

BAB III

LANDASAN TEORI

3.1 Taksonomi

Taksonomi merupakan kaidah dan prinsip yang meliputi pengklasifikasian obyek. Salah satu bentuk Taksonomi adalah klasifikasi atau pengkategorian. Klasifikasi atau pengkategorian diperuntukan untuk sesuatu yang berkelompok dan memiliki keterkaitan satu dengan yang lain. Secara tidak sadar sebetulnya kita sudah terbiasa menggunakan metode ini, misalnya jika kita pergi ke supermarket, untuk mencari apel pasti kita akan mencari di rak buah-buahan. Atau coba cek koleksi MP3 di komputer kita. File MP3 biasanya akan kita susun dalam folder jazz, klasik hingga dangdut (Sugangga, 2008).

AP news Taxonomy adalah sebuah layanan yang digunakan untuk mengekstrak kosakata baku yang luas untuk menjelaskan berita berbahasa inggris. Istilah dalam kosakata mencakup semua hal dari berita seperti: subyek, orang, tempat, organisasi, dan banyak lagi. Ketika pengguna mengirimkan konten ke layanan *tagging service* yang terotomatisasi, data yang dihasilkan adalah data yang diambil dari kosa kata tersebut. Penerbit juga dapat memilih untuk mengintegrasikan *AP news taxonomy* ke dalam sistem penerbitan mereka sendiri untuk mendukung *tagging manual*. Selain standar terminologi dan id yang unik, taksonomi menyimpan berbagai rincian tentang orang, tempat, dan topik di dalamnya. Informasi ini dapat meningkatkan kekuatan pengalaman pencarian,

browsing dan penemuan, atau menampilkan informasi. Sebagai contoh:

- a. Sinonim, akronim, dan variasi ejaan.
- b. Sifat orang, tempat dan hal-hal seperti nomor seragam atlet, garis lintang dan bujur dari lokasi geografis, atau simbol stok waktu untuk sebuah perusahaan.
- c. Hubungan antara konsep, seperti antara dua orang (orang tua dan anak), atau antara orang dan kelompok atau organisasi (atlet dan tim).
- d. Struktur hierarkis untuk subyek dan lokasi geografis, untuk memungkinkan kedua pencarian yang luas dan sempit.

Contoh untuk 5 kategori dari *AP news taxonomy* sebagai berikut:

a. AP subyek

Berbagai topik terstruktur secara hierarkis mulai dari kategori (kejahatan) ke konsep tertentu seperti (senjata api ilegal). Juga mencakup banyak peristiwa yang bernama seperti Academy Awards dan Tour de France. Semuanya memncapai lebih dari 4.000 istilah.

b. AP geografis

Lebih dari 2.000 nama tempat geografis tersusun secara hierarkis seperti benua, wilayah dunia, negara-negara, wilayah, ibukota nasional, kota-kota besar dunia, negara di Amerika Serikat, provinsi Kanada, dan sejumlah kota-kota besar di Amerika Serikat.

c. AP organisasi

Organisasi dan lembaga dari berbagai sektor: organisasi pemerintah, non-profit, tim olahraga, perguruan tinggi dan universitas, kelompok politik dan ideologi, institusi budaya, dan banyak lagi. Lebih dari 1.000 istilah yang berbeda.

d. AP orang

Selebriti, seniman, desainer, penulis, pemimpin bisnis, tokoh politik, tokoh olahraga, royalti, dan pembuat berita terkenal sedunia atau dalam tingkat nasional Amerika Serikat. Terutama mencakup seluruh pembuat berita Amerika Serikat dalam politik, hiburan dan olahraga, termasuk daftar nama lengkap untuk tim olahraga profesional ternama, pemain laki-laki divisi satu NCAA basket dan sepak bola, semua para pejabat Amerika Serikat di tingkat federal dan gubernur, dan semua calon untuk kantor-kantornya. Mencakup lebih dari 90.000 individu.

e. AP perusahaan

Lebih dari 35.000 perusahaan dagang publik, termasuk semua perusahaan dengan perdagangan saham utama pada setiap 70 bursa saham ternama dunia atau perdagangan sebagai ADR di bursa Amerika (Ap, 2008).

3.2 Kolaboratif

Kolaboratif adalah situasi dimana terdapat dua atau lebih orang belajar atau berusaha untuk belajar sesuatu secara bersama-sama. Tidak seperti

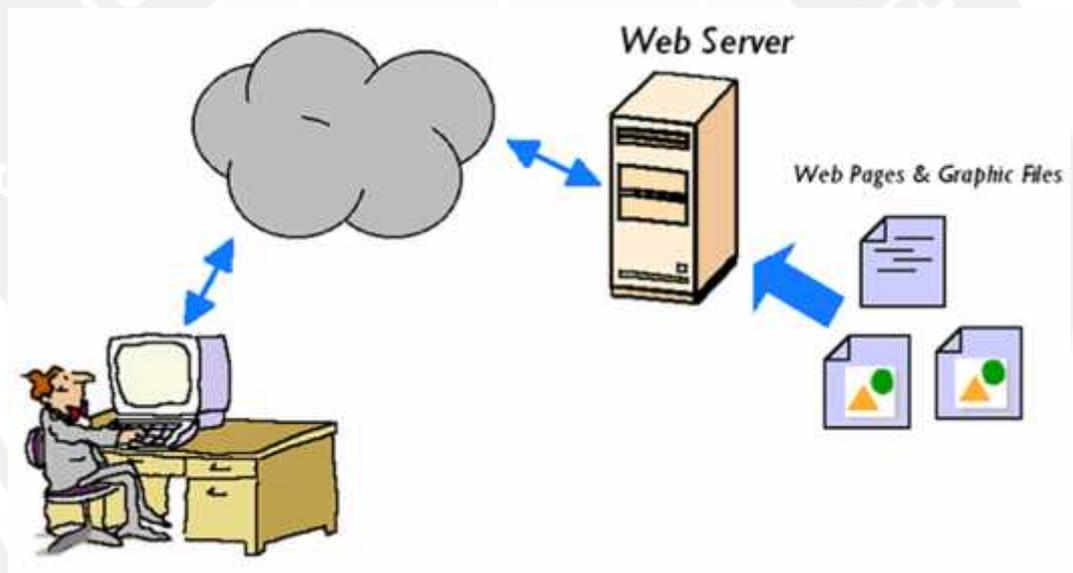
belajar sendirian, orang yang terlibat dalam *collaborative learning* memanfaatkan sumber daya dan keterampilan satu sama lain (meminta informasi satu sama lain, mengevaluasi ide-ide satu sama lain, memantau pekerjaan satu sama lain, dll). Lebih khusus, *collaborative learning* didasarkan pada model di mana pengetahuan dapat dibuat dalam suatu populasi di mana anggotanya secara aktif berinteraksi dengan berbagi pengalaman dan mengambil peran asimetri (berbeda). Dengan kata lain, *collaborative learning* mengacu pada lingkungan dan metodologi kegiatan peserta didik melakukan tugas umum di mana setiap individu tergantung dan bertanggung jawab satu sama lain. Hal ini juga termasuk percakapan dengan tatap muka dan diskusi dengan komputer (*forum online, chat rooms, dll.*). Metode untuk memeriksa proses *collaborative learning* meliputi analisis percakapan dan analisis wacana statistik.

3.3 Taksonomi Kolaboratif

Taksonomi kolaboratif adalah proses pengkategorian suatu data berkelompok dan memiliki keterkaitan satu dengan yang lain. Dalam proses pengkategorian diperlukan suatu tim untuk berdiskusi, mengevaluasi, dan menganalisis suatu data yang berkaitan, sehingga data yang dikelompokkan menjadi data yang terstruktur.

3.4 Web

Web server digunakan sebagai tempat aplikasi web dan sebagai penerima *request* dari *client* (Zahni & Warman, 2013). Web merupakan sebuah perangkat lunak dalam server yang berfungsi menerima permintaan (*request*) berupa halaman web melalui HTTP atau HTTPS dari klien yang dikenal dengan *browser* web dan mengirimkan kembali (*response*) hasilnya dalam bentuk halaman-halaman web yang umumnya berbentuk dokumen HTML (Solichin, 2012).



Gambar 3.1 Arsitektur Web Server

Gambar 3.1 menjelaskan bahwa *Web Server* berguna untuk menampung semua *file* seperti halaman pada *website*, *basis data*, *file gambar*, dan *file web service*.

Didalam *Web Server* terdapat *Web service*, *Web service* adalah kumpulan fungsi dan *method* yang terdapat pada sebuah *web server* yang dapat dipanggil oleh klien dari jarak jauh, kemudian untuk memanggil *method-method* tersebut kita bisa menggunakan aplikasi yang dibuat dengan bahasa pemrograman apa saja yang dijalankan pada *platform* apa saja (Marthasari, et al., 2010). *Web*

service memiliki hasil yang dapat berbentuk XML atau dalam bentuk JSON. Kedua hasil *web service* tersebut dinamakan dengan teknik SOAP dan Rest.

a. SOAP

SOAP singkatan dari *Simple Object Access Protocol* adalah teknik yang digunakan untuk bertukar pesan berbasis [XML](#) melalui [jaringan komputer](#) untuk berkomunikasi dengan program pada OS yang sama maupun berbeda dengan menggunakan [HTTP](#) dan XML sebagai mekanisme untuk pertukaran data.

b. REST

Singkatan dari *Representational State Transfer*, merupakan arsitektur software untuk distribusi hypermedia diantaranya *world wide web*. Secara spesifik REST merujuk kepada suatu koleksi prinsip-prinsip arsitektur jaringan yang menggariskan pendefinisian dan pengalamatan sumber daya (Mellawati, et al., 2012). JSON merupakan salah satu *output* atau keluaran dari teknik REST.

Berikut adalah Tabel perbandingan kedua teknik diatas:

SOAP	REST
Teknologi lama	Teknologi baru
Interaksi Client-Server sangat bergantung	Interaksi Client-Server kurang bergantung
Membutuhkan binary untuk <i>parsing</i> data	Mendukung semua tipe data untuk di <i>parsing</i>
Nilai balikan selalu XML	Nilai balikan fleksibel,

data	tergantung tipe data yang ingin dilempar
Susah untuk dibuat, membutuhkan <i>tools</i> tertentu	Mudah untuk dibuat

Tabel 3.1 Perbandingan SOAP dan REST

Tabel diatas diambil dari dokumentasi jurnal (Wagh & Thool, n.d.)

Basis data yang digunakan adalah MySQL, MySQL merupakan *software* sistem manajemen *basis data* (*Basis data Management System* - DBMS) yang sangat populer digunakan untuk membangun aplikasi *web* sebagai sumber data atau basis datanya. Kelebihan MySQL adalah kemudahannya, dan cepat dalam mengeksekusi *query*. MySQL merupakan sistem manajemen *basis data* yang bersifat *open source* (Sasmita, et al., 2009).

3.5 Tagging

Tagging service merupakan layanan untuk menganalisis isi berita berbahasa Inggris dan secara otomatis mengembalikan metadata yang relevan, menggunakan terminologi standar dari *AP news taxonomy*. Proses yang mengidentifikasi orang, perusahaan, lokasi geografis, organisasi, dan berbagai susunan subyek. Meskipun sistem akan mengenali dan mengembalikan entitas tertentu yang ditemukan dalam konten yang didapatkan (teks yang terekstrak), sistem juga bisa melakukan lebih dari itu, dengan menggunakan aturan semantik yang diciptakan manusia untuk mengidentifikasi topik -

topik yang mungkin sama sekali tidak dapat disebutkan secara jelas di dalam teks. Misalnya, sebuah cerita tentang seorang bintang musik di negara tertentu dapat mengacu pada area subyek "musik negara", bahkan jika kata "musik" tidak muncul dalam cerita. Aturan pengelolaan yang dibuat manusia memperbolehkan untuk kontrol yang lebih tepat atas kinerja layanan (Ap, 2008).

