku [*obsolete*]). Oleh karena itu, perlu diperhatikan metode penyelesaian konflik sebelum terjadi integrasi, hal ini disebut restrukturisasi. Baik restrukturisasi table maupun atribut, proses ini jika diakaitkan dengan schema mapping dan shema matching disebut dengan *schema generator*.

1. Fragmentasi Horizontal : Fragmentasi berdasarkan tupel. Setiap fragment memiliki subset dari tupel relasi, Relasi r dibagi kedalam sejumlah subset $r_1, r_2, r_3, \dots, r_n$, masing-masing berisi dari sejumlah tupel relasi r. Masing-masing tupel relasi r harus merupakan satu dari fragment-fragment tersebut sehingga relasi awalnya dapat dibentuk kembali. Suatu fragment didefinisikan sebagai seleksi pada relasi global r. Sebuah predikat P_i digunakan untuk menyusun fragmen r_i :

$$R_i = P_i(r)$$

Pembentukan kembali dilakukan dengan mengembalikan seluruh fragment :

$$R = \bigcup_{i=1}^n r_i$$

2. Fragmentasi Vertikal : Fragmentasi vertikal dari r (R) melibatkan beberapa subset R_1, R_2, \dots, R_n dari sedemikian sehingga :

$$\bigcup_{i=1}^n R_i = R$$

$$i=1$$

Setiap fragment r_i dari r didefinisikan sebagai :

$$r_i = \pi_{R_i}(r)$$

Pembentukan kembali dengan menggunakan natural join $r = r_1 \bowtie r_2 \bowtie \dots \bowtie r_n$. Fragmentasi vertikal dibuat dengan menambahkan atribut khusus yaitu tuple-id yang merupakan alamat fisik atau logika untuk tupel dan menjadi kunci untuk skema.

- a. Derajat Fragmentasi: Performansi eksekusi suatu query sangat tergantung dari database yang mana dan seberapa besar dari database tersebut didekomposisi dan dialokasikan ke beberapa site. Derajat fragmentasi bisa nol (tidak terfragmentasi sama sekali), fragmentasi horisontal dan fragmentasi vertikal.
- b. Aturan dalam Fragmentasi (sama dengan prinsip normalisasi) :
 - ❖ Komplit : Dekomposisi relation/table R menjadi beberapa fragmen R_1, R_2, \dots, R_n dikatakan komplit kalau setiap item data pada R dapat juga ditemukan di beberapa R_i .
 - ❖ Rekonstruksi : Kalau ada suatu relation/table R didekomposisi menjadi beberapa fragment R_1, R_2, \dots, R_n , maka harus ada operator yang dapat mengembalikan fragmen-fragmen tersebut ke R . ($R = \bigvee R_i, \forall R_i \in F_R$)
 - ❖ Disjointness : Kalau ada suatu relation/table R didekomposisi menjadi beberapa fragment R_1, R_2, \dots, R_n , dan item data d_i ada di R_j , maka d_i harus tidak boleh ada di fragment $R_k (k \neq j)$
- c. Teknik Alokasi fragmen ke beberapa site (Reliability & Efisiensi Query):
 - ❖ Non-replicated (Partitioned Database): Setiap fragmen hanya diletakkan di satu site.
 - ❖ Replicated: Fully Replicated: setiap fragmen ada di setiap site dan Partially Replicated: setiap fragmen ada di beberapa site.

2.3.5. Masalah Integrasi Pada Multidatabase

Integrasi multidatabase dirancang untuk mendapatkan informasi yang terpadu, umumnya bertujuan untuk menggabungkan sistem yang dipilih sehingga akan membentuk satu kesatuan dalam sistem informasi dalam berinteraksi. Pengguna akan

disajikan sebuah logical view data homogen, walaupun secara fisik didistribusikan atau dialokasikan dari sumber data yang heterogen. Untuk itu semua data harus dipresentasikan dari prinsip-prinsip abstraksi yang sama (satu model data global dan satu semantik). Sehingga dihadapkan pada tugas mendeteksi dan resolusi skema data yang berkaitan dengan konflik struktur serta semantiknya. Konflik struktur dan semantik dalam integrasi data base disebabkan adanya beberapa heterogenitas yang berkaitan dengan hardware, sistem operasi, DBMS, data model, schema data, semantik data, middleware, user interface dan kendala aturan bisnis. Beberapa konflik yang dapat terjadi pada integrasi multidatabase yang berkaitan dengan skema relasi maupun keakuratan data yang tidak seragam adalah :

i) Konflik Antar Tabel

(a) Konflik Antar Dua Tabel

(i) Nama Tabel (homonim/sinonim), dapat diselesaikan dengan view.

(ii) Struktur Tabel, seperti jumlah atribut berbeda di dua tabel yang informasinya sama, dapat diselesaikan dengan membuang atribut yang keberadaannya tidak disemua relasi, atau menambah atribut yang kurang pada relasi yang kekurangan atribut.

(iii) Integrity Constrain, misalnya pada dua situs yang terdapat pada tabel yang sama, tetapi isi atribut primary key-nya berbeda, dapat diselesaikan dengan menambahkan primary key tambahan yang berisi informasi situs relasi tersebut disimpan.

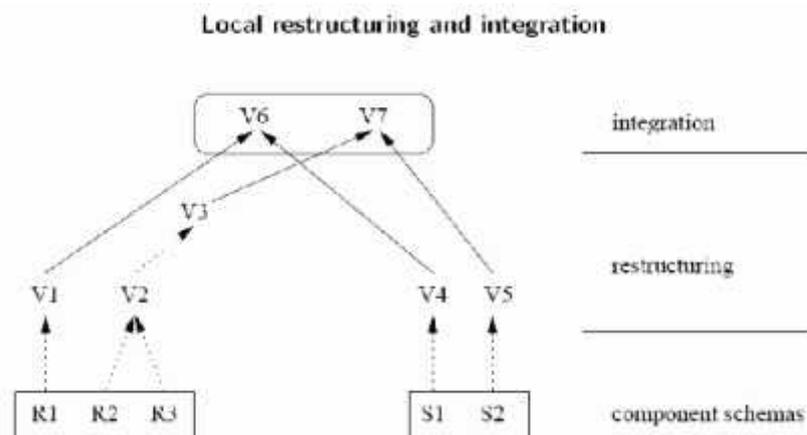
(b) Antar Banyak Tabel, misalnya pada dua komponen basis data jumlah relasinya tidak sama tetapi informasinya sama, dapat diselesaikan dengan penggabungan relasi dan view .

(2) Konflik Antar Atribut

(a) Antar Dua Atribut

- (i) Nama Atribut (Homonomim/Sinonim), dapat diselesaikan dengan penggantian nama (rename) atribut di view.
 - (ii) Integrity Constraint, misal tipe data dapat diselesaikan dengan fungsi-fungsi konversi, seperti `to_char(int)`, atau `to_int(char)` pada view.
- (b) Antar Banyak Atribut, misalnya dalam penyampaian informasi nama orang dalam tabel yang satu digunakan dua atribut (kolom), nama depan dan nama belakang, sementara pada tabel yang lain digunakan satu atribut nama lengkap. Konflik ini dapat diselesaikan dengan penggabungan string atau pemisahan string dengan fungsi substring.
- (3) Konflik Atribut – Tabel, dapat merupakan kombinasi dari permasalahan diatas.

Penyelesaian konflik tersebut diatas yang berkaitan dengan integrasi basis data harus memerlukan analisis yang mendalam akan komponen basis data dan tidak bisa di otomatisasi. Sebelum melakukan integrasi, komponen basis data harus dipersiapkan terlebih dahulu untuk menagani konflik. Proses penyelesaik konflik ini disebut restrukturisasi basis data.



Gambar 4. Bagan Proses Restrukturisasi dan Integrasi Basis Data.

BAB 3

TUJUAN DAN MANFAAT PENELITIAN

Prinsip sinkronisasi dan integrasi data dalam penelitian ini tidak memerlukan keseragaman basis data dan aplikasi antar unit surveilans dalam mengelola data epidemiologi antar unit surveilans. Akan tetapi yang menjadi perhatian adalah bagaimana data kesehatan hasil dari transaksi masing - masing unit surveilans yang beragam bisa di sinkronkan dan selanjutnya diintegrasikan menjadi pusat data epidemiologi pada database epidemiologi dinas kesehatan untuk kepentingan sistem informasi surveilans epidemiologi. Untuk itu tujuan dalam penelitian ini adalah :

- a. Mengidentifikasi masalah dan peluang pengembangan model atau pola kaidah dalam rangka untuk melakukan sinkronisasi dan integrasi data hasil replikasi antar DBMS yang terdistribusi dan beragam dalam suatu pola kaidah basis data yang terdistribusi.
- b. Mendapatkan gambaran tentang teori-teori pendukung tentang sinkronisasi, integrasi, distribusi dan replikasi data yang dapat di upayakan untuk mengembangkan pola kaidah sinkronisasi dan integrasi data antar unit surveilans dalam model aljabar dan notasi relasi.
- c. Diperoleh desain arsitektur pola sinkronisasi dan integrasi data yang terdistribusi secara heterogen untuk mendukung pelaksanaan SST di dinas kesehatan.
- d. Diperoleh diskripsi pola integrasi data yang terdistribusi dari sumber skema local pada DBMS tersebar dan beragam dengan satu target skema global pada DBMS dinas kesehatan sebagai '*data center*' sehingga dapat mendukung pelaksanaan **SST**.

Sedangkan beberapa manfaat penelitiannya adalah sebagai berikut :

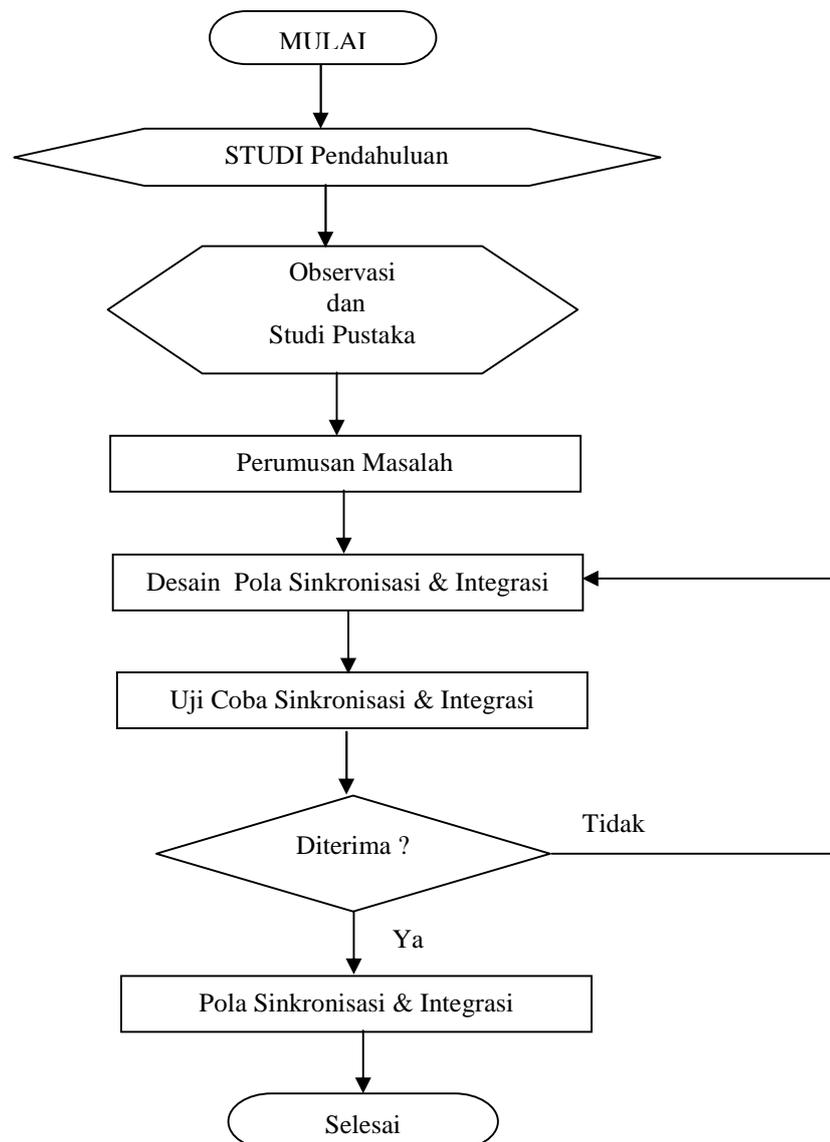
- a. Memberikan diskripsi pola kaidah atau model sinkronisasi dan integrasi data replikasi antar database unit-unit surveilans yang tersebar dan beragam

sehingga kedepan dapat dimanfaatkan oleh dinas kesehatan dalam pelaksanaan penyelenggaraan sistem surveilans terpadu (*SST*) .

- b. Memberikan masukan kepada unit pengelola sistem surveilans terpadu (*SST*) dinas kesehatan tentang pemanfaatan pola kaidah sinkronisasi dan integrasi dalam mengatasi masalah replikasi data antar unit surveilans sehingga bisa mendapatkan data yang sinkron dan sesuai untuk kepentingan sistem informasi epidemiologi.
- c. Meningkatkan kultur dan atmosfir akademik lewat penelitian dalam rangka memanfaatkan ilmu-ilmu atau teori yang bersifat akademik dalam membantu meningkatkan pengembangan teknologi yang bermanfaat bagi lingkungan atau masyarakat.

BAB 4 METODE PENELITIAN

Pada penelitian ini digunakan pendekatan metode studi literatur (library research) dan studi lapangan (field research) untuk mendesain pola sinkronisasi. Adapun tahapan pada metode penelitian dapat dijelaskan dalam diagram dibawah ini :



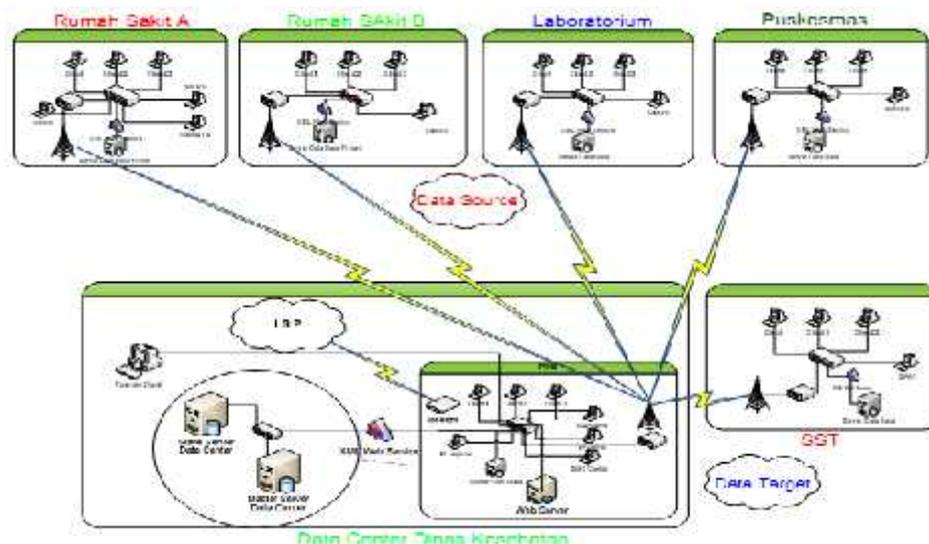
Gambar 5. Tahapan Penelitian

1. Studi Pendahuluan : Pada tahapan ini merupakan kegiatan untuk mengenali lebih lanjut tentang obyek penelitian beserta lingkungan yang terkait dalam rangka mendalami situasi dan kondisi dari sinkronisasi dan integrasi yang akan dikembangkan. Studi pendahuluan dilakukan dengan mengumpulkan informasi mengenai pengelolaan sistem surveilans terpadu (*SST*).
2. Observasi dan Studi Pustaka : Pada tahapan ini akan dilakukan analisis kebutuhan dengan analisa deskriptif dengan cara melakukan kajian pustaka yang terkait dengan konsep sinkronisasi dan integrasi data dan keberadaan data kesehatan yang digunakan untuk surveilans kesehatan pada masing-masing unit surveilans dengan dinas kesehatan.
3. Perumusan Masalah : Tahap selanjutnya setelah mendapatkan permasalahan utama dari obyek penelitian dan dilengkapi dasar teori dari studi pustaka yang mendukung, adalah merumuskan permasalahan yang akan dieksplorasi dalam rangka menemukan pola baru dalam sinkronisasi data kesehatan antar unit surveilans dengan database epidemiologi kesehatan.
4. Desain Pola Sinkronisasi dan Integrasi Data : Pada tahapan ini dilakukan desain pola dan notasi sistem sinkronisasi dan integrasi database sebagai pola integrasi data berdasarkan diagnosis dan identifikasi masalah yang ada, baik dari sisi teori dan teknis maupun implementasi dalam bentuk pola notasi model relasional sinkronisasi dan integrasi sistem distribusi.
5. Pengujian : Pada tahap pengujian ini dilakukan evaluasi terhadap hasil desain dari pola notasi model relasional sinkronisasi dan Integrasi dengan aturan de morgan sehingga nantinya pola sinkronisasi database dapat digunakan secara maksimal dan sesuai (diterima) sebagai suatu pola baru untuk mendukung pengembangan system informasi epidemiologi kesehatan dari lingkungan database distribusi yang heterogen.
6. Hasil Pola Sinkronisasi dan Integrasi (notasi model relasional) : Hasil dari pola notasi model relasional sinkronisasi dan integrasi data ini merupakan sistem sinkronisasi database berbasis proses replikasi data dari masing-masing unit surveilans yang heterogen (puskesmas, laboratorium, rumah sakit) menuju database epidemiologi dinas kesehatan sebagai data center

kesehatan. Sehingga pola ini dapat mempermudah dan mendukung dalam akses data untuk diolah menjadi system informasi epidemiologi kesehatan kabupaten/kota.

BAB 5 HASIL DAN PEMBAHASANYA

Desain arsitektur database untuk replikasi dan integrasi epidemiologi antar unit surveilans (puskesmas, laboratorium, rumah sakit) terletak pada area geografis yang tersebar dan data (source) di peroleh dari sistem aplikasi dan dbms yang beragam (heterogen). Dalam pelaksanaan sistem surveilans terpadu (SST) memerlukan rekayasa untuk keperluan integrasi agar sumber data (source data) dapat di direplikasi atau di distribusikan pada alokasi data senter (target) pada server sistem surveilans terpadu (SST) di dinas kesehatan kota/kabupaten. Rekayasa tersebut untuk menghindari konflik schema dari heteroginitas sumber data surveilans menuju data senter epidemiologi.



Gambar 6. Arsitektur Integrasi Epidemiologi

5.1. Mendeteksi Konflik Antar Skema Unit Surveilans Dengan Skema Data Center

Heterogenitas pada data model integrasi epidemiologi, skema dan level instance akan menyebabkan berbagai macam konflik, konflik tersebut menjadikan permasalahan untuk integrasi dan replikasi data dari sumber data surveilans (puskesmas, poliklinik, rumah sakit) dengan data center pada server sistem surveilans terpadu (SST) di dinas kesehatan sebagai target replikasi. Untuk beberapa konflik level skema terdapat

pada data center dinas kesehatan RL_SST), tetapi digunakan untuk merepresentasikan obyek yang sama. Model konflik ini juga terjadi dari sumber data dari unit surveilan yang lain (puskesmas, laboratorium, poliklinik, dan lainnya).

Gol_Penyakit	Bulan	ICD-X	Nama_Diagnosa	Gol_Umur	Kasus_Lama	Kasus_baru	Jumlah
Infeksi	Januari	A88.1	Infeksi Epidemik V	Muda	6	8	14
Respirasi	Januari	J40	Bronchitis N A	Muda	8	11	19
Simtoma	Januari	R56.0	Vebril Convulsion	Tua	12	17	29
Mata	Januari	H10.9	Conjunctivis Unspe	Tua	8	20	28
Degestif	Januari	K02.9	Dental Caries	Anak	3	13	16

(a) Relasi dari server surveilans_RS_A RL

Kode_Kasus	Bulan	Nama_Diagnosa	Gol_Umur	Kasus_Lama	Kasus_baru	Jumlah
A00-B99: A88.1	Januari	Infeksi Epidemik vertigo	Tua	4	5	9
J00-J99 : J40	Januari	Bronchitis N A	Anak	10	4	14
R00-R99: R56.0	Januari	Vebrile Convulsions	Muda	7	11	18
K00-K99:K02.9	Januari	Dental Caries	Muda	9	12	21

(b) Relasi dari server Data_Center_SST RL

Gambar 8. Relasi Data Surveilans sumber dan target

Dalam rangka untuk mendapatkan petunjuk jenis konflik (deteksi) pada langkah integrasi dari relasi server surveilans RS A dan relasi dari data center SST tersebut, maka kita harus mempertimbangkan proses integrasi. Langkah awalnya adalah mendiskripsikan konflik pada level skema diselesaikan dengan mendefinisikan pemetaan atribut untuk import relasi.

5.2. Representasi Konflik

Sebagai pembahasan untuk representasi resolusi konflik, maka langkah selanjutnya sekenarionya adalah bahwa database pelaporan epidemiologi (sumber, standar laporan epidemiologi RL.2a) pada relasi surveilans RS A harus di integrasikan pada server data center STT. Hubungan yang terstruktur yang ditunjukkan gambar 8. diatas, jelas kita dapat memperkenalkan jenis kasus pasien dari ke dua relasi tersebut.

5.2. Representasi Konflik

Tetapi karena data center SST menggunakan schema sendiri untuk ICD-X , sehingga tidak mungkin dilakukan transformasi dan integrasi langsung. Untuk itu diperlukan pemetaan kode_kasus dengan menggunakan tabel pemetaan seperti gambar 9 tabel mapping dibawah ini.

ICD_Kasus	Gol_Penyakit	ICD_X
A00-B99:A88.1	Infeksi	A88.1
J00-J99 : J40	Respirasi	J40
R00-R99: R56.0	Simtoma	R56.0
K00-K99:K02.9	Digestif	K02.9

Gambar 9. Mapping Relation map_ICD_X

Dengan bantuan tabel pemetaan tersebut diatas maka relasi integrasi impor data dapat didefinisikan sebagai berikut :

```
create table RL_RS_A of RL_type  
as import from surveilans_RS_A.RL
```

```
create table RL_Data_Center of RL-type  
as import from data_center_SST.RL (  
gol_penyakit is @map_icd_x (kode_kasus, icd_kasus, gol_penyakit,  
null),  
icd_x is @map_icd_x (kode_kasus, icd_kasus, icd_x, null );
```

Pada langkah ke dua dari proses integrasi hubungan simantik adalah dengan menggabungkan relasi yang overlap atau tumpang tindih. Tumpang tindih ini bisa terjadi secara vertikal maupun horisontal. Dua jenis konflik dari relasi tersebut adalah terjadi pada jenis konflik struktural dan konflik semantik. Resolusi dari konflik tersebut dengan integrasi schema. Namun karena konflik level instance berkaitan atau berhubungan dengan konflik kelas, maka perlu dibahas keduanya secara bersama-sama dalam model schema transformation.

a. Konflik Struktural transformasi schema

Mewakili fakta dunia nyata dengan konsep pemodelan yang berbeda menghasilkan konflik struktural. Tergantung pada berbagai model data beberapa jenis konflik dapat timbul, tetapi konflik yang paling sering terjadi adalah partisi dan meta konflik. Partisi terjadi ketika relasi yang harus di integrasikan terjadi tumpang tindih vertikal, misal mewakili aspek yang berbeda dari relasi global namun masih mengandung semantik dan atribut yang setara. Meta konflik muncul ketika konsep direpresentasikan sebagai obyek data dalam satu skema, sedangkan konsep tersebut dimodelkan sebagai objek schema (relasi dan atribut). Konflik-konflik tersebut dapat diselesaikan di tingkat skema dengan menggabungkan operator untuk partisi dan restrukturisasi untuk meta konflik. Tetapi kita juga harus berurusan dengan konflik kesetaraan kunci dan konflik nilai atribut juga. Dalam model transformasi schema lebih lanjut dilakukan restrukturisasi melalui transformasi. Dilakukan dengan mengkonversi untuk baris dan kolom secara langsung begitu juga sebaliknya.

Kel-ICD-X	Bulan	Anak	Muda	Tua	Jumlah
Infeksi	Januari	5	7	3	6
Respirasi	Januari	7	5	5	8
Simtoma	Januari	6	7	9	12
Mata	Januari	3	6	6	8
Digestif	Januari	6	8	10	3

(a) Relasi Report_RL

Kel-ICD-X	Bulan	Jumlah	Gol-Usia
Infeksi	Januari	5	Anak
Infeksi	Januaru	7	Muda
Infeksi	Januaru	3	Tua
Respirasi	Januari	7	Anak
Respirasi	Januari	5	Muda
Respirasi	Januari	5	Tua

(b) Relasi Laporan_RL

Gambar 10. Nama dan struktur tabel yang berbeda untuk transformasi konflik integrasi

Untuk contoh kasus relasi dari sumber heterogen dengan perbedaan nama tabel dan struktur tabel tersebut diatas terjadi konflik integrasi akibat nama tabel dan struktur tabel berbeda. Perbedaan struktur tabel terletak pada grouping dan agregasi kolom. Untuk kasus diatas dapat diselesaikan dengan transformasi sebagai berikut :

```
Select *
From report_rl
Transpose to row
      (kel_icd_x, anak, 'anak', bulan),
      (kel_icd_x, muda, 'muda', bulan),
      (kel_icd_x, tua, 'tua', bulan),
As (kel_icd_x, bulan, jumlah, gol_usia);
```

Operasi inverse untuk mentransformasi kedalam baris adalah transformasi kedalam kolom yang mengambil subset dari relasi masukan yang mengandung nilai yang sama dalam kolom tertentu dan membangun suatu output tuple dengan kolom yang mewakili nilai tuple yang berbeda. Dengan cara ini relasi Laporan_RL bisa sesuai dengan struktur relasi diubah kembali menjadi relasi Report_RL.

```
Select *
From laporan_rl
Transpose to column
      (jumlah as anak when gol_usia = 'anak'
      Muda when gol_usia = 'muda'
      Tua when gol_usia = 'tua')
On kel_icd_x, bulan
As (kel_icd_x, bulan, anak, muda, tua, jumlah);
```

Pada bagian dari klausa yang menentukan atribut yang digunakan untuk mengidentifikasi kelompok tupel yang dialihkan tepat kesatu tupel. Operasi ini dilaksanakan mirip dengan operasi GROUP BY, meskipun kelompok yang dihasilkan berubah menjadi satu tupel per kelompok.

b. Konflik Multidatabase

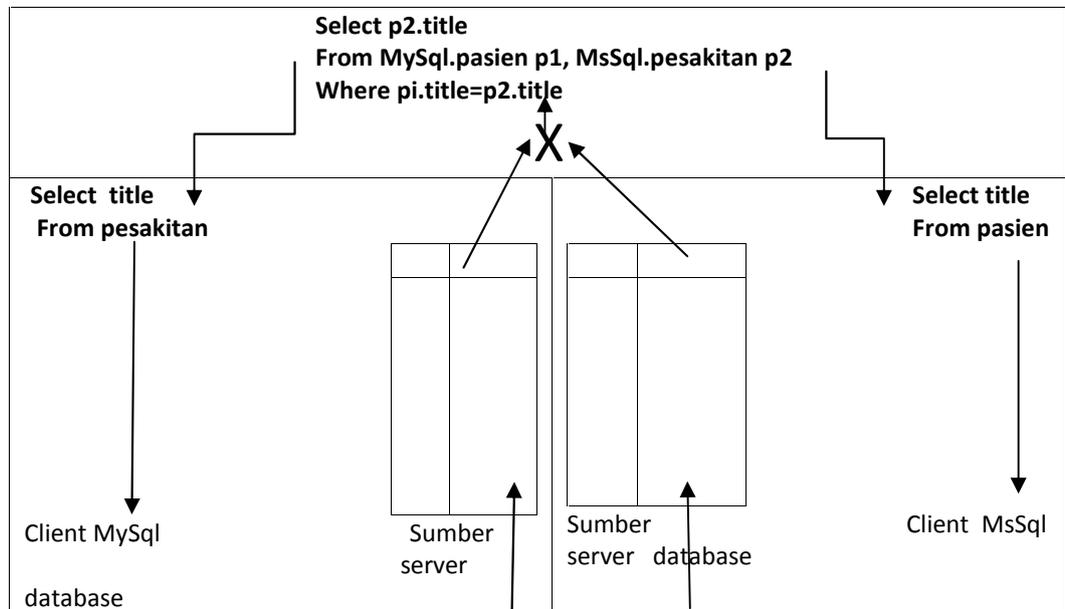
Konflik ini menggambarkan multidatabase schema yang mengharuskan ada akses ke beberapa database relasional. Pada contoh kasus ini melibatkan beberapa database relasional dalam dua dbms yang berbeda (multidatabase). Penggabungan dan pengintegrasian data menggunakan database klien (menggunakan masing masing API DBMS) dalam mengakses database relasional yang berbeda. Hal ini dapat mengatasi perbedaan antara “dialek” SQL yang berbeda. Sebuah aspek yang penting adalah bahwa multidatabase tidak menyembunyikan skema yang berbeda dari database komponen, skema multidatabase hanya persatuan skema dari database komponen (setelah mengubah nama, untuk menghindari konflik nama) seperti contoh kasus integrasi multi database dibawah ini.

Multidatabase Schema	
MySql.Pasien (<u>norm</u> , nama, kelamin, umur, alamat) MySql.Dokter (<u>kode_dkt</u> , spesialis, nama, afiliasi) MsSql.Pesakitan(<u>kdrm</u> , nama, usia, address) MsSql.Paramedis(<u>nmdepan</u> , <u>nmbkkg</u> , rmhsktasal)	
Pasien (<u>norm</u> , nama, kelamin, umur, alamat) Dokter (<u>kode_dkt</u> , spesialis, nama, afiliasi)	Pesakitan(<u>kdrm</u> , nama, usia, address) Paramedis(<u>nmdepan</u> , <u>nmbkkg</u> , rmhsktasal)
DBMS 1. MySql	DBMS 2. MsSql

Gambar 11. Skema Multidatabase

Gambar 11. tersebut menggambarkan penggunaan skema di multidatabase, komponen-komponen dari masing-masing skema dibawa dari masing-masing DBMS dan dapat diakses melalui klien. Masing-masing schema dapat di digabungkan

dan di integrasikan dalam operasional DML dengan cara menambahkan prefiks ke masing-masing nama tabel relasional seperti rancangan pada gambar 12. Dikawatir



Gambar 12. Rancangan Pola DML Multidatabase

5.3. Integrasi Dan Tranformasi (Unit Surveilans Dengan Data Center)

Menganalisis data dari sumber yang heterogen (unit surveilans) membutuhkan akses yang transparan dalam rangka untuk mengambil, menggabungkan, dan mengubah data yang relevan untuk di transformasikan ke data center. Hal ini dapat dilakukan dengan baik dengan memuat data fisik dalam database khusus (misalnya data center dalam lingkungan data warehouse dinas kesehatan atau dedicated database analisi). Manfaat dalam pendekatan integrasi dan transformasi database heterogen membutuhkan dukungan teknologi yang memiliki manfaat :

- Akses transpaaraan kesumber eksternal
- Mengintegrasikan hubungan dari sumber yang berbeda untuk digabung dalam data center.
- Menyelesaikan diskripsi dari konflik structural dan semantic dengan bantuan manipulasi penggantian nama operasi, konversi dan pembuatan fungsi pemetaan.

- Transformasi skema, seperti transposisi hubungan, serta menyelesaikan konflik meta data.
- Menyelesaikan perbedaan data menggunakan fungsi rekonsiliasi dan mendefinisikan agregasi.

Dalam penelitian ini digambarkan ilustrasi menggunakan skenario integrasi dengan data center yang melibatkan dua relasi, relasi RLUnitSurve berisi data kejadian kasus pada unit-unit surveilan pada kecamatan tertentu, RL_Agr berisi data kejadian secara agregat pada wilayah tertentu., skema datanya sebagai berikut :

RLUnitSurve (ICDX, JenisEpi, Kecamatan, Tahun)

RL_Agr(ICDX, Jumlah, Wilayah, Tahun)

5.3.1. Mengakses dan Mengintegrasikan Sumber Heterogen.

Teknologi SQL menyediakan akses hubungan yang diselenggarakan sumber eksternal, misalnya fitur fitur lengkap sistem database, dokumen terstruktur relasional. Akses ini dilakukan melalui adapter yang loadable pada runtime dan menterjemahkan pelayanan permintaan antar protocol multi database global schema. Disini dapat dibedakan dua jenis pandangan : Import memandang hubungan pemetaan dari sumber hubungan view global dan pandangan integrasi menggabungkan beberapa view schema local atau antar tabel. Untuk setiap atribut pemetaan dari source atribut skema local dapat didefinisikan dalam hal :

- Sebagai pengganti nama atribut yang sederhana
- Sebagai nilai konversi dengan menerapkan penggunaan fungsi
- Atau dengan menggunakan nama tabel pemetaan eksplisit untuk mengkonversi nilai-nilai tertentu.

Pandangan integrasi sebagai solusi pertama untuk transformasi data dengan menggunakan skema global dikombinasikan dengan menggunakan operator seperti union, join dan outerjoin.

5.3.2. Skema Tranformasi

Berdasar pada skema global pada data center yan ditransformasi lebih lanjut pada skema dari unit surveilans (RLUnitSurve dan RLAgr) dilakukan operasi seperti proyeksi, union dengan restrukturisasi melalui transposisi. Transporting berarti mengkonversi baris untuk kolom dan sebaliknya. Hasil transformasi dilakukan terhadap kedua relasi tersebut dapat di gambarkan sebagai berikut.

ICDX	Tahun	UnitSA	UnitSB	UnitSC
ICD-1	2008	15	18	22
ICD-2	2008	23	25	28

(a). RLAgrUnit

ICDX	Jumlah	Unit	Tahun
ICD-1	15	UnitSA	2008
ICD-1	18	UnitSB	2008
ICD-1	22	UnitSC	2008
ICD-2	23	UnitSA	2008
ICD-2	25	UnitSB	2008
ICD-2	28	UnitSC	2008

(b). RLAgr

Gambar. 13. Transpormasi Relasi

Hubungan relasi antar unit surveilans yang berisi epidemiaologi kasus (dalam ICDX) jumlah dan wilayah uint surveilans (dari UnitSA, UnitSB, UnitSC). Untuk diproses lebih lanjut misalnya pengelompokan dan agregasi, relasi sesuai dengan kedua relasi tersebut maka diperlukan rancangan fungsi TRANS-POSE-TO-ROW untuk setiap tupelnya sebagai berikut :

```

Select *
from RLAgUnit
transpose to rows
(ICDX, UnitSA, 'UnitSA', Tahun),
(ICDX, UnitSB, 'UnitSB', Tahun),
(ICDX, UnitSC, 'UnitSC', Tahun),
As (ICDX, Jumlah, Wilayah, Tahun);

```

Sedangkan subset kebalikannya dari operasi TRANS-POSE-TO-ROW adalah TRANS-POSE-TO-COLUMNS yang mengambil input dari relasi RLAg sebagai berikut :

```

Select *
Fom RLAg
Transpose to columns
( Jumlah As UnitSA when unit='UnitSA'
      UnitSB when unit='UnitSB'
      UnitSC when unit='UnitSC')
On ICDX, Tahun
As (ICDX, Tahun, UnitSA, UnitSB, UnitSC);

```

Operasi ini dilakukan mirip dengan pengelompokan Operasi Group By, meskipun kelompok (group) yang dihasilkan berubah menjadi satu tuple per kelompok.

BAB 6

KESIMPULAN DAN SARAN

Integrasi instance merupakan aspek penting untuk mengintegrasikan sumberdata heterogen. Pada kasus penelitian ini pemetaan antara obyek dan sumber data yang berbeda harus didefinisikan dan dihilangkan perbedaan-perbedaan (baik struktur, tabel maupun isi) dalam representasi data. Karena itu harus dibahas definis dalam global view hasil integrasi skema. Dengan demikian konflik skema dapat dihindarkan dan dapat dilanjutkan dengan integrasi dan DML antar database (multidatabase). Langkah-langkah utama dalam perancangan integrasi antar databse heterogen meliputi : 1. Seleksi data, dimana dilakukan identifikasi data untuk memilih dan menganalisis atribut dan tuple yang diperlukan diantara skema lokal dan skema global. 2. Integrasi data, menggabungkan data dari berbagai skema yang sudah dipilih atrut dan tupel yang relevan tersebut. 3. Transformasi data, dilakukan transformasi data dengan menggunakan fungsi pada struktur serta persyaratan sintaksis, misalnya normalisasi, scaling dan pemisahan nilai. Dalam makalah ini kami menyajikan dua konflik struktur RL epidemiologi (server surveilans_ RS_A dan Data_Center_SST) untuk menggambarkan konflik. Setelah dilakukan representasi konflik maka bisa dilakukan pemetaan kode_kasus setelah mapping schema (Relation map_ICD_X). setelah proses tersebut maka bisa dilakukan integrasi dan import data serta dapat dilakuka integrasi DML antar database heterogen (multidatabase). Sedangkan untuk menggambarkan rancangan trsformasi menggunakan relasi RLUnitSurv dan RLAgr. Saran dalam penelitian ini adalah dapat di inventasiasai semua kemungkinan konflik integrasi antar database dan schema surveilans sehingga dapa dibuat semacam integrasi dengan schema generator yang dinamis. Tahapan konsentrasi penelitian berikutnya yang perlu diperhatikan adalah berkaitan dengan replikasi antar unit surveilans ke data sencer SST (sistem surveilans Terpadu) sebagai data center, pembersihan data dan reduksi data antara schema lokal dan schema global.

DAFTAR PUSTAKA

- A. Silberschatz, H.F. Korth, S. Sudarshan, 2005, *Database System Concept*, 5th ed., McGraw-Hill Publishing Company, Boston 2005.
- A. Tannenbaum, Marten van Steen, 2008, *Distributed System : Principles and Paradigms*, 5th ed., Prentice Hall, 2008.
- Balanov Alexander and Natalia Janson, 2010, *Shyncronization : From Simple To Complex*, Springer Series in Synergetics Ser, Dewey edition, Springer, ISBN 9783642091285, 2010.
- C.J. Date, 2005, *An Introduction and Database System*, Pearson Education, Addison-Wesley, Boston, USA, 2005.
- Dollimore Jean, 2012, Jose Coulourise and Tim Kindberg, *Distributed System, Concept and Design*. Fifth Edition. Parason Education, Addison-Wesley, Boston, USA, ISBN 9780132143011, 2012.
- Eddy Purwanto, 2012, *Perbandingan Strategi Replikasi Pada Sistem Basisdata Terdistribusi*, Jurnal Universitas Bina Darma Palembang Vol 5. No. X, 2012
- Keputusan Menteri Kesehatan Republik Indonesia No.1479/MENKES/SK/X/2003 tentang Pedoman Penyelenggaraan Sistem Surveilans Epidemiologi Penyakit Menulir Dan Penyakit Tidak Menular Terpadu.
- Retantyo Wardoyo, 2012, *Perbedaan Kode dalam rancangan Database Dan Strategi Penyelesaiannya Untuk Sinkronisasi Data*, Jurnal IPTE-KOM, UGM, Vol. 14. No. 2, 2012.
- Sayed Tossy Messas, 2012, *Sistem Monitoring Dan Notifikasi Epidemologi berbasis Webgis*, Jurnal KITEKTRO. Universitas Syah Kuala, Aceh Vol. 1. No. 11, 2012.
- www.learning.unl.ac.uk/csp003n/lectures/w021-ddb-arch.pdf, diakses : 07-02-2013.

LAMPIRAN - LAMPIRAN

Lampiran 1. : Personalia Tenaga Peneliti Beserta Kualifikasinya

No	Nama/NIDN	Instansi Asal	Bidang Ilmu / Kualifikasi	Alokasi Waktu (jam / minggu)	Uraian Tugas
1.	Muslih, M.Kom	Fak. Ilmu Komputer UDINUS Semarang	Ilmu Komputer/ Database Desainer & DBA	5 jam	<p>Ketua Peneliti</p> <ol style="list-style-type: none"> 1. Mengkoordinir tugas anggota peneliti 2. Mengkoordinasikan dan mengkomunikasikan dengan Obyek penelitian dan pihak eksternal yang terkait 3. Menentukan metode penelitian 4. Menentukan Pola Kaidah Integrasi Schema Local dan Schema Global 4. Menentukan Analisis, Identifikasi dan desain pola kaidah terhadap masalah integrasi data heterogen 5. Merumuskan Hipotesa 6. Menguji, Memantau dan mengendalikan pelaksanaan penelitian 7. Melaporkan hasil penelitian 8. Mempublikasikan
2.	Elkaf Raahmawan Pramudya, M,Kom	Fak. Ilmu Komputer UDINUS Semarang	Ilmu Komputer/ Jaringan Komputer	5 jam	<p>Anggota 1</p> <ol style="list-style-type: none"> 1. Bendahara dan administrasi 2. Menyusun Kuisisioner 3. Mengkoordinasikan dengan mahasiswa 4. Hitung dan analisa pola kaidah sinkronisasi dan integrasi 5. Evaluasi dan Refisi Fungsi sinkronisasi dan untegrasi database heterogen. 6. Menyusun requirement spesifikasi fungsi dan sinkronisasi data epidemiologi.

Lampiran 2. : Publikasi

NO.	Tanggal	Jenis Pulikasi	Tempat /Nama	Keterangan
1.	23 Agustus 2014	Prosiding SNATIF	UMK KUDUS	Sebagai Pembicara
2.	1 Febuari 2014	Jurnal Techno	FIK/UDINUS	Submit

**DESAIN POLA STRUKTUR MAPPING SCHEMA
UNTUK SINKRONISASI DAN INTEGRASI MULTIDATABASE TERDISTRIBUSI
DALAM MENGELOLA DATA EPIDEMIOLOGI**

Muslih^{1*}, Elkaf R, S.S. Nurhendratno²

¹ Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Nakula I, No. 5-11, Kota Semarang

² Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro
Jl. Nakula I, No. 5-11, Kota Semarang

*Email: slametalica301@gmail.com

Abstrak

Pengaruh globalisasi terhadap perusahaan yang semakin luas area bisnisnya, maka organisasi menuntut inovasi teknologi informasi yang dapat mengelola peningkatan jumlah data dan skalabilitas jarak transaksi (antara sistem aplikasi dengan database maupun antar database itu sendiri). Perusahaan yang memiliki cabang diberbagai lokasi yang semakin jauh dan meluas, maka sistem aplikasi komputer memerlukan pilihan arsitektur basis data yang optimal dalam mengimbangi perkembangan bisnis tersebut termasuk yang berkaitan dengan model distribusi dan integrasi datanya. Seperti pada pengelolaan data epidemiologi kesehatan, dimana sumber data tersebar pada database yang ada pada berbagai lokasi rumah sakit dan poliklinik pada suatu wilayah kabupaten atau suatu kota tertentu. Permasalahannya database sumber (source) bersifat heterogen sehingga mengalami potensi konflik (kesulitan) dalam melakukan integrasi menuju pada pusat data epidemiologi (target) pada dinas kesehatan. Potensi konflik yang terjadi adalah ketidakseragaman skema relasi (konflik skema), ketidakakuratan isi (konflik data). Untuk itu dalam integrasi memerlukan analisis database sumber yang bersifat heterogen dengan melakukan strukturisasi dan sinkronisasi sebagai persiapan integrasi data. Dengan permasalahan integrasi antar database distribusi tersebut maka dalam penelitian ini bertujuan mendesain arsitektur database terdistribusi dengan metode replikasi yang akan diimplementasikan pada integrasi database epidemiologi sehingga akan didapatkan sebuah arsitektur database terdistribusi yang bisa mengatasi ketersediaan data pada sistem surveilans terpadu (SST). Untuk mengatasi masalah tersebut maka diperlukan suatu pengembangan arsitektur basis data terdistribusi untuk pola sinkronisasi dan integrasi dengan metode replikasi mapping schema. Metode replikasi mapping schema generator merupakan kemajuan dari rekayasa konsep teknologi DDBMS (Distributed Database Management System) yang mampu melakukan sinkronisasi (captures, routes, transforms) dan integrasi data yang bersifat heterogen secara real-time. Proses utama dalam metode replikasi mapping schema mencakup proses mendeteksi konflik skema, representasi schema matching, replikasi Integration common dan schema mapping. DDBMS mengelola data yang tersebar di site-site dalam sebuah jaringan atau node-node dari sebuah sistem multiprocessor. Prinsip basis data terdistribusi adalah suatu basis data dengan skema global (global schema) yang berada dibawah kendali sistem manajemen basis data (DBMS) terpusat dengan piranti penyimpanan data yang terpisah-pisah dalam skema lokal (local schema) dalam suatu jaringan komputer. Tujuan jangka panjang dalam penelitian ini adalah merancang teknologi dan aplikasi dalam mengembangkan teknik integrasi data dari berbagai ragam aplikasi dan database tanpa harus menyeragamkan aplikasi dan database yang sudah ada. Dengan demikian schema local dapat dipertahankan dalam mendapatkan schema global melalui teori rekayasa sinkronisasi dan integrasi basis data. Sedangkan target khusus yang akan dicapai adalah memperoleh model arsitektur database tersebar untuk integrasi data yang dapat diterapkan dalam mengelola dan mengembangkan sistem informasi epidemiologi terintegrasi pada dinas kesehatan dengan metode replikasi mapping schema. Metode penelitian yang digunakan adalah dengan studi literatur dan studi lapangan. Setelah melakukan studi awal kegiatan penelitian dilanjutkan dengan observasi dan studi pustaka, analisa permasalahan dalam perancangan arsitektur

database. Tahapan berikutnya adalah melakukan desain pola integrasi database dan dilakukan uji integrasi dan replikasi untuk mendapatkan kesimpulan integrasi antar database heterogen. Hasil dalam kasus penelitian ini adalah integrasi antar dua relasi yang terjadi konflik (*surveila_rs_A* dan *data_center_SST*) menggunakan *mapping schema* (*relasi_map_ICD_X*).

Kata kunci: DDBMS, *schema global*, *Schema local*, sinkronisasi

1. PENDAHULUAN

1.1. Latar Belakang

Indikator kesehatan masyarakat sangat erat kaitannya dengan epidemiologi suatu kasus pada suatu daerah tertentu. Epidemiologi adalah wabah penyakit terutama yang menular secara cepat dan tak terduga pada suatu wilayah tertentu. Agar wabah tidak meluas ekskalasinya maka diperlukan sistem monitoring untuk mengembangkan suatu metode dalam menganalisis secara sistematis keadaan dan keberadaan suatu penyakit dalam upaya untuk mengatasi dan menaggulangi secara cepat dan terintegrasi. Untuk itu Departemen Kesehatan telah mengeluarkan keputusan menteri **No. 1479/MENKES/SK/X/2003** tentang : Pedoman Penyelenggaraan Sistem Surveilans Epidemiologi Penyakit Menular Dan Penyakit Tidak Menular Terpadu. Dalam pedoman surveilans tersebut menegaskan diperlukannya suatu Sistem Surveilans Terpadu (*SST*) dengan dukungan basis data yang setandar dimana sistem pengawasan utama epidemiologi meliputi semua unit pelayanan kesehatan (Puskesmas, Laboratorium, Rumah Sakit) di semua Pemerintah Daerah Propinsi dan Pemerintah Daerah Kabupaten/Kota dengan model : Sistem Pencatatan Pelaporan Puskesmas Terpadu (*SP2PT*) dan Sistem Pelaporan Rumah Sakit (*SPRS*). Di tingkat pemerintah daerah pelaksanaan operasional *SST* tersebut sepenuhnya diserahkan kepada dinas kesehatan daerah untuk bisa menjadi sistem informasi epidemiologi dalam rangka mendukung pemberantasan penyakit menular dan tidak menular secara nasional. Direktorat Jendral Pemberantasan Penyakit Menular dan Penyehatan Lingkungan (Dirjen PPM & PL Departemen Kesehatan) sebagai lembaga pemerintah pusat yang mendapat tugas dan bertanggung jawab dalam bidang pengendalian maupun pemberantasan penyakit secara nasional. Namun dalam pelaksanaan dan penyelenggaraan sistem surveilans terpadu (*SST*) tersebut ditingkat kabupaten/kota menghadapi suatu kendala dalam melakukan pengiriman data kesehatan ke dinas kesehatan kabupaten/kota sebagai penanggung jawab kesehatan di tingkat pemerintah daerah. Kendala ini disebabkan karena sumber data unit surveilans (puskesmas, laboratorium, rumahsakit) terletak secara geografis tersebar dan diperoleh dari beragam aplikasi dan database manajemen sistem (*DBMS*) yang beragam (heterogen). Sinkronisasi dan integrasi data antara unit surveilans (puskesmas, laboratorium, rumahsakit) dengan dinas kesehatan kabupaten/kota menjadi masalah utamanya. Hal ini disebabkan karakter masing-masing unit surveilans memiliki ketidak seragaman platform aplikasi dan database (*heterogen*). Untuk itu dibutuhkan suatu pola kaidah sinkronisasi dan integrasi data antar unit surveilans (skema local) dengan dinas kesehatan sebagai "**Data Center**" (skema global) dalam model aljabar relasi dan notasi skema relasi (**relation schema**) serta skema basis data (**database schema**). Sehingga bisa menjadi modal ilmiah dalam membangun basis data terdistribusi untuk mendukung sistem informasi terpadu epidemiologi sebagai pelaksanaan sistem surveilans terpadu (*SST*) tingkat kabupaten/kota.

1.2. Masalah Integrasi Pada Multidatabase

Integrasi multidatabase dirancang untuk mendapatkan informasi yang terpadu, umumnya bertujuan untuk menggabungkan sistem yang dipilih sehingga akan membentuk satu kesatuan dalam sistem informasi dalam berinteraksi. Pengguna akan disajikan sebuah logical view data homogen, walaupun secara fisik didistribusikan atau dialokasikan dari sumber data yang heterogen. Untuk itu semua data harus dipresentasikan dari prinsip-prinsip abstraksi yang sama (satu model data global dan satu semantik). Sehingga dihadapkan pada tugas mendeteksi dan resolusi skema data yang berkaitan dengan konflik struktur serta semantiknya. Konflik struktur dan semantik dalam integrasi data base disebabkan adanya beberapa heterogenitas yang berkaitan dengan hardware, sistem operasi, DBMS, data model, schema data, semantik data, middleware, user interface dan kendala aturan bisnis. Beberapa konflik yang dapat terjadi pada integrasi multidatabase yang berkaitan dengan skema relasi maupun keakuratan data yang tidak seragam adalah :

ii) Konflik Antar Tabel

(a) Konflik Antar Dua Tabel

- (i) Nama Tabel (homonim/sinonim), dapat diselesaikan dengan view.
- (ii) Struktur Tabel, seperti jumlah atribut berbeda di dua tabel yang informasinya sama, dapat diselesaikan dengan membuang atribut yang keberadaannya tidak disemua relasi, atau menambah atribut yang kurang pada relasi yang kekurangan atribut.
- (iii) Integrity Constrain, misalnya pada dua situs yang terdapat pada tabel yang sama, tetapi isi atribut primary key-nya berbeda, dapat diselesaikan dengan menambahkan primary key tambahan yang berisi informasi situs relasi tersebut disimpan.

(b) Antar Banyak Tabel, misalnya pada dua komponen basis data jumlah relasinya tidak sama tetapi informasinya sama, dapat diselesaikan dengan penggabungan relasi dan view .

(2) Konflik Antar Atribut

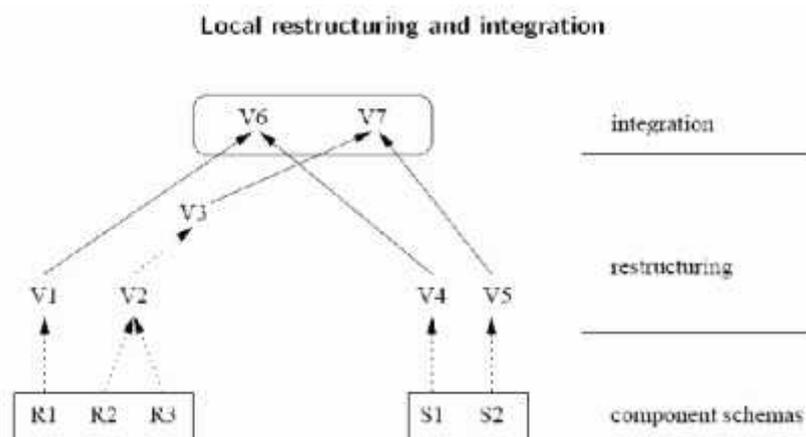
(a) Antar Dua Atribut

- (i) Nama Atribut (Homonim/Sinonim), dapat diselesaikan dengan penggantian nama (rename) atribut di view.
- (ii) Integrity Constraint, misal tipe data dapat diselesaikan dengan fungsi-fungsi konversi, seperti `to_char(int)`, atau `to_int(char)` pada view.

(b) Antar Banyak Atribut, misalnya dalam penyampaian informasi nama orang dalam tabel yang satu digunakan dua atribut (kolom), nama depan dan nama belakang, sementara pada tabel yang lain digunakan satu atribut nama lengkap. Konflik ini dapat diselesaikan dengan penggabungan string atau pemisahan string dengan fungsi substring.

(3) Konflik Atribut – Tabel, dapat merupakan kombinasi dari permasalahan diatas.

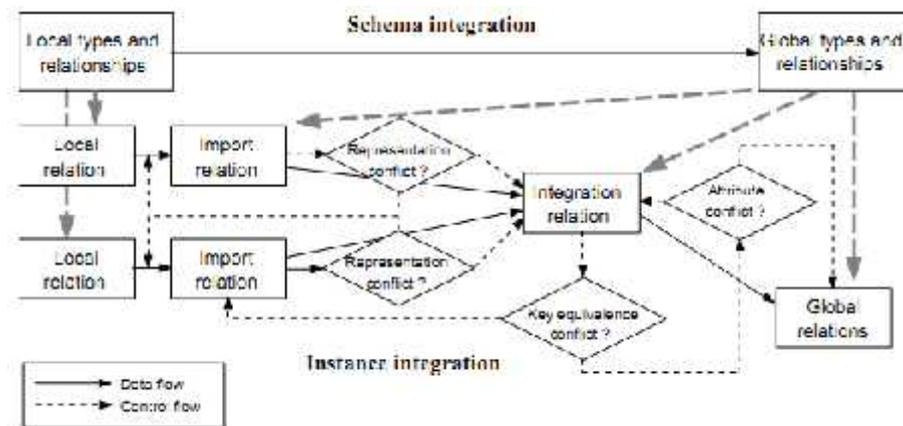
Penyelesaian konflik tersebut diatas yang berkaitan dengan integrasi basis data harus memerlukan analisis yang mendalam akan komponen basis data dan tidak bisa di otomatisasi. Sebelum melakukan integrasi, komponen basis data harus dipersiapkan terlebih dahulu untuk menagani konflik. Proses penyelesaik konflik ini disebut restrukturisasi basis data.



Gambar 1. Bagan Proses Restrukturisasi dan Integrasi Basis Data.

1.3. Replikasi

konflik instance dan untuk mengatasinya dengan melakukan konversi dan resolusi fungsi dimana penerapan impor relasi sebagaimana entended join dan operasi union. Proses ini secara keseluruhan dapat digambarkan seperti di bawah ini.

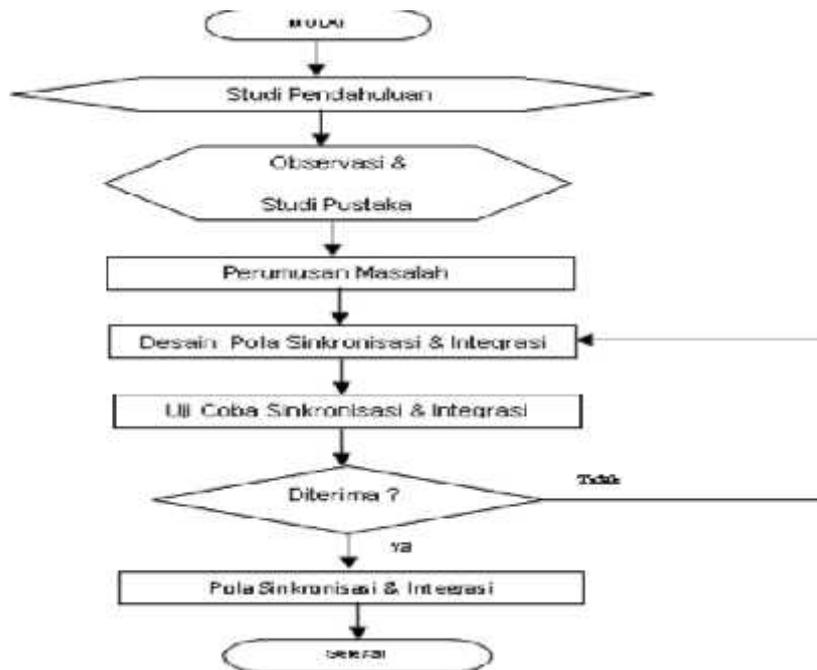


Gambar 3. Proses Integrasi Skema

2. METODOLOGI

Pada penelitian ini digunakan pendekatan metode studi literatur (*library research*) dan studi lapangan (*field research*) untuk mendesain pola sinkronisasi. Adapun tahapan penelitian dapat dijelaskan dalam langkah-langkah dibawah ini :

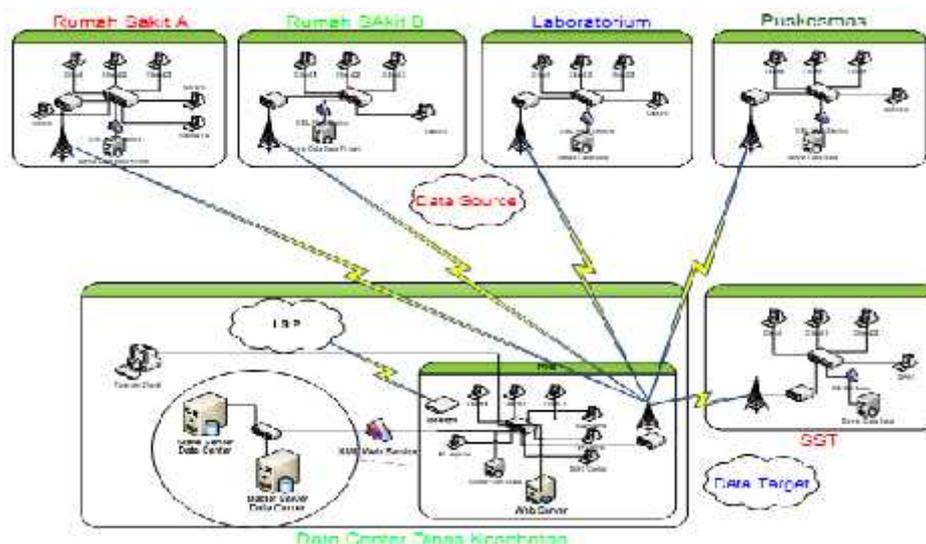
6. Studi Pendahuluan : Pada tahapan ini merupakan kegiatan untuk mengenali lebih lanjut tentang obyek penelitian beserta lingkungan yang terkait dalam rangka mendalami situasi dan kondisi dari sinkronisasi dan integrasi yang akan dikembangkan. Studi pendahuluan dilakukan dengan mengumpulkan informasi mengenai pengelolaan sistem surveilans terpadu (*SST*).
7. Observasi dan Studi Pustaka : Pada tahapan ini akan dilakukan analisis kebutuhan dengan analisa diskriptif dengan cara melakukan kajian pustaka yang terkait dengan konsep sinkronisasi dan integrasi data dan keberadaan data kesehatan yang digunakan untuk surveilans keshatan pada masing-masing unit surveilans dengan dinas kesehatan.
8. Perumusan Masalah : Tahap selanjutnya setelah mendapatkan permasalahan utama dari obyek penelitian dan dilengkapi dasar teori dari studi pustaka yang mendukung, adalah merumuskan permasalahan yang akan dieksplorasi dalam rangka menemukan pola baru dalam sinkronisasi data kesehatan antar unit surveilans dengan database epidemiologi kesehatan.
9. Desain Pola Sinkronisasi dan Integrasi Data : Pada tahapan ini dilakukan desain pola dan notasi sistem sinkronisasi dan integrasi database sebagai pola integrasi data berdasarkan diagnosis dan identifikasi masalah yang ada, baik dari sisi teori dan teknis maupun implementasi dalam bentuk pola notasi model relasional sinkronisasi dan integrasi sistem distribusi.
10. Pengujian : Pada tahap pengujian ini dilakukan evaluasi terhadap hasil desain dari pola notasi model relasional sinkronisasi dan Integrasi dengan aturan de morgans sehingga nantinya pola sinkronisasi database dapat digunakan secara maksimal dan sesuai (diterima) sebagai suatu pola baru untuk mendukung pengembangan system informasi epidemiologi kesahatan dari lingkungan database distribusi yang heterogen.
6. Hasil Pola Sinkronisasi dan Integrasi (notasi model relasional) : Hasil dari pola notasi model relasional sinkronisasi dan integrasi data ini merupakan sistem sinkronisasi database berbasis proses replikasi data dari masing-masing unit surveilans yang heterogen (puskesmas, laboratorium, rumah sakit) menuju database epidemiologi dinas kesehatan sebagai data center kesehatan. Sehingga pola ini dapat mempermudah dan mendukung dalam akses data untuk diolah menjadi system informasi epidemiologi kesehatan kabupaten/kota.



Gambar 4. Diagram Alir Penelitian

3. HASIL DAN PEMBAHASAN

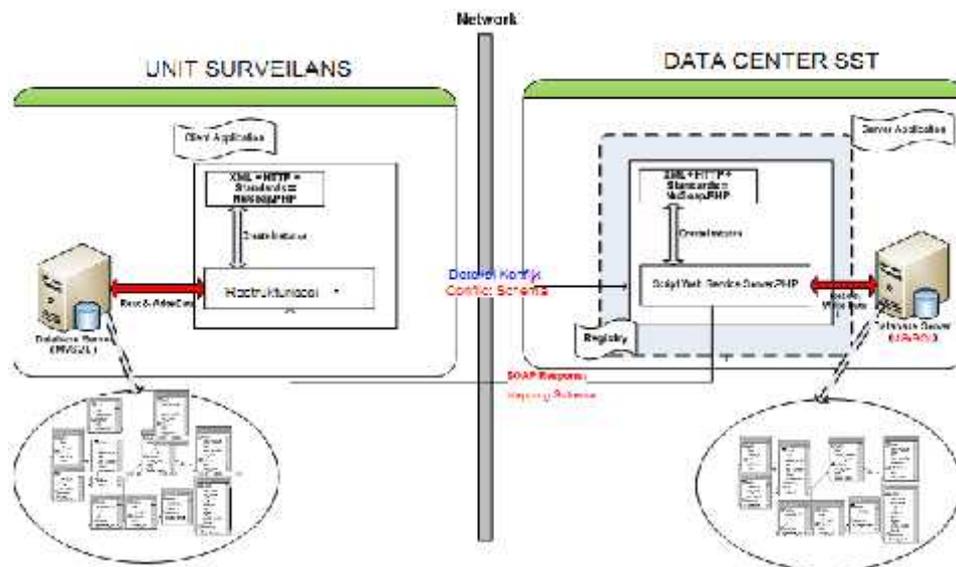
Rancangan arsitektur replikasi dan integrasi epidemiologi antar unit surveilans (puskesmas, laboratorium, rumah sakit) terletak pada area geografis yang tersebar dan data (source) di peroleh dari sistem aplikasi dan dbms yang beragam (heterogen). Dalam pelaksanaan sistem surveilans terpadu (SST) memerlukan rekayasa untuk keperluan integrasi agar sumber data (source data) dapat di direplikasi atau di distribusikan pada alokasi data senter (target) pada server sistem surveilans terpadu (SST) di dinas kesehatan kota/kabupaten. Rekayasa tersebut untuk menghindari konflik schema dari heterogenitas sumber data surveilans menuju data senter epidemiologi.



Gambar 5. Arsitektur Integrasi Epidemiologi

3.1. Mendeteksi Konflik Antar Skema Unit Surveilans Dengan Skemaa Data Center

Heterogenitas pada data model, skema dan level instance akan menyebabkan berbagai macam konflik, konflik tersebut menjadikan permasalahan untuk integrasi dan replikasi data dari sumber data surveilans (puskesmas, poliklinik, rumah sakit) dengan data center pada server sistem surveilans terpadu (SST) di dinas kesehatan sebagai target replikasi. Untuk beberapa konflik level skema terdapat beberapa klasifikasi, pada makalah ini kami batasi pada kasus konflik level instance.



Gambar 6. Deteksi Konflik Skema level instance

Pada percobaan mengatasi konflik instan pada uraian penyelesaian ini, kami menggunakan 2 buah contoh tabel yang memiliki berbagai jenis konflik level instance muncul tidak ketergantungan satu dengan yang lainnya. Sebagai jenis konflik utama menyebutnya sebagai konflik representasi. Hal ini mengacu pada representasi yang berbeda tetapi memiliki nilai data yang sesuai dengan fakta yang sama. Hal ini dapat disebabkan, misalnya oleh pengukuran unit yang berbeda (misal dolar vs rupiah), notasi yang berbeda (misal, firstname lastname vs lastname, firstname), atau perbedaan representasi (misalnya, ISBN dengan strip dan tidak pakai strip). Selama konflik representasi integrasi akan dapat mengakibatkan konflik kesetaraan kunci ketika instance hubungan yang berbeda merujuk pada merujuk pada obyek yang sama tetapi berbeda pula dalam mengidikan obyek tersebut. Hal tersebut menjadi situasi konflik hubungan dan terjadi tumpang tindih semantik. Hal ini memerlukan penyelesaian dengan menambahkan kelas konflik lanjutan yang mengacu pada konflik hubungan tersebut. Dibawah ini merupakan contoh konflik dua relasi dari dua server data sumber epidemiologi yang berbeda, tetapi digunakan untuk merepresentasikan obyek yang sama.

Gol_Penyakit	Bulan	ICD-X	Nama_Diagnosa	Gol_Umur	Kasus_Lama	Kasus_baru	Jumlah
Infeksi	Januari	A88.1	Infeksi Epidemik V	Muda	6	8	14
Respirasi	Januari	J40	Bronchitis N A	Muda	8	11	19
Simtoma	Januari	R56.0	Vebril Convulsion	Tua	12	17	29
Mata	Januari	H10.9	Conjunctivis Unspe	Tua	8	20	28
Degestif	Januari	K02.9	Dental Caries	Anak	3	13	16

(a) Relasi dari server surveilans_RS_A RL

Kode_Kasus	Bulan	Nama_Diagnosa	Gol_Umur	Kasus_Lama	Kasus_baru	Jumlah
A00-B99:	Januari	Infeksi Epidemik	Tua	4	5	9

A88.1		vertigo				
J00-J99 : J40	Januari	Bronchitis N A	Anak	10	4	14
R00-R99: R56.0	Januari	Vebrile Convulsions	Muda	7	11	18
K00- K99:K02.9	Januari	Dental Caries	Muda	9	12	21

(b) Relasi dari server Data_Center_SST RL

Gambar 7. Relasi Data Surveilans sumber dan target

Dalam rangka untuk mendapatkan petunjuk jenis konflik (deteksi) pada langkah integrasi dari relasi server surveilans RS A dan relasi dari data center SST tersebut, maka kita harus mempertimbangkan proses integrasi. Langkah awalnya adalah mendiskripsikan konflik pada level skema diselesaikan dengan mendefinisikan pemetaan atribut untuk import relasi.

3.2. Representasi Konflik

Sebagai pembahasan untuk representasi resolusi konflik, maka langkah selanjutnya sekenarionya adalah bahwa database pelaporan epidemiologi (sumber, standar laporan epidemiologi RL.2a) pada relasi surveilans RS A harus diintegrasikan pada server data center SST. Hubungan yang terserstruktur yang ditunjukkan gambar 7. diatas, jelas kita dapat memperkenalkan jenis kasus pasien dari ke dua relasi tersebut. Tetapi karena data center SST menggunakan schema sendiri untuk ICD-X , sehingga tidak mungkin dilakukan transformasi dan integrasi langsung. Untuk itu diperlukan pemetaan kode_kasus dengan menggunakan tabel pemetaan seperti gambar 8 tabel mapping dibawah ini.

ICD_Kasus	Gol_Penyakit	ICD_X
A00-B99:A88.1	Infeksi	A88.1
J00-J99 : J40	Respirasi	J40
R00-R99: R56.0	Simtoma	R56.0
K00-K99:K02.9	Digestif	K02.9

Gambar 8. Mapping Relation map_ICD_X

Dengan bantuan tabel pemetaan tersebut diatas maka relasi integrasi impor data dapat didefinisikan sebagai berikut :

```
create table RL_RS_A of RL_type
as import from surveilans_RS_A.RL
```

```
create table RL_Data_Center of RL-type
as import from data_center_SST.RL (
gol_penyakit is @map_icd_x (kode_kasus, icd_kasus, gol_penyakit, null),
icd_x is @map_icd_x (kode_kasus, icd_kasus, icd_x, null );
```

Pada langkah ke dua dari proses integrasi hubungan simantik adalah dengan menggabungkan relasi yang overlap atau tumpang tindih. Tumpang tindih ini bisa terjadi secara vertikal maupun horisontal. Dua jenis konflik dari relasi tersebut adalah terjadi pada jenis konflik struktural dan konflik semantik. Resolusi dari konflik tersebut dengan integrasi schema. Namun karena konflik level instance berkaitan atau berhubungan dengan konflik kelas, maka berlu dibahas keduanya secara bersama-sama dalam model schema transformation.

a. Konflik Struktural transformasi schema

Mewakili fakta dunia nyata dengan konsep pemodelan yang berbeda menghasilkan konflik struktural. Tergantung pada berbagai model data beberapa jenis konflik dapat timbul, tetapi konflik yang paling sering terjadi adalah partisi dan meta konflik. Partisi terjadi ketika relasi yang harus diintegrasikan terjadi tumpang tindih vertikal, misal mewakili aspek yang berbeda dari relasi global

namun masih mengandung semantik dan atribut yang setara. Meta konflik muncul ketika konsep direpresentasikan sebagai obyek data dalam satu skema, sedangkan konsep tersebut dimodelkan sebagai objek schema (relasi dan atribut). Konflik-konflik tersebut dapat diselesaikan di tingkat skema dengan menggabungkan operator untuk partisi dan restrukturisasi untuk meta konflik. Tetapi kita juga harus berurusan dengan konflik kesetaraan kunci dan konflik nilai atribut juga. Dalam model transformasi schema lebih lanjut dilakukan restrukturisasi melalui transformasi. Dilakukan dengan mengkonversi untuk baris dan kolom secara langsung begitu juga sebaliknya.

Kel-ICD-X	Bulan	Anak	Muda	Tua	Jumlah
Infeksi	Januari	5	7	3	6
Respirasi	Januari	7	5	5	8
Simtoma	Januari	6	7	9	12
Mata	Januari	3	6	6	8
Digestif	Januari	6	8	10	3

(a) Relasi Report_RL

Kel-ICD-X	Bulan	Jumlah	Gol-Usia
Infeksi	Januari	5	Anak
Infeksi	Januaru	7	Muda
Infeksi	Januaru	3	Tua
Respirasi	Januari	7	Anak
Respirasi	Januari	5	Muda
Respirasi	Januari	5	Tua

(b) Relasi Laporan_RL

Gambar 9. Nama dan struktur tabel yang berbeda untuk transformasi konflik integrasi

Untuk contoh kasus relasi dari sumber heterogen dengan perbedaaan nama tabel dan struktur tabel tersebut diatas terjadi konflik integrasi akibat nama tabel dan struktur tabel berbeda. Perbedaan struktur tabel terletak pada grouping dan agregasi kolom. Untuk kasus diatas dapat diselesaikan dengan transformasi sebagai berikut :

```

Select *
From report_rl
Transpose to row
    (kel_icd_x, anak, 'anak', bulan),
    (kel_icd_x, muda, 'muda', bulan),
    (kel_icd_x, tua, 'tua', bulan),
As (kel_icd_x, bulan, jumlah, gol_usia);

```

Operasi inverse untuk mentransformasi kedalam baris adalah transformasi kedalam kolom yang mengambil subset dari relasi masukan yang mengandung nilai yang sama dalam kolom tertentu dan membangun suatu output tuple dengan kolom yang mewakili nilai tuple yang berbeda. Dengan cara ini relasi Laporan_RL bisa sesuai dengan struktur relasi diubah kembali menjadi relasi Report_RL.

```

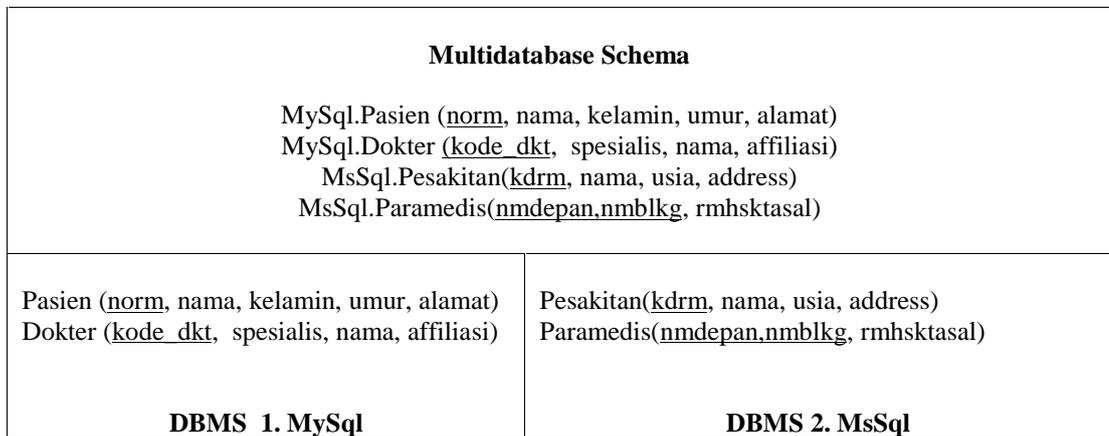
Select *
From laporan_rl
Transpose to column
    (jumlah as anak when gol_usia = 'anak'
    Muda when gol_usia = 'muda'
    Tua when gol_usia= 'tua')
On kel_icd_x, bulan
As (kel_icd_x, bulan, anak, muda, tua, jumlah);

```

Pada bagian dari klausa yang menentukan atribut yang digunakan untuk mengidentifikasi kelompok tupel yang dialihkan tepat kesatu tupel. Operasi ini dilaksanakan mirip dengan operasi GROUP BY, meskipun kelompok yang dihasilkan berubah menjadi satu tupel per kelompok.

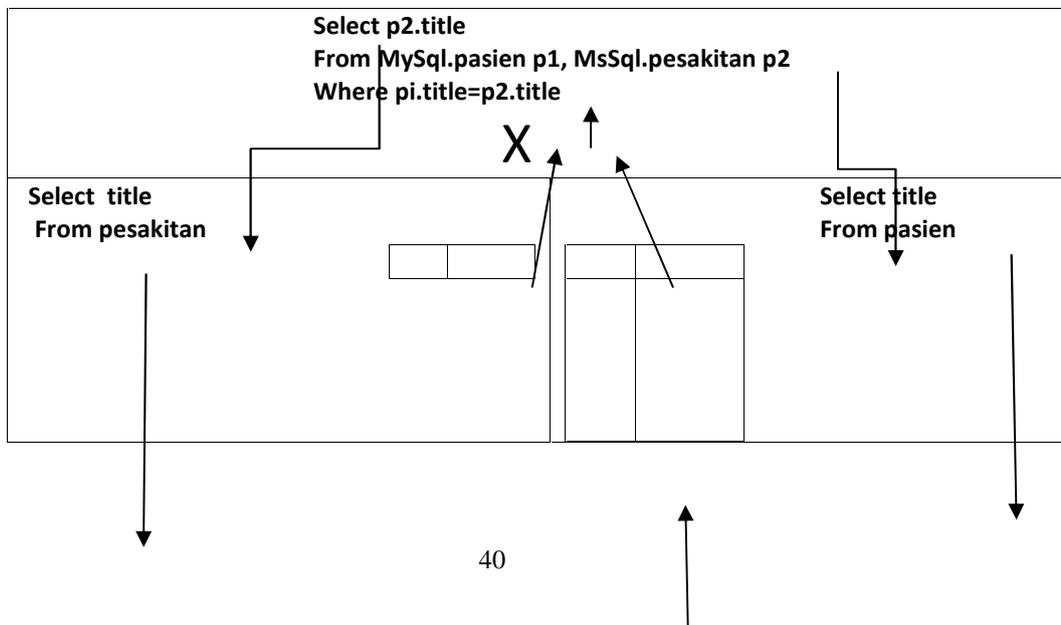
b. Konflik multi database

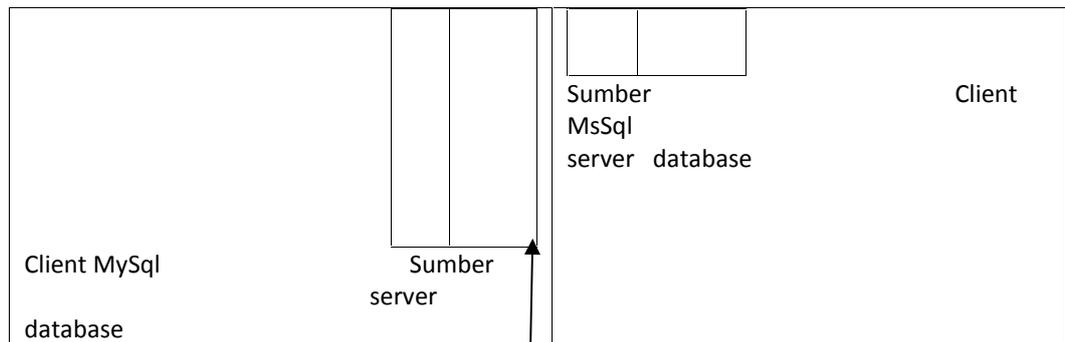
Konflik ini menggambarkan multidatabase schema yang mengharuskan ada akses ke beberapa database relasional. Pada contoh kasus ini melibatkan beberapa datanase relasional dalam dua dbms yang berbeda (multidatabase). Penggabungan dan pengintegrasian data menggunakan database klien (menggunakan masing masing API DBMS) dalam mengakses database relasional yang berbeda. Hal ini dapat mengatasi perbedaan antara “dialek” SQL yang berbeda. Sebuah aspek yang penting adalah bahwa multidatabase tidak menyembunyikan skema yang berbeda dari database komponen, skema multidatabase hanya persatuan skema dari database komponen (setelah mengubah nama, untuk menghindari konflik nama) seperti contoh kasus integrasi multi database dibawah ini.



Gambar 10. Skema Multidatabase

Gambar 10. tersebut menggambarkan penggunaan skema di multidatabase, komponen-komponen dari masing-masing skema dibawa dari masing-masing DBMS dan dapat diakses melalui klien. Masing-masing schema dapat di digabungkan dan di integrasikan dalam operasional dengan cara menambahkan prefiks ke masing-masing nama tabel relasional.





Gambar 11. DML Multidatabase

4. KESIMPULAN

Integrasi instance merupakan aspek penting untuk mengintegrasikan sumberdata heterogen. Pada kasu penelitian ini pemetaan antara obyek dan sumber data yang berbeda harus didefinisikan dan dihilangkan perbedaan-perbedaan (baik struktur, tabel maupun isi) dalam representasi data. Karena itu harus dibahas definis dalam global view hasil integrasi skema. Dengan demikian konflik skema dapat dihindarkan dan dapat dilanjutkan dengan integrasi dan DML antar database (multidatabase). Dalam makalah ini kami menyajikan dua konflik struktur RL epidemiaologi (server surveilans_ RS_A dan Data_Center_SST). Setelah dilakukan representasi konflik maka bisa dilakukan pemetaan kode_kasus setelah mapping schema (Relation map_ICD_X). setelah proses tersebut maka bisa dilakukan integrasi dan import data serta dapat dilakuka integrasi DML antar database heterogen (multidatabase). Saran dalam penelitian ini adalah dapat di inventasisaasi semua kemungkinan konflik integrasi antar database dan sche surveilans sehingga dapa dibuat semacam integrasi dengan schema generator yang dinamis.

DAFTAR PUSTAKA

- [1] Amdahl Corp. Global information sharing software: Selecting a data replicator. <http://amdahl.com/doc/products/storage/gis/mm002623/data.html>, 1998.
- [2] Informix Corp. Enterprise replication: A high performance solution for distributing and sharing information. <http://www.informix.com/informix/whitepapers/entrep.pdf>, 1998.
- [3] Microsoft Corp. Accessing heterogeneous data with microsoft sql server 7.0. <http://www.microsoft.com/sql/interopmigrate/heterodata.htm>, 1998.
- [4] Microsoft Corp. Replication for microsoft sql server 7.0. <http://www.microsoft.com/sql/DeployAdmin/replication.htm>, 1998.
- [5] Sybase Corp. Data movement. <http://www.powersoft.com/products/middleware/dmove.html>, 2000.
- [6] Rick Cattell Graham Hamilton. Jdbc: A java sql api, 1997.
- [7] Eliotte Rusty Harold. XML Bible. IDG Books, 1999.
- [8] Andreas Heuer. Datenbanken: Implementierungstechniken. MITP, 1st edition, 1999.
- [9] IBM. DB2 DataJoiner for Windows NT Systems: Planning Installation an Configuration Guide. IBM Press, 1st edition, 1997.
- [10] IBM. My Mother Thinks I'm a DBA! IBM Press, 1st edition, 1999.
- [11] IBM. Replication Guide and Reference, Version 7. IBM Press, 1st edition, 1999.
- [12] Jamie Jaworski. Java 2 Platform Unleashed. SAMS, 1st edition, 1999.
- [13] B. Niswonger Laura M. Haas, R.J. Miller. Transforming heterogeneous data with database middleware: Beyond integration. Data Engineering Bulletin, 1999.
- [14] Mark Hapner Seth White. Jdbc 2.0 standard extension api, 1998.