

Membangun Aplikasi Natural Language Processing Menggunakan Instant Messenger Untuk Informasi Bencana

Herny Februariyanti,¹⁾ Eri Zuliarso²⁾

Abstract— *Instant messaging is a communication facility to the users of internet chat. By using this facility, the user can communicate by sending a text message to another user. In addition, Instant Messaging also serves to exchange files in a peer to peer. Availability of vast resources and varied, as well as the rapid development of techniques Natural Language Processing (NLP), Information Extraction (IE) and Information Retrieval (IR) greatly influenced the development of QA (Question Answering) system, which used to only be able to answering the questions are limited to a field (domain) specific to based on structured information sources such as databases, up to now can answer different types of questions, referring to a collection of information from unstructured text. This research is utilizing receipts Instant Messenger Google Talk service to provide information to your questions on the online disaster. The system works by utilizing a database of various sites on the Internet. The research is fitted with a QA application system using disaster information crawler results of some catastrophic news site.*

Keywords— disaster information, instant messaging, Question Answering system

1. PENDAHULUAN

Instant Messaging (IM) saat ini mengalami perkembangan yang cukup pesat pada jaringan *user*, karena kemampuannya mengirimkan pesan secara singkat dan cepat antara pengguna telekomunikasi. IM menjadi perangkat yang sangat penting untuk industri di seluruh dunia. IM digunakan di dalam penjadwalan (*scheduling meeting*), pertukaran informasi bisnis dan informasi *client* dan lain-lain. IM telah dikembangkan pada sektor- sektor *private* atau antar *provider* seperti *American Online Instant Messenger* (AIM), MSN dan Yahoo. Pada tahun 1998 muncul protokol IM yang bersifat *open source* yang terkenal dengan protokol Jabber.

Mengantisipasi bencana dalam waktu singkat dapat dilakukan dengan menerapkan sistem peringatan dini. Sistem itu bekerja dengan memanfaatkan basis data dari berbagai situs di Internet. eksistensi TI tidak untuk menghalau suatu bencana alam yang datang secara tiba-tiba melainkan untuk menyampaikan informasi sebelum dan sesudah bencana alam itu terjadi. Teknologi

informasi tersebut tentu tidak berdiri sendiri melainkan terintegrasi dengan berbagai perangkat teknologi canggih lainnya yang dapat memberikan peringatan dini secara sistematis kepada warga yang berdomisili di sekitar kawasan rawan bencana alam. Sistem ini yang kemudian dikenal dengan nama sistem peringatan dini.

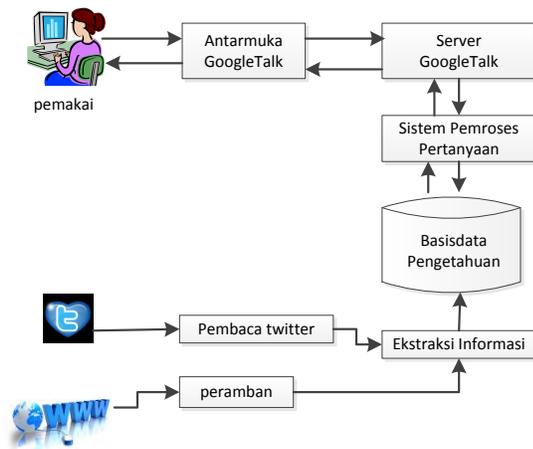
Dalam penelitian ini digunakan Instant Messenger dengan memanfaatkan layanan Google Talk. Untuk dapat memanfaatkan fasilitas ini pemakai harus mempunyai account email gmail (xxxx@gmail.com). Dengan memanfaatkan layanan Google Talk diharapkan Sistem dapat menyebarkan informasi kepada pengguna. Sistem dapat menjawab kebutuhan *request* informasi bencana dari pengguna.

2. ARSITEKTUR SISTEM QUESTION ANSWERING

QA system merupakan sebuah sistem yang mengijinkan user menyatakan kebutuhan informasinya dalam bentuk yang lebih spesifik dan alami, yaitu dalam bentuk *natural language question*, dan tidak mengembalikan daftar dokumen yang harus disaring oleh user untuk menentukan apakah dokumen-dokumen tersebut mengandung jawaban atas pertanyaan, tetapi mengembalikan kutipan teks singkat atau bahkan frase sebagai jawaban [1].

QA system yang dikembangkan dengan tujuan, sumber informasi, dan teknik yang berbeda dapat memiliki arsitektur yang berbeda pula. Gambar 1 menunjukkan arsitektur umum dari QA system yang tersusun atas lima komponen, yaitu: *Piranti antar muka pemakai adalah Google Talk, Basisdata Pengetahuan adalah kumpulan masukan dari system QA, Pemroses Pertanyaan adalah aktor utama Sistem QA yang memproses masukan dari pemakai, mencari jawaban, membuat keluaran dan mencatat ke log transaksi, peramban dan pembaca Twitter digunakan untuk mendapatkan informasi baru tentang bencana dari situs berita online dan Twitter. Ekstraksi informasi digunakan untuk memperoleh informasi atau data yang penting dari dokumen teks bahasa alami yang tidak terstruktur pada domain tertentu.*

Herny Februariyanti, Sistem Informasi, Fakultas Teknologi Informasi, Universitas Stikubank (UNISBANK) Semarang, (hernyfeb@gmail.com)



Gambar 1 Arsitektur Sistem QA

2.1. Instant Messaging

Instant messaging merupakan fasilitas komunikasi *chatting* untuk para pengguna internet. Dengan menggunakan fasilitas ini, *user* dapat berkomunikasi dengan cara mengirimkan pesan berupa *text* dengan *user* lain. Selain itu, *Instant Messaging* juga berfungsi untuk tukar menukar file secara *peer to peer*. Saat ini *Instant Messaging* yang terkenal adalah *Yahoo Messenger* dan *MSN Messenger*. Namun kita dapat membuat *Instant Messenger* sendiri menggunakan protokol *Jabber* serta aplikasi yang *Open Source* dan serta memodifikasi *client* sesuai dengan keinginan kita. Bahkan saat ini, *Google* mengembangkan *Instant Messengernya*, *Gtalk*, yang juga berbasis protokol *XMPP* (*Jabber*).

2.2. Protokol Jabber

Jabber adalah sebuah protokol *XML* yang terbuka untuk pertukaran *message* dan *presence* yang *real-time* antara dua *user* di dalam jaringan *Jabber*. Banyak kegunaan teknologi *Jabber*, pada awalnya teknologi *Jabber* bersifat *asynchronous*, *platform IM* yang dapat digunakan secara luas dan jaringan *IM* berdasarkan fungsinya hampir sama dengan sistem *IM* yang resmi seperti *AOL Instant Messaging* (*AIM*) dan *Yahoo Instant Messaging* [2].

Sebagai usaha menjadikan *Jabber* sebagai protokol standar *Instant Messaging*, pada Juni 2000 komunitas *Jabber* telah mempublikasikan protokol tersebut sebagai *Request for Comment* (*RFC*) pada *Internet Engineering Task Force* (*IETF*) sebagai bagian dari standar *Instant Messaging and Presence Protocol* (*IMPP*), tetapi *IMPP* ini tidak berjalan sukses. Pada bulan Mei 2001, *Jabber Community* dan *Jabber Inc.* membuat *Jabber Software Foundation* untuk menyediakan asisten organisasi secara langsung (*direct organizational assistance*) dan asisten teknis secara tidak langsung terhadap komunitas *Jabber*. Pada tahun 2002, *Internet Engineering Steering Group* (*IESG*) menyetujui formasi

Extensible Messaging and Presence Protocol Working Group (*XMPP*) dengan *Internet Engineering Task Force* (*IETF*).

Ruang lingkup *working group* adalah untuk mengeksplorasi dan dimana protokol tersebut digunakan, memodifikasi protokol yang sudah ada agar dapat memenuhi *RFC 2799* seperti persyaratan yang ditentukan dalam spesifikasi *Common Presence and Instant Messaging* (*CPIM*). Fokus utama *working group* adalah membuat *XML stream* termasuk *stream* pada level *security* dan autentikasi, elemen data dan *namespace* yang dibutuhkan untuk mencapai dasar *IM* dan *Presence* [3].

XMPP working group menerbitkan *XMPP Core Internet-Draft* sebagai dokumen yang menggambarkan fitur-fitur utama *Extensible Messaging* dan protokol *Presence*. Makalah *XMPP* ini memuat protokol *Jabber* yang bekerja pada sistem keamanan *client-server* dan *server-server*.

2.3. Crawler

Web Crawler, juga sering dikenal sebagai *Web Spider* atau *Web Robot* adalah salah satu komponen penting dalam sebuah mesin pencari modern. Fungsi utama *Web Crawler* adalah untuk melakukan penjelajahan dan pengambilan halaman-halaman *Web* yang ada di *Internet*. Hasil pengumpulan situs *Web* selanjutnya akan diindeks oleh mesin pencari sehingga mempermudah pencarian informasi di *Internet*.

Mendesain sebuah *crawler* yang baik saat ini menemui banyak tantangan. Secara eksternal, *crawler* harus mengatasi besarnya situs *Web* dan link jaringan. Secara internal *crawler* harus mengatasi besarnya volume data. Sehubungan dengan terbatasnya sumber daya komputasi dan keterbatasan waktu, maka harus hati-hati memutuskan *URL* apa yang harus di scan dan bagaimana urutannya.

Crawler dapat dengan mudah mengambil ribuan berita dari situs berita online. Namun volume informasi terlalu besar untuk menangkap semua informasi yang dibutuhkan. Untuk itu digunakan teknik pencarian informasi seperti *Topik Deteksi* dan *Pelacakan untk* dapat membantu mengatur beritayang berkaitan dengan bencana alam [4].

2.4. Twitter

Twitter adalah sebuah situs web yang dimiliki dan dioperasikan oleh *Twitter Inc.*, yang menawarkan jejaring sosial berupa mikroblog sehingga memungkinkan penggunaanya untuk mengirim dan membaca pesan yang disebut *kicauan* (*tweets*). *Kicauan* adalah teks tulisan hingga 140 karakter yang ditampilkan pada halaman profil pengguna. *Kicauan* bisa dilihat

secara luar, namun pengirim dapat membatasi pengiriman pesan ke daftar teman-teman mereka saja.

Karakteristik penting dari Twitter adalah sifat real-time. Sebagai contoh, ketika gempa bumi terjadi, banyak orang membuat posting Twitter (tweet) terkait gempa, yang memungkinkan deteksi gempa kejadian segera, hanya dengan mengamati tweet [5].

2.5. Pengolahan Bahasa Alami

Shan Wang, 2000 [6] mengembangkan system query berbahasa alami Cina untuk mengakses database yang diberi nama Nchiql. Nchiql didisain sebagai system query ke system basisdata Cina dan dapat digunakan ke aplikasi yang berbeda pada DBMS yang berbeda. Ada tiga sub-sistem dalam Nchiql: *intelligent user interface*, *domain-knowledge extraction system* dan *query statement processing system*. Pemakai memasukkan kalimat query dalam bahasa alami melalui antarmuka yang cerdas. Menggunakan teknologi pemrosesan bahasa alami, Nchiql menterjemahkan kalimat query ke kalimat SQL sesuai dengan semantic basisdata yang diekstrak oleh system ekstrak dan disimpan dalam system kamus. Setelah eksekusi kalimat SQL oleh DBMS berhasil, hasil query dikirimkan ke antarmuka pemakai dan ditampilkan dalam cara yang ramah untuk pemakai (*user friendly*).

Andayani, 2002 [7], telah melakukan penelitian Query Bahasa Indonesia untuk basis data akademik. Database yang digunakan adalah relasional. Aplikasi pemrosesan bahasa alami untuk query basis data ini menggunakan input dalam bahasa Indonesia. Input berkisar pada pertanyaan-pertanyaan untuk mengakses informasi atau data dari basisdata. Dengan demikian, meskipun bahasa Indonesia sudah mempunyai grammar dan aturan produksi, akan tetapi dalam aplikasi ini perlu ditentukan aturan produksi yang secara khusus menangani pola-pola pertanyaan pada input.

Yun Yao Li, 2005 [8], mengembangkan sebuah antarmuka bahasa alami untuk mengquery database XML. NaLIX adalah sebuah antarmuka query bahasa alami interaktif yang berlaku umum ke basisdata XML. Sistem NaLIX dapat menerima kalimat bahasa Inggris sebarang sebagai input query, yang mencakup agregasi (*aggregation*), bersarang (*nesting*), dan harga gabungan (*join*), among other things. Query ini diterjemahkan, menjadi ekspresi XQuery sehingga dapat dievaluasi ke basisdata XML. Penterjemahan dilakukan melalui pemetaan secara gramatikal dimana bahasa alami diuraikan menjadi token kemudian dilakukan pendekatan korespondensi dengan elemen-elemen dalam XQuery sehingga didapat hasil XML.

2.6. Basisdata XML

Bourret, 2005[9] Basisdata XML adalah sistem perangkat lunak yang digunakan untuk menyimpan data yang membolehkan data untuk diimpor, diakses dan diekspor dalam format XML. Basisdata XML mempunyai keunggulan lebih baik dibandingkan dengan system basisdata relasional jika data yang akan disimpan berupa dokumen. Dengan basisdata XML juga memungkinkan untuk melakukan penelusuran isi dokumen.

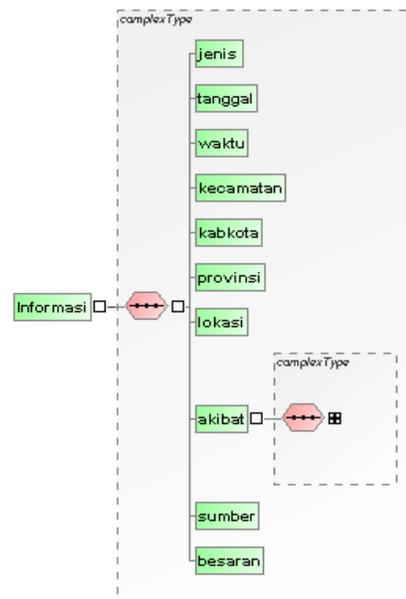
Ada dua kelas utama basisdata XML yang ada :

- XML-enabled. Kelas ini memetakan semua XML ke basisdata tradisional (seperti basisdata relasional), menerima XML sebagai input dan membuat XML sebagai keluaran.
- Native XML (NXD). Model internal dari basisdata tergantung pada XML dan menggunakan dokumen XML sebagai dasar unit penyimpanan.

3. HASIL DAN PEMBAHASAN

3.1. Struktur Data

Untuk dapat menyimpan data informasi bencana dalam format XML, maka terlebih dahulu dibuat XML Schema. XML Schema dibuat untuk menjaga konsistensi masukan data. XML Schema informasi bencana untuk data yang disimpan dapat dilihat pada gambar 2 sebagai berikut :



Gambar 2. XML Schema Informasi Bencana

Contoh data dalam format XML :

```
<?xml version="1.0" encoding="UTF-8"?>
<Informasi>
<jenis>gempa bumi</jenis>
```

<tanggal>22 Agustus 2013</tanggal>
 <waktu>2.45 WITA</waktu>
 <kabkota>Denpasar</kabkota>
 <provinsi>Bali</provinsi>
 <sumber>BMKG</sumber>
 <besaran>5,4 SR</besaran>
 </Informasi>

3.2. Sistem Pemrosesan Pertanyaan

Sistem pemroses pertanyaan terdiri dari beberapa modul program :

1. GoogleTalk.java

Merupakan program utama, untuk mengelola pembacaan pesan. Program menggunakan thread yang akan memantau apakah ada pesan yang masuk. Apabila ada pesan yang masuk selanjutnya akan diproses di ThreadPesan.java.

2. JabberSmackAPI.java

Berisi modul untuk berkomunikasi dengan GoogleTalk menggunakan pustaka JabberSmackApi.

3. ThreadPesan.java

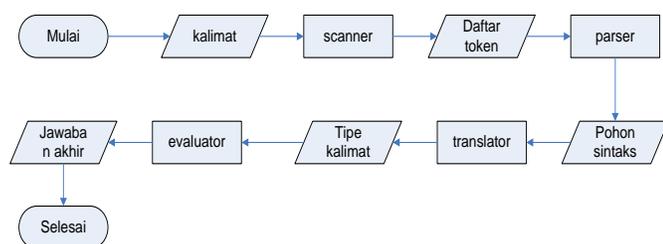
Digunakan untuk mengelola pesan. Pengelolaan dilakukan dengan melakukan parsing, penterjemahan menjadi perintah Xquery, mengeksekusi perintah XQuery dan mengelola hasil query.

4. SendMessage.Java

Digunakan untuk mengirim pesan jawaban dari hasil query.

3.3. Struktur Pengolah Bahasa Alami

Komponen pengolah bahasa alami untuk Aplikasi Pengolah Bahasa Alami Untuk Query Basisdata XML ditunjukkan pada Gambar 3.



Gambar 3. Komponen pengolah bahasa alami untuk Aplikasi Pengolah Bahasa Alami Untuk Query Basisdata XML

Setiap kalimat bahasa alami, berupa kalimat berbahasa Indonesia, yang dimasukkan akan melewati proses yang dilakukan oleh scanner, parser, translator dan evaluator sebelum mendapatkan jawaban akhir. Scanner akan melakukan pemeriksaan bentuk kalimat dan mengelompokkannya menjadi daftar token yang kemudian diteruskan ke proses berikutnya yang

dilakukan oleh parser. Dalam proses ini parser melakukan pelacakan terhadap token-token tersebut untuk dibandingkan dengan daftar token yang telah ditetapkan. Translator akan menterjemahkan hasil parsing untuk mengecek kesesuaian struktur kalimat dengan pola atau aturan produksi yang. Selanjutnya, hasil proses yang sesuai dengan pola kalimat ini akan diteruskan ke evaluator.

3.4. Penyusunan Aturan Produksi

Aplikasi pemrosesan bahasa alami untuk query basis data ini menggunakan input dalam bahasa Indonesia. Input berkisar pada pertanyaan-pertanyaan untuk mengakses informasi atau data dari basisdata. Dengan demikian, meskipun bahasa Indonesia sudah mempunyai grammar dan aturan produksi, akan tetapi dalam aplikasi ini perlu ditentukan aturan produksi yang secara khusus menangani pola-pola pertanyaan pada input. Berdasarkan pola keteraturan pertanyaan-pertanyaan, maka aturan produksi dengan symbol awal <query> ditentukan seperti dalam gambar :

```

<query> → <frase atribut><frase kondisi>
<frase atribut> → <atribut> <ekor atribut>**
<ekor atribut> → <kata sambung> <frase atribut>
| kosong
<frase kondisi> → kosong <frase 1> | <atribut> <ekor
kondisi>**<frase 2>
<ekor kondisi> → <data><frase 3> | <data> <kata
sambung><frase kondisi>** <frase 4>
| <atributkondisi> <data><frase 5>
| <atributkondisi> <data><frase
kondisi>**<frase 6>
<atributkondisi> → 'bukan'|'tidak'|'sebelum'|'sesudah'
<kata sambung> → ', '| dan
<atribut> → jenis|tanggal|waktu|kabkota|
Provinsi|lokasi|korban|meninggal|luka-
luka|mengungsi
<data> → tergantung dari data yang ada pada query
  
```

Berdasarkan contoh-contoh pertanyaan input dalam aplikasi ini, ada 7 tipe query yang diidentifikasi. Setiap tipe mempunyai notasi dan pola input yang berbeda. Notasi tipe query akan digunakan lebih lanjut dalam proses implementasi. Pola input menentukan penggolongan tipe query berdasarkan input yang sesuai dengan pola tersebut.

Simbol-simbol yang digunakan dalam penulisan pola input adalah sebagai berikut :

[T/P] : adalah kata Tanya atau kata perintah yang mengawali input, yang dalam proses selanjutnya dapat diabaikan.

Kata tersebut diapit tanda [] yang berarti bersifat opsi.

- [Plk] : satu atau lebih kata-kata pelengkap.
- <atribut> : nama atribut yang terdapat dalam daftar token, atau kata-kata yang merupakan sinonim atribut.
- <bukan> : kata-kata yang mempunyai arti sama dengan “bukan”
- <opr> : symbol atau kata-kata yang berfungsi sebagai operator
- <int> : kata-kata yang menjelaskan tentang intensitas, yang dibedakan menjadi maks dan min.

Ketujuh tipe query yang diidentifikasi adalah sbb:

- a. Tipe q_a (query-atribut)
Tipe query yang hanya berisi satu atribut yang akan ditampilkan. Query ini merupakan tipe yang paling sederhana, yang hanya memuat atribut yang ditanyakan.
Tipe ini mempunyai pola input, yaitu :
 - i. [T/P] <atribut> [Plk]
Contoh : Apa bencana yang terjadi
 - ii. [T/P] [Plk] <atribut>
Contoh : Dimana terjadi gempa bumi
- b. Tipe q_aa (query-atribut-atribut)
Tipe query yang berisi beberapa atribut yang akan ditampilkan. Tipe query ini memuat beberapa atribut yang ditanyakan. Untuk memisahkan satu atribut dengan atribut berikutnya digunakan kata sambung ‘dan’ atau tanda baca koma ‘,’.
Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> (<ktsambung> <atribut>)*
Contoh : Tampilkan bencana apa, tanggal dan provinsi
- c. Tipe q_a_opr (query-atribut-atribut-operator)
Tipe query berisi satu atribut yang akan ditampilkan dan satu kondisi.
Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> [Plk] <atribut> <data>
Contoh : Berapa korban gempa bumi Nabire
- d. Tipe q_a_opr (query- atribut-atribut-atribut-operator- atribut-operator)
Tipe query berisi beberapa atribut yang akan ditampilkan dan beberapa kondisi.
Tipe query ini memuat beberapa atribut yang ditanyakan. Untuk memisahkan satu atribut dengan atribut berikutnya digunakan kata sambung ‘dan’ atau tanda baca koma ‘,’. Demikian juga untuk memisahkan satu atribut kondisi dengan atribut kondisi berikutnya

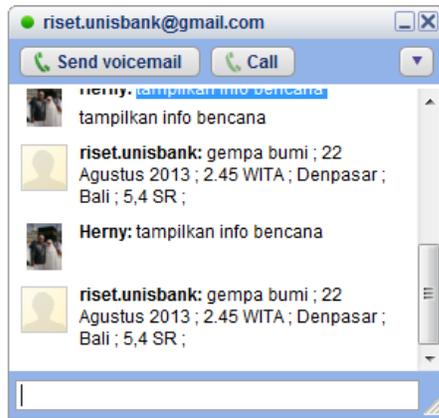
digunakan kata sambung ‘dan’ atau tanda baca koma ‘,’.

Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> (<ktsambung> <atribut>)* [Plk]
<atribut> <data> <ktsambung> <atribut><data>

Contoh :
Berapa korban yang meninggal dan korban yang luka saat tanah longsor di Kabupaten Aceh Tengah tanggal 2 Juli 2013

- e. Tipe q_operator (query-atribut-operator-<data>)
Tipe query berisi beberapa atribut yang akan ditampilkan dan kondisi operator ‘sebelum’ atau ‘sesudah’.
Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> (<ktsambung> <atribut>)* [Plk]
<atribut><op_tahun><data>
Contoh :
Tampilkan gempa yang terjadi di Provinsi Papua sebelum tahun 2010
- f. Tipe q_bukan (query-atribut-bukan-data)
Tipe query berisi beberapa atribut yang akan ditampilkan dan kondisi operator “bukan” atau “tidak” atau “selain”.
Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> (<ktsambung> <atribut>)* [Plk]
<atribut><op_bukan><data>
Contoh :
Tampilkan gempa yang terjadi bukan di Provinsi Papua
- g. Tipe q_a_bukan (query-atribut-atribut-bukan-data)
Tipe query berisi beberapa atribut yang akan ditampilkan dan kondisi operator “bukan” atau “tidak” atau “selain”.
Tipe ini mempunyai pola input, yaitu :
[T/P] <atribut> (<ktsambung> <atribut>)* [Plk]
<atribut><op_bukan><data>
Contoh :
Tampilkan gempa dan tanggal kejadian yang terjadi bukan di Provinsi Papua

Gambar 4.a. dan 4.b menampilkan contoh tampilan layar Google Talk saat berinteraksi dengan Sistem. akan menjawab secara otomatis sesuai dengan pertanyaan pengguna.



Gambar 4.a. Contoh 1 Tampilan antarmuka dan response dari Sistem QA



Gambar 4.b. Contoh 2 Tampilan antarmuka dan response dari Sistem QA

4. KESIMPULAN

Kontribusi utama dari penelitian ini adalah mengadaptasi teknik menjawab FAQ ke sistem yang menjawab pertanyaan menggunakan data yang terstruktur, yakni menggunakan basisdata XML.

Sistem yang dibuat dapat melakukan query pada basisdata XML dengan menggunakan kalimat bahasa Indonesia sederhana.

Pemanfaatan Google Talk sebagai antarmuka memberi alternative untuk mendapatkan akses informasi secara real time.

5. SARAN

Dalam penelitian ini setiap pertanyaan akan langsung dijawab secara real time. Dalam penelitian berikutnya akan dibuat sistem instant mesenger menggunakan model dialog. Dengan model ini jawaban atas pertanyaan akan mempertimbangkan pertanyaan sebelumnya.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Direktorat Penelitian dan Pengabdian Masyarakat Dirjen DIKTI yang telah memberikan dukungan financial terhadap penelitian ini. Penelitian ini didanai HIBAH Bersaing Tahun Anggaran 2013.

DAFTAR PUSTAKA

- [1] Sneiders, E. "Automated Question Answering: Template-Based Approach. Ph.D. thesis." , Department of Computer and Systems Sciences, Stockholm University and the Royal Institute of Technology, Sweden ,2002.
- [2] Shigeoka, Iain, "Instant Messaging in Java The Jabber Protocols", Manning Publications Co.,2002.
- [3] Cover, Robin, "IETF Charters Extensible Messaging and Presence Protocol(XMPP) Working Group.," WWW page, 2002 <http://xml.coverpages.org/>.
- [4] Yang, C, Wei, C., "Tracing the Event Evolution of Terror Attacks from On-line News", Lecture Notes in Computer Science Volume 3975, 2006
- [5] Sakaki,T., Okazaki, M., "Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors", WWW2010, April 26-30, 2010.
- [6] Wang, Shan 2000, "Nchiql: A Chinese Natural Language Query System to Databases, Proceedings of the 1999 International Symposium on Database Applications in Non-Traditional Environments,IEEE.
- [7] Andayani, Sri, 2002, Qubin : Query Bahasa Indonesia untuk basis data akademik, Tesis magister Komputer, Program Pasca sarjana Ilmu Komputer, UGM
- [8] Yunyao Li, 2005, NaLIX: an Interactive Natural Language Interface for Querying XML, *SIGMOD*, Baltimore, Maryland, USA
- [9] Bourret, Ronald ,<http://www.rpbouret.com>
- [10] Loucopoulos, P., Zicari, R. (eds.), "Conceptual modeling, databases, and CASE: an integrated view of information systems development.", John Wiley & Sons , 1992
- [11] Tsopoulos, I, 1995, Natural Language Interface, *Journal of Natural languages Engineering*, Cambridge University Press.
- [12] World Wide Web Consortium, "Extensible Markup Language (XML) 1.0 (Second Edition)", W3C Recommendation, October 2000.
- [13] World Wide Web Consortium, "XQuery: A Query Language for XML", W3C Working Draft, February 2000.