

# Pentingnya Peranan Bahasa dalam Interoperabilitas Informasi berbasisan Komputer karena Keragaman Semantik

I Wayan Simri Wicaksana, Lintang Yuniar Banowosari, Lily Wulandari, Setia Wirawan  
Mahasiswa Doktoral Ilmu Komputer Universitas Gunadarma  
{iwayan, lintang, lily, setia}@staff.gunadarma.ac.id

## Abstrak

*Sumber informasi semakin bertambah secara dramatis pada dekade terakhir ini dikarenakan teknologi internet. Besarnya jumlah sumber informasi juga melahirkan keragaman dari sumber informasi tersebut. Keragaman timbul karena dari perbedaan domain keilmuan, negara, bahasa, dan sebagainya. Sehingga masalah untuk interoperabilitas informasi menghadapi tantangan baru khususnya dari keragaman konsep.*

*Banyak pihak yang berpikir secara terkotak bahasa adalah bagian dari ilmu sosial dan komputer adalah bagian dari ilmu eksak, ternyata paradigma seperti ini tidaklah tepat lagi. Karena untuk interoperabilitas pada era Internet dengan keragaman semantik, peran bahasa sangatlah penting untuk mengatasi keragaman semantik dalam pencarian informasi.*

*Pada paper ini akan diketengahkan bagaimana peran WordNet sebagai salah satu acuan untuk sistem interoperabilitas dalam mengukur kesamaan semantik terhadap sebuah konsep. Dari hasil pengukuran tersebut, keragaman semantik dapat diatasi untuk meningkatkan terjadinya pertukaran informasi.*

Kata kunci : keragaman semantik, similaritas, WordNet, LCH

## 1. Pendahuluan

Sumber informasi semakin bertambah secara dramatis pada dekade terakhir ini dikarenakan teknologi Internet. Besarnya jumlah sumber informasi juga melahirkan keragaman dari sumber informasi tersebut. Keragaman timbul karena perbedaan domain keilmuan, negara, bahasa, dan sebagainya. Sehingga masalah untuk interoperabilitas informasi menghadapi tantangan baru khususnya dari keragaman konsep.

Memasuki era globalisasi dan teknologi informasi, bahasa Indonesia tidak saja dilihat sebagai aset kebudayaan melainkan merupakan sarana perhubungan dan aset di bidang ekonomi, politik, dan strategi hubungan global, misalnya semakin dipelajarinya bahasa Indonesia di Jepang, Australia, Amerika, dan lain lain. Dengan demikian bahasa Indonesia telah dipelajari dan diajarkan, khususnya untuk kepentingan politik, ekonomi dan pengembangan hubungan global

Banyak pihak yang berpikir secara terkotak bahasa adalah bagian dari ilmu sosial dan komputer adalah bagian dari ilmu eksak, ternyata paradigma seperti ini tidaklah tepat lagi. Karena untuk interoperabilitas pada era Internet dengan keragaman semantik, peran bahasa sangatlah penting untuk meningkatkan kualitas dalam pencarian informasi.

Mengacu kepada [7], semantik (*semantic*) adalah cabang linguistik yang mempelajari secara khusus tentang arti, perubahan arti, dan prinsip hubungan antara kata dan artinya. Semantik mencakup pembuatan arti pada taksonomi sederhana. Semantik juga mempelajari tentang tanda dan simbol. Dari sudut pandang ilmu informasi, semantik juga memberikan efek pada komunikasi informasi dalam dan antara bahasa maupun domain. Bahkan pada bidang

komputer dikembangkan dengan istilah '*semantic web*' [8], '*semantic integration*', '*semantic interoperability*' dan ontologi.

Pada pertukaran informasi perbedaan konsep definisi dari sebuah kata tidak hanya terkait dengan bahasa, negara saja [1]. Tetapi juga terkait dengan domain informasi, contohnya kalau kita bicara kata 'penyimpanan' maka dari domain komputer akan berpikir adalah memori atau *hard-disk*, tapi kalau dari domain manufaktur bisa diartikan sebagai gudang. Hal ini jelas sangat besar perbedaan maknanya, sehingga pencarian informasi di Internet yang hanya berdasarkan arti dasar akan membawa hasil dari pertukaran atau pencarian informasi yang tidak tepat.

Tulisan ini akan dibagi menjadi beberapa bagian, pada bagian 2 yaitu tentang latar belakang pengertian semantik dari Ilmu Bahasa, dan semantik dari Ilmu Komputer serta pengantar mengenai *WordNet*. Bagian 3 akan menguraikan sebuah model pendekatan untuk mengatasi permasalahan perbedaan semantik dengan memanfaatkan *WordNet*, beberapa model perhitungan dalam menghitung similaritas beserta contoh pemanfaatannya. Bagian 4 adalah kesimpulan dan rencana ke depan dari riset ini.

## 2. Latar Belakang

Tujuan dari interoperabilitas informasi adalah memberikan sebuah pandangan kepada berbagai sumber informasi untuk dapat diakses dengan benar. Pada umumnya keragaman sumber informasi terjadi pada tingkat sintaktik, yaitu pendefinisian data; tingkat struktural adalah klasifikasi atau taksonomi data dan tingkat semantik adalah pada konsep data.

### 2.1 Semantik dari Ilmu Bahasa

Mengacu pada [9], **semantik** adalah ilmu yang mempelajari simbol-simbol dan artinya, terutama simbol dalam bahasa.

Kita mengetahui bahwa semantik itu mengkaji berbagai arti sebagai unsur intern dari struktur bahasa antara lain arti leksikal/harfiah, arti gramatikal, arti kolokasi/*makna ambiguitas* [10]. Ditinjau dari sudut arti leksikal, makna ambiguitas yaitu kemungkinan penafsiran arti lebih dari satu atau makna ganda pada satu kata, gabungan kata.

Semantik (dari bahasa Yunani *semantikos*, atau “arti yang signifikan”, diturunkan dari *sema*, tanda) adalah pembelajaran mengenai arti dari suatu *term*. Semantik kadang merupakan kebalikan dari sintak, yang mana semantik berhubungan dengan “arti” sementara sintak berhubungan dengan struktur/pola yang “diekspresikan” (sebagai contoh ditulis atau diucapkan). Semantik adalah satu bidang dari linguistik yang secara tradisional didefinisikan sebagai ilmu yang mempelajari arti dari (bagian dari) kata, frasa, kalimat dan teks.

### 2.2 Semantik dari Ilmu komputer

Kata atau istilah semantik sudah lama digunakan dalam area ilmu komputer. Tapi pada tulisan ini kata semantik yang digunakan dikaitkan dengan teknologi pengembangan *web* yaitu *semantic web*.

*Semantic web* diperkenalkan dengan penggunaan pendekatan semantik untuk memecahkan permasalahan interoperabilitas informasi. Ontologi telah menjadi alat yang menarik dan menantang untuk pendekatan semantik. Interoperabilitas semantik dicapai menurut hubungan

antar terminologi ke ontologi yang berseberangan, seperti menggunakan *sinonim*, *hyponyms* dan *hypernym*.

*Semantic Web* telah mengantar evolusi WWW ke tingkat pemanfaatan yang lebih baik. Ada dua visi dalam pengembangan *web* ke depan, yaitu membuat *web* semakin baik sebagai media kolaborasi, dan kedua *web* semakin dapat dipahami oleh mesin. Dengan memberikan anotasi data akan membuat informasi yang lebih dapat dipahami oleh mesin. Untuk mengembangkan *semantic web* [11] beberapa hal dibutuhkan seperti :

- Mengembangkan bahasa dan terminologi, dimana bahasa yang digunakan untuk mengkespresikan sesuatu yang dapat membuat mesin lebih bisa memahami dengan meta-informasi untuk dokumen.
- Mengembangkan *tool* dan arsitektur baru yang menggunakan bahasa dan terminologi tersebut untuk mengakses, mengubah dan integrasi informasi.
- Mengembangkan aplikasi yang memberikan sebuah tingkat pelayanan baru kepada pemakai dengan *semantic web*.

*Semantic Web* adalah satu proyek yang secara intensif membuat satu media yang universal untuk pertukaran informasi dengan memberikan arti (semantik), dalam satu cara yang dapat dimengerti oleh mesin, ke dalam isi dari dokumen yang ada di *web*. Saat ini di bawah arahan dari pembuat *web*, Tim Berners-Lee dari World Wide Web Consortium, *semantic web* memperluas kemampuan dari WWW melalui penggunaan standar, bahasa *markup* dan *tool* pemrosesan yang berkaitan.

## 2.3 WordNet

Pendekatan yang dilakukan untuk mengatasi permasalahan perbedaan semantik salah satunya adalah dengan memanfaatkan *WordNet* [2] yang merupakan sebuah leksikal database elektronik. *WordNet* dikembangkan untuk bahasa Inggris oleh Universitas Princeton di Amerika.

*WordNet* adalah sistem referensi leksikal *online* yang rancangannya terinspirasi oleh teori psikolinguistik dari memori leksikal manusia. Kata benda, kata kerja, kata sifat dan kata keterangan dalam bahasa Inggris diorganisir menjadi himpunan sinonim, dimana masing masing merepresentasikan satu konsep leksikal. Relasi yang berbeda menghubungkan himpunan sinonim.

Pada *WordNet* beberapa informasi dapat dicari seperti persamaan kata, lawan kata, arti kata (*glossary*), singkatan bahkan juga sampai kepada beberapa hal yang penting untuk sistem informasi seperti:

- taksonomi, 'matahari' adalah bagian (*subClass*) dari 'tata surya'
- agregasi, 'genteng' adalah bagian (*part of*) dari 'rumah'
- kemiripan, [anjing,kucing] > [anjing, pohon]

## 3. Menghitung Similaritas

Sampai saat ini kita sudah menggunakan istilah keterhubungan semantik secara bebas, dan kadang-kadang disebut juga persamaan semantik. Sebelum kita mendiskusikan berbagai

ukuran yang sudah kita pelajari secara detil, kita perlu memperjelas perbedaan antar dua terminologi ini .

Dua konsep dapat terkait tanpa menjadi serupa, maka keterkaitan harus dilihat sebagai dugaan yang lebih umum dibanding persamaan. Sebagai contoh, dua konsep mungkin terkait sebab mereka adalah lawan kata, tetapi mereka nampaknya tidak akan dipertimbangkan sebagai serupa.

Kita menggunakan istilah persamaan di dalam suatu pengertian yang sangat spesifik, yang mengacu pada suatu hubungan antar konsep yang didasarkan pada informasi ketika ditemukan dalam suatu hirarki *is-a*. Di dalam kasus *WordNet*, pertimbangan keterbatasan persamaan ini untuk antar pasangan kata benda atau pasangan kata kerja, karena hirarki konsep di dalam *WordNet* tidak mencampur jenis kata. Sebagai perihal praktis, hanya hirarki kata benda yang cukup luas untuk memungkinkan perbedaan *fine grained* antar konsep terkait.

### 3.1. Peran WordNet

Dari kemampuan *database* pada *WordNet*, banyak pengembangan lebih lanjut dilakukan untuk mengukur kesamaan semantik sebuah konsep. Beberapa pendekatan untuk analisis kata pada bidang komputer telah diketengahkan seperti pada [3,4,5,6]. Pendekatan yang dilakukan adalah dengan menghitung perpotongan antara konsep 1 dan konsep 2 terhadap arti kata, taksonomi dan vektor. Pada pendekatan yang kami gunakan adalah dengan menambahkan peran ontologi pada domain terkait.

Sebagai ilustrasi, misalkan kita mengirimkan sebuah permintaan informasi ke beberapa institusi (asumsikan semua institusi memiliki sumber informasi elektronik yang dapat diakses oleh publik) di Depok untuk mencari berapa jumlah **tenaga kerja** (*employee*). Misalkan permintaan informasi dikirimkan ke berbagai institusi seperti perusahaan swasta, kantor pemerintah, lembaga pendidikan. Untuk kantor pemerintah tenaga kerja diistilahkan dengan pegawai (*employee*), beberapa pabrik menggunakan istilah buruh (*labor*), sementara perusahaan swasta diistilahkan pekerja (*worker*), dan di universitas memakai kata dosen (*lecture*). Kalau kita hanya mengacu kepada *query* “berapa jumlah **tenaga kerja** di Depok / *how many employee at Depok*”. Maka informasi yang bisa dijawab berdasarkan pendekatan *keyword* adalah hanya untuk kantor pemerintahan, sedangkan dari institusi lainnya akan memberikan informasi dengan nilai nol. Walaupun kita tahu bahwa antara *employee*, *labor*, *worker*, *lecture* adalah hal yang sama. Dengan melakukan perhitungan kesamaan semantik kita bisa mengatasi permasalahan ini. Dimana kita selain menghitung nilai kesamaan semantik, kita juga menentukan nilai batas minimum untuk kesamaan semantik yang kita cari. Dari contoh ini jelas dengan pendekatan kesamaan semantik kita bisa mendapatkan informasi yang lebih baik.

### 3.2 Beberapa Model Perhitungan

Ada berbagai macam metode pengukuran keterkaitan/keterhubungan dan persamaan (*Measures of relatedness and Similarity*), yakni :

- *Path Based Measures* (Rada, Leacock and Chodorow/LCH, Wu and Palmer/WUP, Hirst and St.Onge)

- *Information Content Measures* (Resnik, Jiang and Conrath, Lin)
- *Gloss Based Measures*
- *Gloss Vectors*

Paper ini mengacu pada pengukuran *Path Based Measures* khususnya Leacock dan Chodorow (LCH).

### 3.2.1 Ukuran *Path Based*

Ketika diberi suatu hirarki *is-a*, satu makna menentukan tingkatan bagi dua konsep yang terkait adalah untuk menghitung banyaknya *edge*/tepi antar mereka, atau untuk menemukan panjang alur paling pendek antar dua konsep. Pada prinsipnya alur berdasarkan ukuran (*path based measures*) dapat berlaku bagi taksonomi manapun. Sehingga, di dalam evaluasi yang bersifat percobaan yang kita coba untuk mempekerjakan ukuran panjang lintasan (dan ukuran isi informasi/*information content measures*, yang akan diuraikan kemudian) baik dengan kata benda dan kata kerja. Dalam banyak kasus ini tidak berhasil dengan kata kerja, yang mana diharapkan karena hirarki katakerja di dalam *WordNet* berlimpahan dan dangkal. Sebagai hasilnya sangat sedikit konsep kata kerja yang benar-benar menduduki hirarki yang sama dan di sana akan jarang terjadi alur antar konsep kata kerja. Bagaimanapun, adalah penting untuk mencatat bahwa penggambaran ini lebih pada suatu batasan dalam *WordNet* dibanding sesuatu yang tidak bisa dipisahkan di dalam ukuran ini.

Sayangnya, panjang lintasan adalah paling sesuai ketika mereka mempunyai suatu penafsiran yang konsisten sepanjang taksonomi atau jaringan. Ini adalah bukan kasus dengan *WordNet*, karena konsep yang lebih tinggi dalam suatu hirarki lebih umum dibanding yang berada di bawahnya. Sehingga, suatu panjang jalur antara dua konsep umum dapat memberi kesan perbedaan yang besar sedangkan antara dua konsep spesifik mungkin tidak. Sebagai contoh, tikus dan binatang pengerat dipisahkan oleh suatu panjang jalur satu, yang mana jarak yang sama yang memisahkan *fire iron* / setrika api dan *implement*.

Fakta bahwa panjang lintasan dapat ditafsirkan dengan cara yang berbeda tergantung pada di mana mereka terjadi dalam *WordNet* telah mendorong pengembangan sejumlah ukuran berdasar pada panjang lintasan yang menyertakan berbagai faktor koreksi.

### 3.2.2 Leacock dan Chodorow

Model Leacock Dan Chodorow [6] dihubungkan dengan pendekatan Rada, Et.Al., akan didasarkan pada panjang alur yang paling pendek antara konsep kata benda dalam suatu hirarki *is-a*. Alur yang paling pendek adalah yang meliputi jumlah konsep intermediate/antara yang paling sedikit. Nilai ini diskala oleh kedalaman hirarki  $D$ , di mana kedalaman digambarkan sebagai panjang alur yang terpanjang dari suatu node daun/*leaf* ke node akar hirarki. Sehingga, ukuran persamaan mereka digambarkan sebagai berikut:

$$lch = \log\left(\frac{(2 * D)}{(length(c1, c2))}\right) \quad (1)$$

dimana :

c1 = konsep1

c2 = konsep2

$length(c1,c2)$  = panjang lintasan yang paling pendek (yaitu., jumlah minimum *edge* antara dua konsep)

D = Maksimum *depth* dari taksonomi (Jumlah terbesar kedalaman node antara dua konsep)

Pngukuran dengan Leacock-Chodorow mengasumsikan adanya sebuah top node yang mewakili semua node, dan akan selalu memberikan nilai lebih besar dari nol, sepanjang dua konsep yang akan dibandingkan terdapat di WordNet.

### 3.3. Contoh Penerapan

Berikut ini diilustrasikan contoh penghitungan keterkaitan/keterhubungan antar kata *teacher-employee-worker* dengan menggunakan metode LCH seperti yang dirumuskan pada persamaan (1). Langkah-langkah yang dilakukan adalah sebagai berikut :

1. Mencari panjang lintasan dari c1 dan c2 dimana c1 adalah *teacher* dan c2 adalah *employee*.
2. Masukkan c1 dan c2 ke dalam *WordNet*, dengan hasil seperti pada gambar 1 dan 2.
3. Hitung panjang lintasan mulai dari entity c1 ke c2. Dari gambar 1 dan 2 untuk contoh *teacher-employee* didapatkan *length* sebesar 7 (denga jalur : *teacher-educator-professional- adult-person-worker-employee*)
4. Sedangkan untuk mencari nilai D dapat diperoleh dengan membandingkan jumlah *depth* dari c1 dan c2, sebagai contoh untuk *teacher* akan memiliki *depth* sebesar 10 (dengan jalur : *teacher-educator-professional-adult-person-organism-living thing-object-physical entity-entity*), sedangkan untuk kata *employee* sebesar 8 (dengan jalur : *employee-worker-person-organism-living thing-object-physical entity-entity*), sehingga yang digunakan untuk D adalah nilai 10.
5. Perhitungannya akan menjadi seperti berikut ini:

$$lch = \log\left(\frac{(2 * 10)}{length(9,7)}\right)$$

$$lch = \log\left(\frac{20}{7}\right)$$
$$= 0,45$$

Ulangi langkah 1-5 untuk *teacher-worker* dan *employee-worker*. Sehingga didapatkan untuk *teacher-worker* adalah 0,52 (*length*=6, D=10) dan untuk *employee-worker* = 0,90 (*length*=2, D=8). Hasil pencarian kata di WordNet dapat di lihat pada gambar 1, 2 dan 3, serta ditampilkan dalam grafik pohon di gambar 4.

**teacher**,instructor -- (a person whose occupation is teaching)

- => educator, pedagogue, pedagog -- (someone who educates young people)
- => professional, professional person --(a person engaged in one of the learned professions)
- => adult, grownup -- (a fully developed person from maturity onward)
- => person, individual, someone, somebody, mortal, soul -- (a human being; "there was too much for one person to do")
- => organism, being -- (a living thing that has (or can develop) the ability to act or function independently)
- => living thing, animate thing -- (a living (or once living) entity)
- => object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- => causal agent, cause, causal agency -- (any entity that produces an effect or is responsible for events or results)
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Gambar 1. Output WordNet Ver 2.1 Off-line untuk kata *teacher*

**employee** -- (a worker who is hired to perform a job)

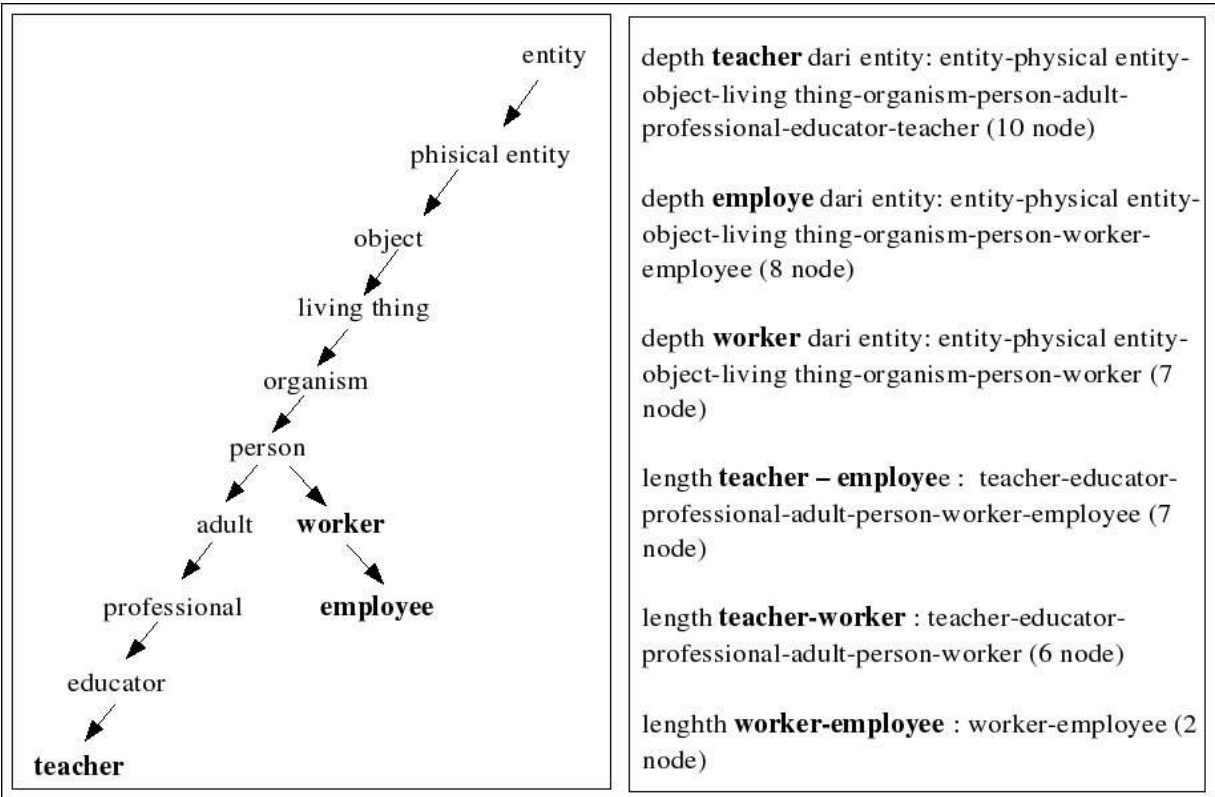
- => worker -- (a person who works at a specific occupation; "he is a good worker")
- => person, individual, someone, somebody, mortal, soul -- (a human being; "there was too much for one person to do")
- => organism, being -- (a living thing that has (or can develop) the ability to act or function independently)
- => living thing, animate thing -- (a living (or once living) entity)
- => object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- => causal agent, cause, causal agency -- (any entity that produces an effect or is responsible for events or results)
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Gambar 2. Output WordNet Ver 2.1 Off-line untuk kata *employee*

**worker**-- (a person who works at a specific occupation; "he is a good worker")

- => person, individual, someone, somebody, mortal, soul -- (a human being; "there was too much for one person to do")
- => organism, being -- (a living thing that has (or can develop) the ability to act or function independently)
- => living thing, animate thing -- (a living (or once living) entity)
- => object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- => causal agent, cause, causal agency -- (any entity that produces an effect or is responsible for events or results)
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Gambar 3. Output WordNet Ver 2.1 Off-line untuk kata *worker*



Gambar 4. Tampilan grafik pohon dari gambar 1-3

Dalam implementasi kita dapat menentukan nilai batasan kesamaan semantik, misalkan kita kirim query "how many number of employee at Depok city?". Pada sistem informasi kita tentukan nilai batasannya, kalau kita *setting* misalkan a nilainya harus di atas 0,50 ; maka *teacher* tidak dihitung sebagai *employee*, tapi kalau kita hitung dengan batasan 0,40, maka *teacher* akan



dihitung sebagai *employee*. Dengan memanfaatkan semantik dari linguistik analysis kita bisa melakukan interoperabilitas information lebih baik dibandingkan dengan pendekatan *keyword* semata.

#### 4. Kesimpulan dan Rencana Ke Depan

WordNet dapat digunakan sebagai salah satu acuan untuk sistem interoperabilitas dalam mengukur kesamaan semantik sebuah konsep. Hasil pengukuran kesamaan semantik tersebut dapat kita gunakan untuk mengatasi keragaman semantik sehingga pertukaran informasi dapat terjadi lebih baik.

Semakin besar nilai hasil perhitungan menunjukkan semakin besar tingkat kesamaan semantik sebuah konsep dan sebaliknya semakin kecil nilai hasil pengukuran menunjukkan semakin kecil tingkat kesamaan semantik sebuah konsep. Sehingga dengan cara menaikkan atau menurunkan nilai batas sebuah pencarian, kita dapat memperoleh informasi yang paling sesuai.

Penelitian ini menggunakan WordNet berdasarkan pada bahasa Inggris, sehingga segala sesuatunya mengacu pada aturan dan ketentuan yang berlaku pada tata bahasa Inggris. Seperti kita ketahui aturan dan keketentuan bahasa satu dengan bahasa lainnya berbeda sehingga dalam mencari makna atau kesamaan sebuah konsep harus mengacu sebaiknya mengacu pada bahasa yang bersangkutan. Oleh karena itu perlu dikembangkan sebuah leksikal database elektronik yang mengacu pada bahasa Indonesia untuk memecahkan masalah keragaman semantik yang ada pada bahasa Indonesia yang banyak menyerap kata asing dan daerah yang cukup beragam di tanah air tercinta ini.

#### 5. Daftar Pustaka

- [1] Aris Budi Wuryanto, "Pendekatan Silang Budaya sebagai Pencitraan Budaya Indonesia Melalui Pengajaran BIPA", 2005, akses Mei 2005, <http://www.ialf.edu/kipbipa/papers/ArisBudiWurianto.doc>
- [2] WordNet homepage, akses Januari 2005, <http://WordNet.princeton.edu>
- [3] Michelizzi, Jason., "Similarity and other current activities", 2005, akses Juni 2005 <http://www.d.umn.edu/~tpederse/Group04/jm-slides-sep-9.pdf>
- [4] "The use of WordNet in text analysis", 13 Feb 2002, <http://tcc.itc.it/people/bentivogli/corpora/WN-text-analysis.html>
- [5] Burgun. Anita, Bodenreider. Olivier, "Comparing term, concepts and semantic classes in WordNet and the Unified Medical Language System, in Proc. NAACL'2001 Workshop, 2001, pp.77-82. <http://lhncbc.nlm.nih.gov/lch/docs/published/2001/pub2001027.pdf>
- [6] Tsang. Vivian, Stevenson. Suzanne, "Calculating Semantic Distance between Word Sense Probability Distributions", 2005, akses April 2005, <http://www.cnts.ua.ac.be/conll2004/pdf/08188tsa.pdf>
- [7] akses Juni 2005, [http://lazdynas.homeip.net:9673/IST4BALT/Members/ZigmasB/Ontology\\_taxonomy/semantik\\_web\\_glossary](http://lazdynas.homeip.net:9673/IST4BALT/Members/ZigmasB/Ontology_taxonomy/semantik_web_glossary)
- [8] akses Februari 2005, <http://www.w3.org/2001/sw/>
- [9] akses Maret 2005, <http://id.wikipedia.org/wiki/Semantik>.

- [10] akses Juni 2005, <http://www.pikiran-rakyat.com/cetak/0704/25/khazanah/wisatabahasa.htm>
- [11] Berners-Lee, T., *Weaving the Web, The Original Design and ultimate Destinity of the World Wide, Web*, Harper, 1999
- [12] akses Januari 2005, [http://en.wikipedia.org/wiki/Semantic\\_web](http://en.wikipedia.org/wiki/Semantic_web)
- [13] Pedersen, Ted and Banerjee, Satanjeev, and Patwardhan, Siddharth "*Maximizing Semantic Relatedness to Perform Word Sense Disambiguation*", University of Minnesota Supercomputing Institute Research Report UMSI 2005/25, March 2005, akses Mei 2005, <http://www.msi.umn.edu/general/Reports/rptfiles/2005-25.pdf>,