

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/316644260>

# Semantic Search Engine dan Penjawab Otomatis Berbasis Twitter untuk Permintaan Data Publikasi BPS

Conference Paper · October 2015

---

CITATIONS

0

---

READS

22

2 authors, including:



[Takdir Takdir](#)

Sekolah Tinggi Ilmu Statistik

6 PUBLICATIONS 4 CITATIONS

SEE PROFILE

# *Semantic Search Engine* dan Penjawab Otomatis Berbasis *Twitter* untuk Permintaan Data Publikasi BPS

Fakhriyanto

Jurusan Komputasi Statistik  
Sekolah Tinggi Ilmu Statistik  
Jakarta, Indonesia  
11.6654@stis.ac.id

Takdir

Jurusan Komputasi Statistik  
Sekolah Tinggi Ilmu Statistik  
Jakarta, Indonesia  
takdir@stis.ac.id

**Abstract**— *Search engine* merupakan *software* yang memudahkan pengguna untuk menemukan hal-hal sesuai dengan topik yang diinginkan. Pada kasus tertentu, yakni pencarian dokumen yang berupa teks, metode pencarian tradisional, yakni berdasarkan struktur kata/kalimat, yang diterapkan pada *search engine* masih relevan digunakan. Namun, pada kasus khusus dimana hal yang ingin ditemukan oleh pengguna tidak diacu oleh sejumlah teks/kata yang sesuai, penggunaan *search engine* konvensional dapat memberikan hasil yang tidak tepat. Pada penelitian ini, kami mengadopsi metode pencarian berbasis *semantic* untuk diterapkan pada layanan permintaan data di Badan Pusat Statistik (BPS). Ekstraksi *semantic* dari publikasi statistik, yang mengandung tabel, angka, beserta penjelasannya, yang dihasilkan oleh BPS yang akan menjadi *knowledge* dari *search engine* untuk melayani permintaan data oleh pengguna yang selama ini dilayani secara manual oleh petugas melalui *email*. Hasil analisis kami menunjukkan bahwa efisiensi dapat dilakukan dengan mengganti mekanisme *email* dengan penjawab otomatis berbasis *Twitter*®. *Core search engine* yang kami ciptakan kemudian diintegrasikan dengan *Twitter*® agar dapat digunakan secara interaktif.

**Keywords**— *semantic; search engine; data; statistics*

## I. PENDAHULUAN

Badan Pusat Statistik (BPS) adalah salah satu lembaga pemerintahan yang bertanggung jawab kepada Presiden. Data di BPS terbagi menjadi publikasi, data olahan, dan data mentah. Publikasi adalah data yang diterbitkan untuk masyarakat umum dalam bentuk *Portable Document Format (pdf)*. BPS menerbitkan beragam hasil publikasi dari penelitiannya, data dan informasi yang dipublikasi oleh BPS berupa laporan tahunan, triwulan serta laporan-laporan statistik yang telah dibukukan seperti statistik propinsi dan publikasi indeks harga di beberapa kota di Indonesia. Data olahan adalah data yang berasal dari data mentah yang melalui proses pengukuran dan verifikasi berdasarkan fakta-fakta yang terjadi sehingga bisa menghasilkan informasi.

Salah satu Subdirektorat di BPS yang melayani permintaan data untuk umum yaitu Subdirektorat Layanan dan Promosi Statistik. Subdirektorat ini menjadi pintu gerbang masyarakat dalam mencari data yang diinginkan. Tingkat kepuasan

masyarakat dalam mencari data, sangat dipengaruhi oleh subdirektorat tersebut. Dalam melayani permintaan data, konsumen dapat menggunakan layanan *email* atau datang langsung ke Kantor BPS Pusat, Ruang Konsultasi Statistik yang berada di Gedung 2 lantai 3. Permintaan data melalui *email* dilakukan dengan mengirimkan surel ke *bpsHQ@bps.go.id*. Setiap permintaan data yang masuk akan mendapatkan jawaban kepastian layanan paling lama 5 (lima) hari kerja terhitung setelah permintaan dikirimkan, hal ini sesuai dengan UU No.14 Tahun 2008 tentang Keterbukaan Informasi Publik.

Berdasarkan survei pendahuluan yang kami lakukan, permintaan data melalui *email* selama ini kurang efektif. Hal ini dipengaruhi oleh respon dari masing-masing petugas. Ketika konsumen melakukan permintaan data maka konsumen harus menunggu respon dari petugas. Jika petugas yang mendapatkan *email* permintaan data tidak mengetahui letak data maka perlu bertanya kepada staff yang lain. Proses ini membutuhkan *effort* dan waktu yang banyak untuk menjawab *email* permintaan data konsumen. Kurang cepatnya respon permintaan data akan berpengaruh kepada tingkat kepuasan konsumen.

Proses yang berjalan saat ini pada Subdirektorat Layanan dan Promosi Statistik adalah petugas mencari data yang dibutuhkan konsumen secara manual pada *file* publikasi satu per satu dengan memeriksa apakah data yang dibutuhkan tersedia atau tidak. Format publikasi yang terdiri dari tabel dan angka yang memerlukan ketelitian agar data yang disampaikan ke konsumen sesuai dengan yang diinginkan.

Penelitian ini merancang suatu *semantic search engine* untuk menangani pencarian data pada publikasi BPS yang memiliki karakteristik khusus, yakni terdiri dari tabel dan angka. Entitas yang terdapat pada publikasi direpresentasikan menjadi sebuah *knowledge* dan dipetakan dengan data yang bersesuaian agar *search engine* dapat melakukan pencarian secara *semantic*. Untuk keperluan otomatisasi layanan, sistem penjawab otomatis yang diintegrasikan dengan media sosial *Twitter*®. Hasilnya, pertanyaan yang dibuat oleh konsumen dapat dijawab secara otomatis oleh sistem secara *real time*.

## II. PENELITIAN TERKAIT

### A. NELL (Never-Ending Learning Language)

*NELL* adalah sistem pembelajaran mesin *semantic* yang dikembangkan oleh tim peneliti di *Carnegie Mellon University* yang didukung dan dibiayai oleh *DARPA*, *Google*, dan *NSF* [1], [2], [3].

Sejak awal 2010, tim peneliti *Carnegie Mellon* telah menjalankan *NELL* sepanjang waktu, memilah-milah ratusan juta halaman *web* mencari hubungan antara informasi yang sudah tahu dan apa yang ditemukan melalui proses pencarian - untuk membuat koneksi baru dengan cara yang dimaksudkan untuk meniru cara manusia belajar informasi baru. Tujuan dari *NELL* dan sistem pembelajaran *semantic* lainnya, seperti sistem *Watson IBM*, adalah untuk dapat mengembangkan cara menjawab pertanyaan yang diajukan oleh pengguna dalam bahasa alami tanpa campur tangan manusia dalam proses.

Penelitian ini mencoba mengikuti pola pembentukan *knowledge* yang dilakukan pada *NELL*, yakni dalam bentuk graf relasi antarentitas, untuk merepresentasikan informasi yang terdapat pada publikasi BPS. Berbeda dengan *NELL* yang mengolah dokumen teks yang berisi dekripsi/ Pernyataan tentang fakta, *semantic search engine* yang dikembangkan pada penelitian ini dikhususkan untuk publikasi BPS yang terdiri dari tabel dan angka sehingga perlu intervensi manusia untuk membantu mesin membentuk *knowledge*.

### B. Sistem Tanya Jawab

Sistem tanya jawab untuk otomatisasi pelayanan terhadap konsumen telah banyak diteliti sebelumnya. Salah satunya adalah sistem tanya jawab (*QA*) yang interaktif (*IQA*) untuk menangani pelayanan konsumen berupa interaksi antara perusahaan dengan konsumen [4]. Sistem tersebut dirancang untuk memiliki kemampuan berkomunikasi secara aktif dengan pengguna. Untuk melakukan interaksi tersebut, state (*information state*) dijaga oleh system dengan *flag active topic/active node*. Topik yang dimaksud oleh penanya (pengguna) dideteksi melalui ekstraksi *keywords* yang digunakan. *Keywords* tersebut kemudian dikelompokkan untuk mengetahui jenis topik yang diinginkan. Penerapan *IQA* tersebut pada bahasa Indonesia juga telah diteliti [5].

Pada penelitian ini, pendekatan yang dilakukan dalam mendeteksi topik atau publikasi yang dimaksud pengguna adalah melalui tingkat prioritas keyword yang diinputkan. Keyword yang memiliki path yang lebih dekat dengan root akan memiliki prioritas yang tinggi.

### C. Penelitian pada Subdirektorat Layanan dan Promosi Statistik

Penelitian sebelumnya telah dilakukan untuk memaksimalkan pelayanan dalam memberikan layanan konsultasi dan pencarian data yang diberikan oleh BPS kepada masyarakat. Penelitian sebelumnya memfokuskan pada penyediaan sarana konsultasi kepada BPS menggunakan teknologi bahasa alami [6] dan pelayanan pencarian publikasi semiotomatis melalui layanan *email* [7]. Pada penyediaan sarana konsultasi kepada BPS menggunakan teknologi bahasa

alami difokuskan untuk menjawab pertanyaan-pertanyaan konsultasi dari konsumen.

## III. ANALISIS PERMASALAHAN DAN SISTEM USULAN

### A. Proses Pelayanan Permintaan Data

Saat ini sistem yang sedang dijalankan oleh BPS di Subdirektorat Layanan dan Promosi Statistik untuk melayani permintaan dari konsumen adalah melalui *email* dan datang langsung ke BPS [8]. Dibutuhkan waktu dan *effort* yang besar untuk memperoleh data yang diinginkan konsumen dengan cara berkunjung langsung ke kantor BPS, baik ke Kantor Pusat, maupun Propinsi/Kabupaten. Oleh karena itu, BPS juga menyediakan layanan permintaan data dan konsultasi melalui *email*. Bagi konsumen yang ingin meminta data melalui layanan email dapat mengirimkan *email* ke *bpsHQ@bps.go.id*.

*Email* yang masuk ke *inbox* petugas setiap hari berkisar 10-15 *email*. *Email* tersebut dapat berupa permintaan data atau konsultasi. Petugas memeriksa *email* yang masuk satu per satu lalu mencari data yang dibutuhkan oleh konsumen. Data yang dibutuhkan konsumen biasanya terdapat pada publikasi. Jika tidak terdapat dalam publikasi maka petugas memeriksa ketersediaan data mentah. Koleksi publikasi terdapat di komputer masing-masing petugas yang di-*share* melalui *network sharing*. Setelah ditemukan data yang dibutuhkan konsumen maka petugas meng-*attach* data/publikasi tersebut yang kemudian akan dikirimkan ke konsumen melalui *email*.

Untuk meningkatkan pelayanan, saat ini semua publikasi dapat diunduh gratis di *website* BPS sesuai dengan PP PNBPN no. 7 tahun 2015. Publikasi yang diunduh melalui situs BPS dilindungi dengan *watermark*. Dengan kebijakan ini, maka masyarakat dapat dengan mudah mengakses data yang terdapat pada publikasi BPS.

TABLE I. LAPORAN JUMLAH PERMINTAAN DATA TAHUN 2015 DI BPS

No	Jenis Pelayanan	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Agus	Sep	Okt	Nov	Des
1	Telepon	360	435	258	301	267	263	186	80	0	0	0	0
2	Katalog Mikro Data	3	14	4	29	13	4	7	1	0	0	0	0
3	Forum	2	2	1	0	11	8	26	20	0	0	0	0
4	Datang Langsung	241	234	318	428	495	390	240	165	0	0	0	0
5	Email	299	266	239	258	218	211	166	72	0	0	0	0
6	Fax	0	0	1	0	0	0	0	0	0	0	0	0
7	Surat / Pos	1	3	5	3	9	6	5	6	0	0	0	0
8	Surat Eksternal	15	7	8	17	4	5	5	4	0	0	0	0
9	Surat Internal	4	3	5	0	2	2	0	0	0	0	0	0
10	Pemohon Rp0,00	4	1	6	3	2	0	2	0	0	0	0	0
TOTAL		929	966	847	1041	1021	889	637	348	0	0	0	0

\* data terupdate tanggal 11 Agustus 2015

Pada tabel laporan di atas terlihat bahwa *email* merupakan permintaan data yang paling banyak dilakukan setelah datang langsung dan melalui telepon. Hal ini menunjukkan media berupa email termasuk pilihan yang digemari. Email merupakan alternatif yang baik dalam pelayanan permintaan data karena konsumen tidak perlu datang langsung ke BPS sehingga terdapat efisiensi waktu, biaya, dan tenaga. Dengan menggunakan email pelayanan permintaan data menjadi praktis. Hanya dengan menggunakan sebuah *device* seperti *PC Desktop*, *laptop*, *smartphone*, atau sejenisnya, konsumen dapat langsung melakukan permintaan data dengan mengirimkan ke *bpsHQ@bps.go.id*.

Dalam melaksanakan tugasnya, terutama dalam melakukan penyiapan bahan dan pelayanan konsultasi statistik, Subdirektorat Layanan dan Promosi Statistik mengalami kesulitan. Kesulitan yang dialami antara lain ketika ada permintaan data dan harus mencari data tersebut di dalam publikasi. Publikasi yang tersimpan di dalam komputer petugas tidak terstruktur dan dalam jumlah yang banyak sehingga menyulitkan petugas untuk mencarinya. Selain itu kesulitan yang dialami petugas adalah kalimat-kalimat permintaan data dari konsumen yang tidak terstruktur sehingga membutuhkan waktu untuk menangkap permintaan data yang dimaksud oleh konsumen tersebut. Hal ini berakibat proses pelayanan permintaan data yang dibutuhkan konsumen memakan waktu yang lama.

### B. Sistem Usulan

Dari masalah-masalah yang telah dipaparkan, maka untuk meningkatkan layanan, dibutuhkan suatu sistem yang dapat memberikan solusi untuk mengatasi masalah tersebut antara lain:

1. Dibutuhkan suatu sistem yang dapat mempermudah petugas untuk menemukan data/publikasi yang tepat.
2. Kebutuhan terhadap suatu sistem yang tidak mengandalkan pengetahuan dan pengalaman dari petugas.
3. Kebutuhan untuk manajemen koleksi publikasi agar tersusun dengan baik.
4. Kebutuhan untuk membalas permintaan data konsumen dengan cepat.

Solusi yang ditawarkan pada penelitian ini adalah berupa suatu sistem yang dapat memberikan jawaban otomatis untuk permintaan data publikasi. Pada sistem ini output yang dikeluarkan terbatas pada judul dan halaman publikasi data yang dibutuhkan oleh konsumen. Keuntungan dari sisi petugas adalah mengurangi atau menghilangkan *effort* petugas dalam menangkap isi permintaan data yang dibutuhkan oleh konsumen serta mencarinya pada publikasi secara manual. Dari sisi konsumen keuntungan yang didapatkan adalah konsumen dapat segera mengetahui letak data yang dibutuhkan, yakni berupa judul dan halaman publikasi yang relevan. Konsumen tidak perlu mencari secara manual data yang dimaksud pada publikasi yang tersedia di *website* BPS. Selain itu konsumen dapat memperoleh data yang dibutuhkan secara *real time*. *Core* dari sistem yang dibangun berupa *semantic search engine*.

## IV. RANCANGAN SISTEM

Knowledge base dari system yang dibangun mengikuti pola semantic dari publikasi yang disusun oleh BPS secara umum.

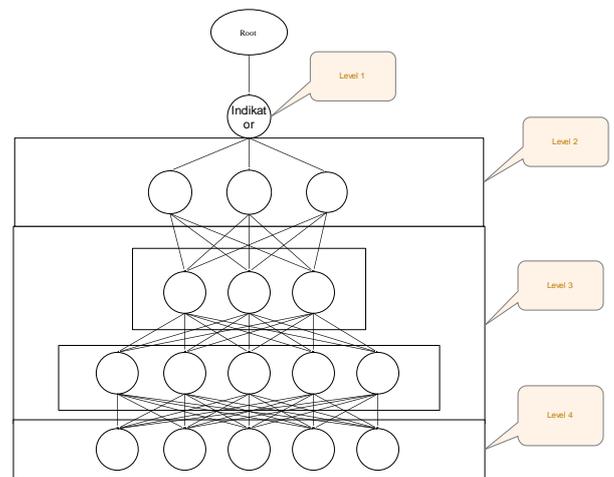


Fig. 1. Model konseptual sistem *semantic*

Fig. 1 menunjukkan model penyusunan korpus secara general dimana *root* adalah node paling atas yang berisi judul dan halaman publikasi. *Root* tersebut memiliki sejumlah relasi dengan entitas lain yang dibagi kedalam beberapa level.

- a. Level pertama adalah indikator, indikator adalah inti dari konten pada tabel Misalnya pdb, pdrb, ipm, dan sejenisnya.
- b. Level kedua adalah penjelas dari indikator dimana pada level ini berisi keterangan dari indikator misalnya pdb atas dasar harga berlaku, pdb atas dasar harga konstan, pdb menurut lapangan usaha industri pengolahan, dan lain-lain. Kata-kata setelah indikator tersebut merupakan penjelas tambahan dari indikator yang lebih spesifik.
- c. Level ketiga adalah wilayah dimana pada level provinsi ini terdiri dari node node provinsi dan node kabupaten/kota.
- d. Level keempat adalah dimensi waktu dimana pada level ini node berisi tahun, periode, dan satuan waktu lainnya seperti trwiulan, caturwulan, semester.

Representasi knowledge yang dirancang pada sistem yang dibangun dapat dijelaskan melalui contoh pada Fig. 2.

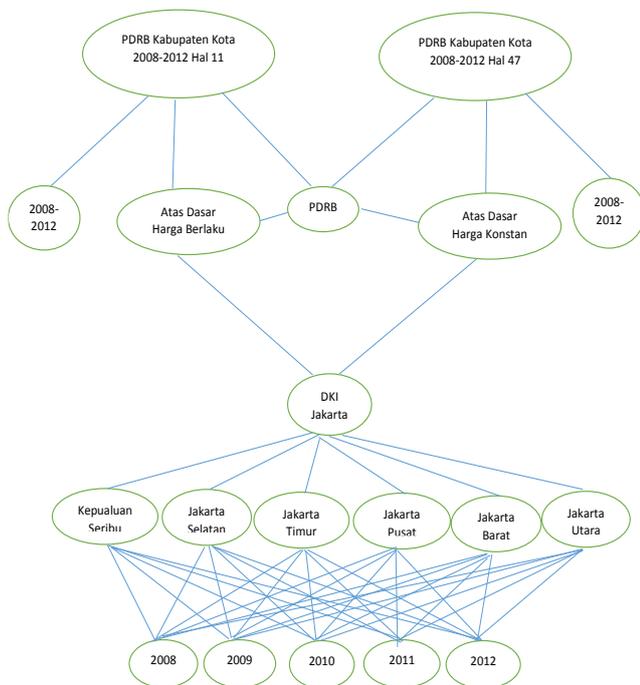


Fig. 2. Contoh representasi *knowledge*

Pada graf Fig. 2 terlihat bahwa publikasi PDRB Kabupaten/Kota 2008-2012 halaman 11 dan publikasi PDRB Kabupaten/kota 2008-2012 halaman 47 ada sebagai hasil akhir dari relasi graf. Publikasi tersebut mengandung informasi indikator PDRB dimana indikator PDRB itu sendiri memiliki dua keterangan yaitu atas dasar harga berlaku dan atas dasar harga konstan. Keterangan PDRB atas dasar harga berlaku dan PDRB atas dasar harga konstan memiliki relasi dengan provinsi DKI Jakarta dimana provinsi DKI Jakarta terpecah lagi menjadi level Kabupaten/Kota. Setiap level tersebut memiliki relasi dengan relasi provinsi DKI Jakarta. Pada level Kabupaten/Kota juga memiliki relasi terhadap tahun.

Pemodelan graf direpresentasikan dalam bentuk *relational database*. Sistem *semantic search engine* memiliki basis pengetahuan (*knowledge base*) yang terdiri dari 3 entitas/tabel yaitu tabel *keyword*, tabel relasi, dan tabel *page info*. Tabel *keyword* berisi kata-kata penting yang berada pada publikasi dimana setiap kata tersebut memiliki definisi yang disebut sebagai *tag* atau penanda. Tabel relasi adalah tabel yang berisi relasi-relasi dari *id* pada tabel *keyword*. Relasi pada tabel *keyword* ini kemudian membentuk suatu jaringan semantik. Tabel *page info* adalah tabel yang berisi informasi tentang judul, halaman, dan *link/URL* publikasi.

Untuk menemukan publikasi yang dimaksud oleh penanya (pengguna), *query* dilakukan secara rekursif untuk tiap kata/frase yang mungkin terbentuk dari kalimat pertanyaan. *Query* dilakukan mulai dari kata pertama pada kalimat pertanyaan ke seluruh entitas yang terdapat pada database, kemudian dilanjutkan dengan frase yang dibentuk dari 2 kata pertama, 3 kata pertama, hingga 5 kata pertama yang merupakan batasan panjang frase maksimal yang ditetapkan pada penelitian ini. Hal yang sama dilakukan dengan *starting point* kata kedua, kata ketiga, dan seterusnya. Entitas yang

memiliki kedekatan dengan *root* akan memiliki skor yang lebih tinggi.

## V. HASIL DAN PEMBAHASAN

Pada eksperimen yang dilakukan, jumlah node entitas (tidak termasuk node root, contoh node entitas seperti *ipm*, *pdrb*, *pdn*, *amh*, *jabar*, *jatim*, dan lain-lain) yang dibangun di dalam sistem sebanyak 167 entitas. Sedangkan jumlah node root (informasi mengenai publikasi seperti judul publikasi dan halaman publikasi) sebanyak 28 entitas. Jumlah relasi yang disusun sebanyak 416 relasi.

Pengujian dilakukan dengan memasukkan 6 *email* permintaan data yang diperoleh dari Subdirektorat Layanan dan Promosi Statistik dengan hasil yang diperoleh sebagai berikut.

TABLE II. HASIL PENGUJIAN I

No. Email	Hasil	Persentase
1.	1 dokumen yang relevan dari 1 yang ditampilkan	$1/1 * 100\% = 100\%$
2.	4 dokumen yang relevan dari 5 yang ditampilkan	$4/5 * 100\% = 80\%$
3.	4 dokumen yang relevan dari 5 yang ditampilkan	$4/5 * 100\% = 80\%$
4.	4 dokumen yang relevan dari 5 yang ditampilkan	$4/5 * 100\% = 80\%$
5.	4 dokumen yang relevan dari 5 yang ditampilkan	$4/5 * 100\% = 80\%$
6.	2 dokumen yang relevan dari 2 yang ditampilkan	$2/2 * 100 = 100\%$

Terlihat bahwa rata-rata hasil persentase yang didapatkan pada pengujian terhadap keenam *email* adalah sebesar 86.67%. Hasil tersebut telah menunjukkan bahwa sistem telah akurat yang baik.

Selanjutnya dilakukan pengujian lagi dengan menghilangkan kalimat pembuka dan penutup pada *email*. Dengan demikian hanya konten intinya saja yang diproses oleh mesin pencari. Perbandingan hasil yang diperoleh adalah sebagai berikut.

TABLE III. HASIL PENGUJIAN II

No. Email	Dengan Kalimat Pembuka dan Penutup	Tanpa Kalimat Pembuka dan Penutup
1.	$1/1 * 100\% = 100\%$	$1/1 * 100\% = 100\%$
2.	$4/5 * 100\% = 80\%$	$4/5 * 100\% = 80\%$
3.	$4/5 * 100\% = 80\%$	$4/5 * 100\% = 80\%$
4.	$4/5 * 100\% = 80\%$	$4/5 * 100\% = 80\%$
5.	$4/5 * 100\% = 80\%$	$4/5 * 100\% = 80\%$
6.	$2/2 * 100 = 100\%$	$2/2 * 100 = 100\%$

Jika diambil konten intinya saja, ternyata hasil yang diperoleh sama. Oleh karena pembuka dan penutup tidak memberikan pengaruh terhadap keakuratan permintaan data melalui sistem *semantic search engine*, maka penulis mengajukan sebuah wadah baru yang relevan dengan kondisi tersebut, yakni integrasi *search engine* dengan media sosial

twitter. Keuntungan yang diperoleh dengan integrasi tersebut diantaranya adalah:

1. Lebih interaktif,
2. Pertanyaan yang diajukan lebih terfokus pada intinya karena *twitter* membatasi penggunaannya untuk melakukan *twit* sebanyak 140 karakter.
3. *Real-time*. Dalam hal ini konsumen tidak perlu lagi menunggu petugas untuk mencarikan data karena sistem yang akan mencari data yang dibutuhkan secara otomatis dan menjawabnya melalui Twitter.
4. Kemudahan dalam permintaan data. Hal ini dikarenakan *twitter* merupakan salah satu media sosial yang ramai penggunaannya dan sering dikunjungi.

Berikut adalah salah satu contoh percakapan *twitter* hasil ujicoba sistem.



Fig. 3. Contoh percakapan pada *twitter* hasil ujicoba

Jika data tersedia maka *twitter* langsung menjawab publikasi yang dibutuhkan oleh konsumen. Dalam hal ini *output* jawaban yang ditampilkan di *twitter* adalah Judul dan halaman yang direkomendasikan oleh sistem *semantic search engine*. Jika data tidak tersedia maka sistem akan langsung memforward *twit* dari konsumen ke *email* petugas untuk ditindaklanjuti.

## VI. KESIMPULAN

Sistem usulan yang dibangun dapat menjawab otomatis permintaan ketersediaan data publikasi. Berdasarkan hasil uji coba, sistem dapat merekomendasikan ketersediaan publikasi dengan akurat. Namun, basis pengetahuan perlu diinput secara manual mengikuti semantik dari kumpulan publikasi data yang ada. Penelitian selanjutnya diharapkan dapat memberikan solusi untuk mengotomatisasi dokumen publikasi kedalam bentuk representasi *knowledge*. *IQA* yang memiliki sifat yang lebih interaktif juga diharapkan dapat dikembangkan lebih lanjut. Dengan integrasi dengan *twitter*, pertanyaan pengguna dapat lebih fokus dan direspon dengan *real time*.

## REFERENSI

- [1] T. Mitchell, W. Cohen, E. Hruschka, P. Talukdar, J. Betteridge, A. Carlson, B. Dalvi, M. Gardner, B. Kisiel, J. Krishnamurthy, N. Lao, K. Mazaitis, T. Mohamed, N. Nakashole, E. Platanios, A. Ritter, M. Samadi, B. Settles, R. Wang, D. Wijaya, A. Gupta, X. Chen, A. Saparov, M. Greaves, and J. Welling, "Never-Ending Learning," in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*, 2015.
- [2] A. Carlson, J. Betteridge, B. Kisiel, B. Settles, E. R. H. Jr., and T. M. Mitchell, "Toward an Architecture for Never-Ending Language Learning," in *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence (AAAI 2010)*, 2010.
- [3] A. Carlson, J. Betteridge, E. R. Hruschka Jr., and T. M. Mitchell, "Coupling Semi-supervised Learning of Categories and Relations," in *Proceedings of the NAACL HLT 2009 Workshop on Semi-Supervised Learning for Natural Language Processing*, 2009, pp. 1–9.
- [4] A. Purwarianti and Z. Hakim, "INTERACTIVE QUESTION ANSWERING FOR CUSTOMER SERVICE REPRESENTATIVE USING," pp. 1–6, 2012.
- [5] M. C. Wijanto and A. Purwarianti, "Indonesian dialogue management system using information state for independent closed domain," in *Computational Intelligence and Cybernetics (CYBERNETICSCOM), 2013 IEEE International Conference on*, 2013, pp. 26–30.
- [6] M. Safi'udin, "Sistem Tanya Jawab Berbahasa Alami Domain Informasi Statistik sebagai Media Layanan Konsultasi Statistik di Badan Pusat Statistik," 2013.
- [7] H. Indrajati, "Sistem Aplikasi Layanan Permintaan Data Olahan dan Publikasi BPS Melalui Email dengan Pencarian Semiotomatis," 2014.
- [8] Badan Pusat Statistik, "Website BPS." [Online]. Available: <http://bps.go.id>. [Accessed: 06-Apr-2015].