

PENDEKATAN 'AGREEMENT' BERBASIS SEMANTIK PADA P2P UNTUK INTEROPERABILITAS INFORMASI SPASIAL

I Wayan Simri Wicaksana¹⁾, Suryo Guritno²⁾, Lily Wulandari³⁾

¹⁾ Ilmu Komputer, Universitas Gunadarma - Universite de Bourgogne
Jl. Margonda Raya no.100, Depok 16424, Indonesia
Aile de l'ingeneur, BP 47870, 21078 Dijon CEDEX, France
email : iwayan@staff.gunadarma.ac.id, iwayan@u-bourgogne.fr

²⁾ FMIPA, Universitas Gadjah Mada,
Jl. Bulak Sumur, Yogyakarta 55281, Indonesia
email : suryoguritno@ugm.ac.id

³⁾ Ilmu Komputer, Universitas Gunadarma
Jl. Margonda Raya 100, Depok 16424, Indonesia
email : lily@staff.gunadarma.ac.id

Abstrak.

Currently, information sources contain spatial information, such as address, and coordinate. In the last decades, diversity of geographic information has increased. Internet contributes to create diversity and availability of GIS.

GIS represents traditional database and geographic, as a result more complex data structure and semantic are needed. The level of diversity of GIS need a special approach to tackle the complexity.

This paper address a approach based on 'semantic agreement' which utilize common ontology and Peer to Peer (P2P) architecture as a combination approach.

P2P allows peers to group in same common interest in a community, which called clustering. The clustering develops semantic overlay network (SON). SON reduces heterogeneity of semantic.

Our architecture has three types of peers. Super peer publishes common ontology as reference, Provider peer contains sharing data which represented by export schema. Agreement is a key about understanding between export schema and common ontology. Refer to this agreement, interoperability can be delivered, such as for discovery sources, sending query or handling the answer.

This paper emphasizes in how to develop agreement and how to represent the agreement in implementation on the community.

Kata Kunci

P2P, interoperabilitas, semantik, ontologi, agreement

1. Pendahuluan dan Latar Belakang

Informasi spasial atau geografi sangatlah dibutuhkan saat ini, karena sebagian besar informasi yang tersedia mengacu kepada informasi geografi, hal yang paling sederhana adalah seperti informasi alamat.

Karena dianggap penting, banyak pihak melakukan pengembangan GIS. Sehingga pada saat ini, tersedia berbagai produk perangkat lunak, file format dan metoda. Banyaknya pihak yang terlibat, tidak dapat dihindari untuk terjadinya keragaman atau variasi. Keragaman yang kerap dibicarakan pada GIS adalah masalah representasi peta dengan raster terhadap vektor, dan konversi antar berbagai jenis file.

Dalam berbagai kegiatan seperti kebijakan pada e-Government, kerap membutuhkan penggunaan berbagai sumber informasi spasial. Kendala umum adalah kesulitan untuk mendapatkan pandangan yang umum dari bervariasinya sumber informasi spasial. Perbedaan terjadi dari level rendah seperti modeling peta, format file, unit yang digunakan, skala yang digunakan, hingga ke tingkat lebih atas seperti semantik. Pada tingkat semantik dapat terjadi sebuah obyek geografi yang sama memiliki arti (konsep / view) yang berbeda, atau sebaliknya obyek yang berbeda memiliki representasi yang sama.

Sebagai contoh, sungai sebagai obyek geografi, representasi atau view akan berbeda antara berbagai domain, seperti domain transportasi, ekologi, peternakan, pembangkit energi. Misalkan di domain transportasi akan menitik beratkan pada ukuran fisik dan kecepatan aliran sungai, kalau disisi domain peternakan akan melihat jenis hewan apa saja yang sesuai untuk dapat dibiakkan di sungai yang bersangkutan. Sehingga representasi obyek sungai akan berbeda.

Dasar yang melandasi untuk berbagi pakai informasi GIS adalah :

- Untuk mengembangkan informasi GIS dibutuhkan biaya yang besar dan memakan waktu yang lama.
- Menghindari redudansi pengumpulan data yang sama.
- Meningkatkan pemanfaatan berbagai sumber informasi spasial.

Pendekatan paling sering dilakukan dalam pemanfaatan berbagai pakai informasi spasial adalah pada tingkat pertukaran file, sehingga masalah konversi file adalah isu yang paling umum dibahas. Pada paper ini akan dilihat permasalahan di lapis atas dari konversi file yaitu pemahaman akan konsep.

Karena disadari tingginya perbedaan semantik dari berbagai domain, maka kami melakukan pendekatan dengan memanfaatkan arsitektur P2P, karena kemampuan P2P untuk mengklompokkan peer yang memiliki minat yang sama. Sehingga pembentukan sebuah komunitas pada domain atau interest yang sama dapat terwujud pada arsitektur P2P. Dilain hal P2P memiliki karakter untuk menhandel lingkungan yang dinamis, dimana peer dapat bergabung dan berpisah terhadap komunitas pada sembarang waktu, otonomi dari masing-masing peer adalah tinggi, dan terdesentralisasi.

Pada paper ini akan menguraikan tentang riset kami yang sedang berjalan, dimana pendekatan dengan 'agreement' berbasis semantik akan diuraikan pada bagian 2, dan bagian 3 akan memberikan ilustrasi representasi dari agreement. Bagian terakhir akan memberikan kesimpulan dari paper ini.

2. Agreement berbasis Semantik

Tujuan utama dari penelitian kami adalah untuk mengembangkan pendekatan untuk berbagai pakai informasi spasial dengan memanfaatkan arsitektur P2P dengan memperhatikan isu utama sebagai berikut :

- bagaimana merepresentasikan data peer pada tingkat semantik
- bagaimana memungkinkan terjadinya interoperabilitas semantik tanpa adanya *global view* dengan pendekatan 'agreement' antara peer
- bagaimana mengembangkan interoperabilitas semantik pada P2P

Untuk mencapai tujuan di atas diperlukan beberapa komponen untuk digunakan. Pada pendekatan 'agreement' akan melihat kepada model interoperabilitas semantik dan web semantik. Untuk semantik, pendekatan ini akan memanfaatkan model dari ontologi sebagai bentuk representasi pengetahuan dan referensi dari komunitas.

2.1. Ontologi

Ontologi semakin meningkat digunakan pada sistem interoperabilitas untuk menangkap makna dan hubungan antara konsep dari berbagai domain. Gruber [7] mendefinisikan sebuah ontologi adalah sebagai sebuah spesifikasi yang eksplisit dari konseptualisasi sebuah entitas dunia nyata pada sebuah domain

tertentu. Sebuah ontologi adalah sebuah vokabulari yang terdiri dari arti-arti dan hubungannya antara makna tersebut. Ontologi difokuskan pada beberapa riset terakhir adalah menangani isu pendefinisian, perancangan dan representasi.

Konteks dan metadata pada semantik dapat digunakan untuk memungkinkan terjadinya pertukaran data. Sebuah konteks merepresentasikan secara eksplisit batasan dan asumsi hubungan ke dalam definisi dan data. Berbagai model telah digunakan untuk mendefinisikan konteks, seperti 'logic rules', 'conceptual graph'. metadata dan istilah pada sebuah ontologi. Metadata adalah data yang menguraikan konten (tipe, format, semantik) dari sumber data. Beberapa meta data telah coba dikembangkan untuk data spasial. Sebagai contoh CSDGM (*content standard for digital geospatial metadata*) yang didefinisikan oleh US federal, CEN TC 287 yang didefinisikan oleh Komite Eropa untuk standarisasi, dan ISO TC 211 didefinisikan oleh *international organization*.

Semi-structure data seperti XML adalah digunakan untuk memodelkan sistem informasi yang berorientasi web sebagai fasilitas interoperabilitas. XML telah diterima secara luas sebagai standard *de facto* untuk pemodelan dan pertukaran data di lingkungan Internet. XML adalah sebuah bahasa tekstual yang menyediakan diskripsi struktur dari informasi. Sehingga ini memungkinkan untuk mendefinisikan *grammar* dengan berbasis XML Schema (XMLS). Beberapa turunan dari XML telah banyak dikembangkan seperti GML (*Geography Markup Language*), dan GPSML (*Global Positioning System Markup Language*). XML merupakan dasar sintaks dari tingkat di atasnya seperti model dari RDF dan OWL.

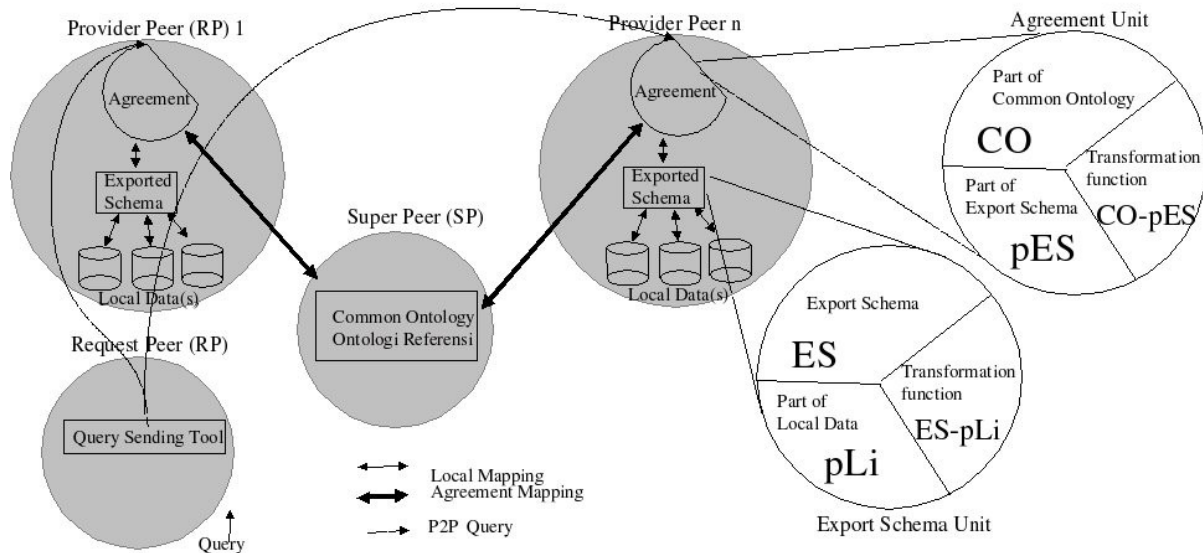
2.2. Model 'Agreement'

Arsitektur 'agreement' berbasis semantik terdiri dari beberapa lapis/*context* sebagai berikut :

- *Community context* adalah merepresentasikan konsep umum dari komunitas yang di representasikan dalam sebuah ontologi pada *Super peer* (SP) terdiri dari ontologi referensi atau *common ontology* (CO) yang menyediakan taksonomi untuk mendiskripsikan domain yang bersangkutan. Representasi CO akan menggunakan RDFS dan OWL.
- *Local context* adalah merepresentasikan lokal data di *Provider peer* (PP) ke bentuk skema ekspor (ES). Lokal data dapat terdiri dari model database relasional atau XML. Representasi skema ekspor akan menggabungkan XMLS dan RDFS.
- *Agreement context* adalah merepresentasikan persamaan pemahaman dari skema ekspor di PP terhadap *common ontology* di SP. Agreement adalah merupakan sebuah kunci utama dalam interoperabilitas semantik untuk GIS pada arsitektur P2P. Agreement akan direpresentasikan memanfaatkan tiga level dari XMLS, RDFS dan OWL.

- *Request peer* (RP) adalah peer yang mengirimkan query ke PP. Dalam penulisan query akan berbasiskan pada 'agreement' dan *common ontology*. RP tidak terletak pada salah satu *context*, tetapi dalam memanfaatkan berbagi pakai informasi akan mengacu kepada tiga *context* tersebut.

Secara ringkas komponen dari peer agreement berbasis semantik dapat dilihat pada gambar 1.



Gambar 1. Model Pendekatan Peer Agreement

Pendekatan agreement adalah berbeda dengan mapping yang secara umum dilakukan. Pada 'agreement' antara super peer dan provider peer tidak harus setuju atas semua konsep, sedangkan pada mapping yang umum adalah mencoba untuk mencari nilai mapping dari semua konsep.

Perbedaan bisa terjadi antara skema ekspor dan *common ontology* karena:

- Perbedaan class, seperti sinonim, homonim, perbedaan metode.
- Perbedaan hirarki, seperti generaliasi, perbedaan agregasi, perbedaan domain
- Perbedaan attribute, seperti perbedaan unit, nilai default, perbedaan tipe data.

Kunci utama pada proses 'agreement' adalah 'agreement unit' yang merepresentasikan 'agreement' antara ontologi referensi di SP dan skema ekspor di PP. Karena 'agreement' akan terdiri dari subset beberapa 'agreement unit'. Proses pembentukan peer 'agreement unit' terdiri dari dua tahap.

Tahap pertama adalah untuk melakukan 'match' pada elemen node (NME) antara SP dan PP. Elemen node yang dimaksud dapat berupa class, attribute atau link di export schema dan common ontology. Misalkan pada export schema ada sebuah class (elemen node) 'secondary road' sementara di common ontology class (elemen node) tersebut tidak ada. Sehingga perlu dihitung antara 'secondary street' di export schema dengan class yang ada di common ontology.

Proses *match* atau NME dilakukan berdasarkan analisis linguistik dan *tool* (WordNet). Analisis linguistik banyak digunakan pada *Natural Language Processing* seperti untuk membandingkan kata berdasarkan kesamaan bunyi, sinonim, singkatan, dan sebagainya. WordNet adalah merupakan pengembangan database bahasa yang mengikuti model taksonomi yang sudah memberikan bentuk dan pemahaman konsep. WordNet dapat melakukan perhitungan semantik antar dua kata, seperti contoh dapat dihitung hubungan semantik antara kata 'anjing' dan 'kucing' akan lebih dekat dibandingkan kata 'anjing' dan 'apel'. Tugas ini untuk manusia relatif mudah dilakukan, tetapi untuk mesin masih sulit dilakukan terutama untuk mencapai otomasi penuh.

Dengan WordNet, informasi yang didapat bisa mengetahui apakah sebuah kata lebih umum atau khusus (*generalization/specification*), sebuah kata adalah bagian (*part of*) dari kata lain. Dengan bantuan WordNet akan memungkinkan untuk membantu orang yang tidak ekspert pada domain tertentu untuk melakukan 'agreement', lebih lanjut bisa dikembangkan menuju 'agreement' ke arah semi otomatis.

Sebuah model dari NME akan terdiri dari tupel

$$\langle m_{ID}, N_{PP}^k, N_{SP}^l, R_{nc} \rangle \quad (1)$$

dimana :

m_{ID} adalah identitas NME

N_{PP}^k adalah node element k-th pada provider peer

N_{SP}^l adalah node element l-th pada super peer

R_{nc} adalah perhitungan kesamaan semantik antar node element. Perhitungan menggunakan WordNet dengan metoda perhitungan PathLength. Nilai yang dimiliki akan antara 0-1, dimana 1 adalah kedekatan tertinggi.

Pada tahap ke dua, penyesuaian berikutnya dilakukan pada elemen *class*, *properties*, *relation*, *instances*. Sebuah 'agreement unit' direpresentasikan dalam tupel:

$$\langle SMC_{ID}, ES_{PP}^i, type_{PP}^i, CO_{SP}^j, type_{SP}^j, trans \rangle \quad (1)$$

dimana :

SMC_{ID} : adalah ID unit,

ES_{PP}^i : adalah i-th ES di PP,

$type_{PP}^i$: adalah tipe elemen i-th ES,

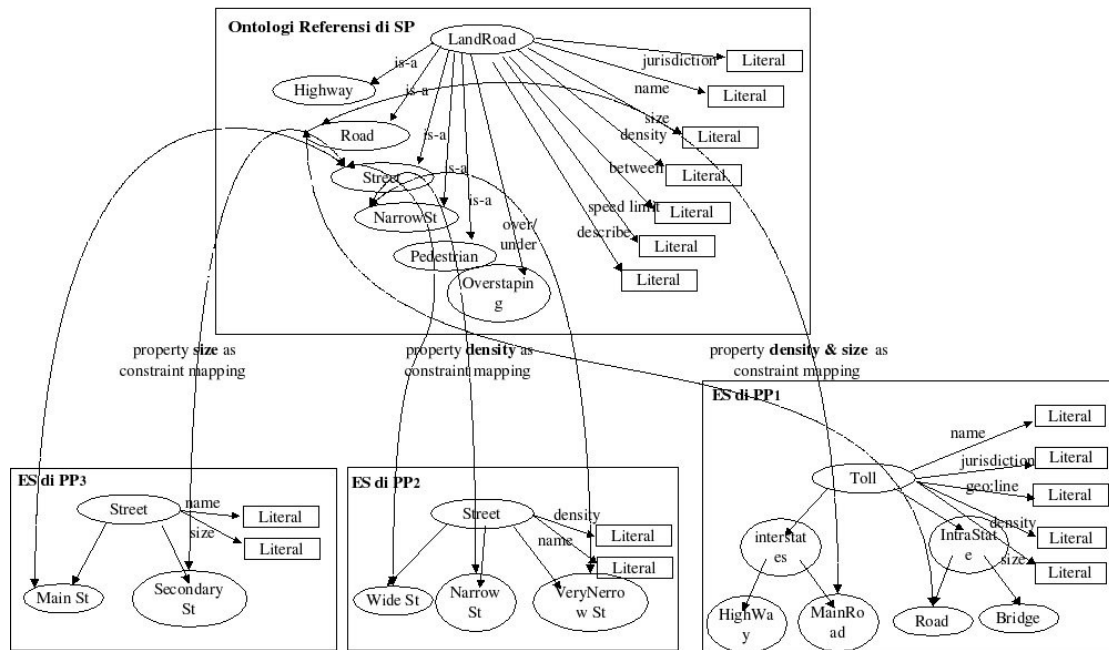
CO_{SP}^j : adalah element j-th CO,

$type_{SP}^j$: adalah tipe elemen j-th di CO, dan

$trans$: adalah fungsi transformasi antara ES dan CO, bentuk transformasi bisa pada turunan pemetaan (*class*, *relation*, *property*) dan pemetaan operasi dan batasan (*instances*).

Transformasi pada persamaan diatas dapat dilakukan dari prinsip sederhana pemetaan satu-satu, menggunakan look-up tabel hingga menggunakan persamaan fungsi yang kompleks. Secara umum transformasi dapat dikategorikan ke dalam :

- Pemetaan derivatif seperti SUPER (land transport SUPER road), SUB (road SUB land transport), and EQUAL (main street EQUAL road).



Gambar 2. Contoh Peer 'Agreement' pada Kasus Komunitas GIS Land Road

- Pemetaan operasi, ini adalah biasanya digunakan bersama dengan pemetaan derivatif dan konstraint. Contoh pemetaan operasi adalah MERGE (name EQUAL last name MERGE family name).
- Pemetaan konstrain, digunakan untuk memberikan batasan dari instances ke dalam konsep. Seperti $>$, $<$ dan sebagainya. Contoh adalah batas kecepatan di jalan biasa $<$ batas kecepatan di jalan tol.

Melihat dari katagori di atas, pada model kami akan menggunakan transformasi dengan aturan :

Pemetaan Class

- **subClass**, C1 adalah subClass C2 jika $ext(C1)$ adalah $ext(C2)$, tetapi tidak semua $ext(C2)$ adalah $ext(C1)$.
- **Equivalent**, C1 equivalent terhadap C2 jika $ext(C1)$ adalah $ext(C2)$ dan $ext(C2)$ adalah $ext(C1)$.

Pemetaan Property

- **subProperty**, P1 adalah sebuah subProperty P2 jika setiap anggota P1 adalah juga anggota P2 tetapi anggota P2 tidak harus semuanya anggota P1.
- **Equivalent**, P1 adalah equivalent terhadap P2 jika $P1(x,y)=P2(x,y)$.
- **Inverse**, P1 adalah inverse dari P2 jika $P1(x,y)=P2(y,x)$

Pemetaan Instances

- konversi unit, seperti mil ke km
- konversi tipe database, seperti tanggal ke text
- konversi datablock, seperti alamat terhadap nama jalan, kode pos dan nama kota.

3. Contoh Kasus

Gambar 2 memperlihatkan contoh sederhana dengan beberapa PP yang memiliki modelisasi obyek geografi yang berbeda. Pada PP1, PP2, PP3 *road* didefinisikan berturut-turut dengan klasifikasi berdasarkan area yang dicakup, densitas kendaraan yang lewat (*density*) dan ukuran dari jalan (*size*).

Mengacu pada gambar 2, misalkan administrator PP3 ingin bergabung dengan komunitas land transport. Maka administrator dari PP3 akan dilakukan agreement ke common ontology di super peer.

Langkah pertama akan menghitung berdasarkan NME, sebagai ilustrasi pada PP3 kita ambil node (class) *secondary street*, sementara di common ontology tidak ada . Maka kita coba menghitung dengan WordNet antara *secondary street* terhadap *road*, *street* dan *narrow street*, hasil perhitungan dapat direpresentasikan dalam bentuk :

```
<m1, Secondary Street, Road, 0.25>
<m2, Secondary Street, Street, 0.50>
<m3, Secondary Steet, Narrow Street, 0.40>
```

Dengan bantuan perhitungan NME kita bisa memilih kandidat untuk melakukan perhitungan agreement unit dari 'Secondary Street' di skema eksport terhadap 'Street' di common ontology.

Perhitungan agreement unit akan memiliki nilai :

```
<SMCA,Secondary Street, Class, Street, Class, if 3
m< size(CO) < 5m>
```

Maksud dari agreement di atas adalah, 'Secondary Street' pada class di skema eksport akan sama dengan 'Street' di class dari common ontology jika attribute di common ontology (CO) bernilai di antara 3-5 meter. Karena konsep yang digunakan oleh PP3 adalah pembagian kelas jalan berdasarkan ukuran lebar jalan. Sementara pada CO tidak tampak pembagian tipe jalan berdasarkan apa. Pada provider peer lain dapat saja pembagian berdasarkan konsep lainnya, seperti pada PP1 dan PP2.

Agreement unit di atas dapat direpresentasikan dengan RDFS dan OWL (list representasi disederhanakan) sebagai berikut:

```
<owl:Class rdf:ID="SMCA">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="&pp3;Secondary Street">
        <owl:hasValue rdf:resource="&pp3;size">
      </owl:Restriction>
    </owl:equivalentClass>
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="&co;Street">
        <owl:onProperty rdf:resource="&co;size">
          <owl:hasValue>between 3-5m</owl:hasValue>
        </owl:Restriction>
      </owl:equivalentClass>
    </owl:Class>
```

Ilustrasi pemanfaatan dari 'agreement' akan melihat sebuah contoh berikut ini. Misalkan ada sebuah query "cari *secondary street*" yang dikirim oleh sebuah RP, dimana RP memiliki persepsi akan *secondary street* yang berbasiskan dari *size*. PP3 dapat merespon query karena memiliki view yang sama, tetapi PP1(*area cakupan jalan*) dan PP2 (*traffic density*) tidak dapat menjawab secara langsung. Diasumsikan, terjadi sebuah negosiasi antara RP (*size*) dan PP2 (*traffic density*). PP2 akan menginformasikan ke RP bahwa klasifikasinya berbasiskan *density* bukan *size*, sehingga RP dapat mendefinisikan ulang konsepnya melalui 'unit agreement' $\langle SMC_{ID}, density_{PP}, attribute_{PP}, size_{SP}, attribute_{SP}, trans_{SP} = \text{konversi 'size' ke 'density'} \rangle$. Konversi dapat dilakukan melalui *look-up table*, suatu fungsi matematik, atau dengan mendefinisikan ulang persepsi/ *view* dari RP. Misalkan untuk ukuran lebar jalan 3-5 meter akan dianggap sama dengan kepadatan lalu lintas (*traffic density*) untuk 10-20 kendaraan per jam. Sehingga RP dapat menulis ulang query-nya ke PP2 menjadi "cari *secondary street* dengan *traffic density antara 10-20 kendaraan per jam*", query ini merupakan hasil negosiasi di atas. Setiap terjadi sebuah negosiasi, tidak selalu dihasilkan persamaan konsep, kalau tidak terjadi kesamaan konsep, maka query yang bersangkutan akan dibatalkan.

4. Kesimpulan

Dengan ekstensifnya penggunaan XML sebagai *markup language* untuk mendukung interoperabilitas ,

menjadi sangat menarik mengembangkan model untuk memperkaya representasi informasi di atas XML. Pendekatan 'agreement' berbasis semantik untuk P2P dapat sebagai salah satu model untuk interoperabilitas dalam mengatasi keragaman sumber informasi. Pada pendekatan ini keragaman sintaktik, struktur dan semantik akan di resolusi pada lingkungan yang tidak memiliki *global view*. Pendekatan ini dapat dimanfaatkan kepada lingkungan informasi spasial yang memiliki struktur model dan semantik lebih kompleks.

Kontribusi dari penelitian kami adalah untuk mengembangkan model pendekatan dengan 'agreement' berbasis semantik yang menggunakan ontologi khususnya untuk informasi spasial di lingkungan P2P.

Rencana kami berikutnya adalah mengembangkan prototipe dari pendekatan kami untuk membuktikan pendekatan 'agreement', serta untuk menyempurnakan negosiasi untuk memperbaiki respon terhadap query.

REFERENSI

- [1] ANDROUTSELLIS-THEOTOKIS Stephanos, "A Survey of Peer-To-Peer File Sharing Technologies", White Paper, Athens University of Economics and Business, Greece, 2002,

http://www.eltrun.aueb.gr/whitepapers/p2p_2002.pdf

- [2] BARKAI David, "An Introduction to Peer-To-Peer Computing", 2001,

<http://www.intel.com/update/departments/initech/it02012.pdf>

- [3] CRESPO Arturo, Garcia-Molina Hector, "Routing Indices For Peer-To-Peer Systems", 2001,

http://www-db.stanford.edu/~crespo/publications/crespoa_ri.pdf

- [4] LÖSER Alexander, Siberski Wolf, Wolpers, Wolfgang Nejd, "Information Integration in Schema-Based Peer-To-Peer Networks", 2003,

http://cis.cs.tu-berlin.de/~alooser/publications/Caise03_submission_Loeser_Nejdl_Wolpers_Siberski.PDF

- [5] PANTI Maurizio, Penserini Loris, Spalazzi Luca, "A Pure P2P Approach To Information Integration", 2001,

http://citeseer.ist.psu.edu/cache/papers/cs/26827/ftp://zszsirspeople.epfl.ch/zszabererzscitationszsztr_2002_02_19.pdf/a-pure-p-p.pdf

- [6] ZISMAN A, Kramer J, "Toward Interoperability in Heterogeneous Database Systems", Imperial College Research Report, London, 1995,

- [7] GRUBER T, "Toward Principles for The Desig of Ontologies used for Knowledg Sharing", Int

Journal of Human Computer Studies, 43:907-928, 1995

I Wayan Simri Wicaksana, mendapatkan gelar S.Si dalam bidang Fisika di Universitas Indonesia, Jakarta pada 1988. Pendidikan S2 dilakukan di University of Technology Swinburne, Melbourne Australia dengan mendapatkan gelar M.Eng (CIM) pada 1992. Sejak awal 2003 terdaftar sebagai mahasiswa Doktoral Ilmu Komputer di Universitas Gunadarma, Jakarta dan sejak akhir 2004 terdaftar sebagai mahasiswa Doktoral Informatik di Univesité de Bourgogne, Dijon Perancis. Selain sebagai staf tetap di Univesitas Gunadarma, juga terlibat pada berbagai proyek pemerintah dan swasta sebagai konsultan TI. Anggota dari Ikatan Profesi Komputer Indonesia (IPKIN), Kelompok Pengguna Linux Indonesia (KPLI) Jakarta dan Tim Pandu.

Suryo Guritno, mendapatkan gelar S1 dari Univesitas.....pada tahun.....Dilanjutkan program S2 dengan gelar di Univesitas diselesaikan tahun.....Jenjang PhD didapatkan dari Univesitas.....di tahun.....Gelar Profesor bidang statistik diterima pada tahun.....Tugas sebagai pengajar tetap dilakukan di Univesitas Gadjah Mada, dan menjadi dosen terbang di beberapa perguruan tinggi di Indonesia.

Lily Wulandari, memperoleh S,Kom dan M.M.SI dari Univesitas Gunadarma, Jakarta pada tahun..... dan Sejak 2004 terdaftar sebagai mahasiswa S3 Ilmu Komputer Univesitas Gunadarma. Saat ini sebagai staf pengajar di Univesitas Gunadarma.