

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Konten situs yang bagus menjadi tidak bermanfaat jika tidak ada yang membacanya. Miskin pengunjung karena konten tidak ditemukan dengan baik oleh mesin pencari menjadi salah satu masalah terbesar untuk situs-situs baru. Situs lama dan mempunyai nama besar memang bukan menjadi masalah karena reputasi yang mereka miliki.

Konten yang bagus merupakan salah satu aspek yang menjadi pertimbangan sebuah situs tampil di halaman depan mesin pencari atau bahkan pada nomer satu hasil pencarian. Tetapi masih banyak yang menjadi pertimbangan oleh mesin pencari. Konten "Bagus" bersifat subyektif yang sukar diukur dengan komputasi. Ukuran konten "Bagus" didapat dari umpan balik pengguna dan reputasi dari situs yang bersangkutan.

Selain bagus, panjang konten, umur konten, kemiripan konten dengan kata kunci dan keunikan konten. Konten salinan biasanya mempunyai nilai SERP yang lebih rendah dari konten aslinya. Penentunya adalah umur dokumen, semakin tua maka semakin baik reputasinya.

Keragaman konten, jumlah konten dan keterkaitan diantaranya pada suatu situs juga menentukan nilai SERP. Konten-konten yang unik dan masih dalam suatu pokok permasalahan biasanya dapat terhubung dengan baik.

Butuh usaha yang besar untuk membangun konten yang unik dan bagus, sehingga tidak dapat dilakukan dengan cepat. Generator konten merupakan salah satu alternatif untuk meningkatkan SERP situs dengan cepat. Konten-konten dapat dihasilkan secara otomatis dengan bantuan feed atau rss dari mesin pencari.

Diharapkan dengan jumlah konten yang banyak, semakin mudah situs untuk tampil di halaman mesin pencari sehingga pengunjung situs dapat ditingkatkan dengan mudah. Walaupun demikian generator konten merupakan cara sementara

untuk membuat situs menjadi terkenal, konten yang bagus dan unik merupakan kunci utama dari keberhasilan suatu situs menarik pengunjung. Pengunjung setia situs tetap menjadi satu-satunya sumber lalu-lintas situs terbaik.

1.2 PERUMUSAN MASALAH

Berdasarkan latar belakang di atas, maka permasalahan yang dapat dirumuskan adalah bagaimana membuat aplikasi generator konten pada suatu situs yang dapat bekerja secara otomatis.

1.3 BATASAN MASALAH

Dalam penelitian ini ada beberapa pembatasan masalah yang dilakukan, yaitu: hanya melakukan implementasi aplikasi pada lingkungan CMS wordpress dan ujicoba pada 2 situs lama dengan jumlah pengunjung kurang dari 100 orang perhari.

1.4 TUJUAN DAN MANFAAT PENELITIAN

Tujuan yang ingin dicapai dalam penelitian ini adalah merancang dan membuat aplikasi generator konten berbasis web yang dapat membangun konten pada situs ber-CMS wordpress dengan tujuan meningkatkan SERP situs. Manfaat yang diharapkan dari penelitian ini adalah : Dapat meningkatkan utilitas dan kemanfaatan suatu situs di internet dengan melihat dari jumlah pengujung di situs tersebut.

1.5 METODOLOGI PENELITIAN

1.5.1. Obyek Penelitian

Obyek penelitian dari penelitian ini adalah artikel yang terdapat pada situs <http://lyricofsong.info> dan <http://lowongankerja-terbaru.info>

Data Yang diperlukan

Merupakan data yang mendukung dalam penelitian ini meliputi data primer dan data sekunder.

Data primer : Data yang diperoleh langsung dari internet dalam hal ini adalah hasil pencarian dari mesin pencari.

Data Sekunder : Data yang diperoleh dengan membaca dan mempelajari referensi mengenai pengambilan konten dan generator konten dari mesin pencari.

1.5.2 Teknik Pengumpulan Data

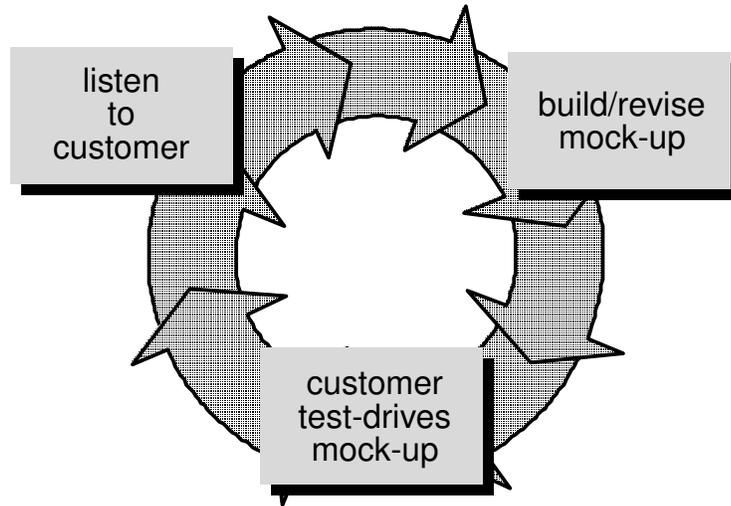
Pengumpulan data dimaksudkan agar mendapatkan bahan-bahan yang relevan, akurat dan reliable. Maka teknik pengumpulan data yang dilakukan dalam penelitian ini adalah sebagai berikut :

Observasi : Dengan melakukan pengamatan dan pencatatan secara sistematis tentang hal-hal yang berhubungan kemampuan generator konten.

Studi Pustaka : Dengan pengumpulan data dari bahan-bahan referensi, arsip, dan dokumen yang berhubungan dengan permasalahan dalam penelitian ini.

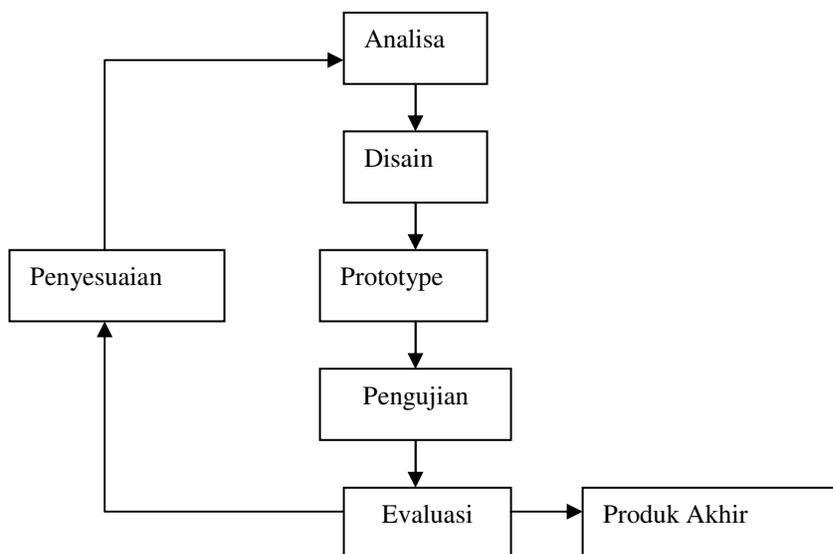
1.5.3. Metode Pengembangan

Penelitian ini menggunakan model *prototyping* . di dalam model ini sistem dirancang dan dibangun secara bertahap dan untuk setiap tahap pengembangan dilakukan percobaan-percobaan untuk melihat apakah sistem sudah bekerja sesuai dengan yang diinginkan. Sistematika model *prototyping* terdapat pada gambar 1.1, dan pada gambar 1.2 memperlihatkan tahapan pada prototyping



Gambar 1.1. Sistematika Prototyping (Pressman, 1997:40)

Berikut adalah tahapan yang dilakukan pada penelitian ini dengan metode pengembangan prototyping.



Gambar 1.2. Tahapan Protoyping

Analisa : Pada tahap ini dilakukan analisa tentang masalah penelitian dan menentukan pemecahan masalah yang tepat untuk menyelesaikannya. Dalam tahap ini juga dilakukan penyusunan Software Requirement Specification.

Disain : Pada tahap ini dibangun rancangan sistem dengan beberapa diagram bantu DFD dan ER-D. Perancangan alur program juga dilakukan tahap ini dengan menggunakan diagram Flow Chart.

Prototype : Pada tahap ini dibangun aplikasi berbasis web yang sesuai dengan disain dan kebutuhan sistem.

Pengujian: Pada tahap ini dilakukan pengujian hasil dari proses indek pada aplikasi yang dibangun.

Evaluasi : Pada tahap ini dilakukan evaluasi apakah performa aplikasi sudah sesuai dengan yang diharapkan, apabila belum maka dilakukan penyesuaian-penyesuaian secukupnya.

Penyesuaian : Tahap ini dilakukan apabila pada evaluasi performa aplikasi kurang memadai dan dibutuhkan perbaikan, tahap ini melakukan penyesuaian dan perbaikan pada aplikasi sesuai dengan kebutuhan

1.6 Sistematika Penulisan

Sistematika penulisan terdiri dari empat bab yang masing-masing bab menguraikan hal-hal yang berbeda.

Bab I Pendahuluan

Pada bab ini diuraikan mengenai permasalahan yang dibahas secara umum yang meliputi : latar belakang, perumusan dan pembatasan masalah, tujuan dan manfaat penelitian, metode pengumpulan data, metodologi penelitian serta sistematika penelitian.

Bab II Landasan Teori

Dalam Bagian ini memuat hal-hal teoritis yang ada hubungannya dengan penyelesaian masalah dalam penelitian ini. Pada bab ini di uraikan antara lain tentang mysql, php dan generator konten

Bab III Perancangan Sistem

Dalam bagian ini dibahas tentang rancang bangun aplikasi. Rancangan meliputi struktur program dan Flowchart.

Bab IV Implementasi

Dalam bagian ini dibahas tentang langkah-langkah implementasi untuk perangkat lunak aplikasi yang telah selesai di rancang pada bab III. Disini disertakan juga kode sumber dari fungsi-fungsi utama.

Bab V Kesimpulan Dan Saran

Pada bab ini berisi kesimpulan dan saran dari penelitian ini. Selain hasil penelitian berupa perangkat lunak aplikasi, juga dihasilkan juga saran-saran untuk penelitian lebih lanjut pada bidang generator konten.

BAB II

LANDASAN TEORI

2.1 Internet

2.1.1 Pengertian Internet

Internet dalam bahasa Inggris merupakan singkatan “International Networking”. Pengertian internet secara umum adalah jaringan komputer yang ada di seluruh dunia di mana setiap komputer memiliki alamat (internet Address) yang dapat digunakan untuk mengirim data atau informasi. Dalam hal ini komputer yang dulunya berdiri sendiri menjadi dapat berhubungan langsung dengan host – host atau komputer – komputer yang lainnya. Bentuk data dapat ditransmisikan melalui internet mencakup teks, suara, udara, video, piranti lunak.

Menurut Ause (1997 : 1), internet merupakan sekumpulan jaringan yang saling terhubung dengan jaringan lain menggunakan bahasa yang dikenal dengan TCP/IP.

Sedangkan menurut Ellsworth (1995 : 437), internet adalah jaringan komunikasi digital yang menghubungkan jaringan – jaringan yang lebih kecil dari banyak negara di seluruh dunia. Internet menggunakan protokol standar yang disebut TCP/IP.

Dari beberapa pengertian internet di atas dapat ditarik kesimpulan bahwa internet adalah merupakan suatu jaringan komunikasi digital global yang dapat menembus batas geografis dan menghubungkan banyak komputer di berbagai negara dengan menggunakan suatu bahasa atau protokol standar yang dikenal dengan nama TCP/IP.

2.2 Protokol internet

2.2.1 Pengertian protokol internet

Protokol dalam dunia komunikasi data komputer digunakan untuk mengatur bagaimana sebuah komputer berkomunikasi dengan komputer lain. Komputer yang

terhubung ke internet berkomunikasi dengan menggunakan protokol yang sama, karena protokol ini berfungsi mirip dengan bahasa.

2.2.2 Hypertext Transport Protokol (HTTP)

Salah satu protokol yang sering digunakan dalam dunia internet adalah HTTP (Hypertext Transport Protokol). Protokol HTTP ini digunakan untuk berbagai jenis layanan dalam WWW (World Wide Web) di jaringan TCP/IP. Protokol HTTP juga dapat digunakan untuk berkomunikasi antara web browser dan web server satu sama lain.

HTTP akan kita gunakan jika pemakai hendak mengakses suatu website tertentu. HTTP memiliki tugas yaitu untuk mentransfer dokumen atau file berupa hypertext yang dalam pelaksanaannya dikenal dengan nama HTML.

Dengan demikian HTTP akan mentransfer HTML ke browser dari server tempat HTML tersebut di simpan. Protokol HTTP berifat request response, yaitu dalam protokol ini client menyampaikan pesan request ke server dan server kemudian akan memberikan respon yang sesuai dengan request tersebut.

Protokol HTTP ini pada dasarnya keseluruhan beroperasi tanpa sepengetahuan pemakai, tidak setiap pemakai diwajibkan perlu tahu TCP/IP bila pemakai hanya sekedar menggunakan internet atau web untuk kebutuhannya.

2.2.3 Transmission Control Protokol / Internet Protokol (TCP/IP)

Internet beroperasi menggunakan satu set protokol yang mengontrol dan mengarahkan data di dalam jaringan. Protokol – protokol ini disebut sebagai TCP/IP.

Jaringan besar yang menyusun internet memberikan peluang bagi penggunaanya supaya dapat saling berkomunikasi dengan menggunakan dua protokol yaitu TCP dan IP.

Protokol TCP/IP adalah suatu tipe protokol yang di gunakan untuk melakukan komunikasi data dan informasi di internet. Sedangkan protokol sendiri adalah suatu kesatuan prosedur atau bahasa yang memungkinkan 2 atau lebih sistem yang berbeda

dapat saling berkomunikasi. Protokol ini merupakan suatu protokol terbuka dimana protokol ini dapat di terapkan dan menghubungkan berbagai sistem tanpa memandang spesifikasi ataupun tipe mesin komputer yang digunakan.

Dalam membawa suatu informasi pada internet merupakan tanggung jawab TCP, di mana TCP memenggal informasi menjadi paket – paket yang berisi data untuk ditransfer dan di susun ulang di tempat tujuan. Lalu IP bertugas memastikan pengiriman data yang akurat ke alamat yang benar.

TCP/IP terdiri dari beberapa layer. Berikut merupakan fungsi dari masing – masing layer TCP/IP adalah :

Physical Layer : Bagian ini berfungsi melewati data yang di kirim melalui media fisik seperti konektor dan kabel.

Data Link Layer : Bagian ini berfungsi mempaketkan data ke dalam bentuk frame.

Internet Protokol : Berfungsi meroute data antar sistem.

TCP: TCP berfungsi meneruskan data dari link layer dan mengubahnya ke dalam bentuk paket.

Application and Service : Bagian ini berfungsi meneruskan paket ke software aplikasi yang biasa digunakan oleh user.

2.3 Layanan Internet

2.3.1 World Wide Web (WWW)

World Wide Web (WWW) adalah jaringan komputer yang terdiri dari client dan server dengan menggunakan software khusus membentuk sebuah jaringan yang disebut jaringan client-sever.

WWW juga merupakan jaringan dokumen yang sangat besar yang saling dihubungkan satu sama lain, satu set protokol yang mendefinisikan bagaimana sistem bekerja dan menstransfer data, dan sebuah perangkat lunak yang membuatnya bekerja dengan mulus.

WWW ada 2 hal penting yaitu web server dan web browser. Informasi yang di letakkan di WWW disebut “homepage” dan setiap homepage memiliki alamatnya

masing – masing. WWW menggunakan teknik hypertext dan multimedia yang membuat internet mudah digunakan dan di jelajahi.

2.3.2 Electronic Mail (E-mail)

E-mail merupakan cara pengiriman surat atau pesan secara elektronik. E-mail juga merupakan penggunaan teknologi pasar elektronik yang memungkinkan pengguna komputer untuk berkomunikasi dengan pengguna komputer lainnya dengan berbagai tujuan. E-mail menjadi salah satu alasan mengapa komputer saling terhubung. Transfer E-mail yang lebih cepat adalah server menstransfer E-mail dengan menggunakan STMP (Single Mail Transfer Protokol).

Dengan E-mail dapat mengirim file – file berupa program, gambar, grafik, video dan lain sebagainya. Serta dapat juga mengirim ke lebih dari 1 orang sekaligus pada saat bersamaan tanpa mengenal batas ruang dan waktu.

2.3.3 Feed Back

Merupakan pesan umpan balik dari konsumen yang berisi penilaian terhadap suatu proses layanan yang diberikan oleh perusahaan.

2.3.4 Uniform Resource Locator (URL)

URL adalah suatu sarana yang digunakan untuk menentukan lokasi informasi pada suatu web server. URL merupakan cara standar untuk menentukansitus atau halaman pada internet.

URL sama halnya dengan alamat dalam surat biasa yang terdiri dari kode pos dan alamat serta nomor jalan. Begitu juga dengan URL, URL memberikan informasi yang tersedia melalui internet dengan cara standar yang mana menentukan elemen internet seperti lokasi server, dokumen, file dan lain – lainnya.

Format umum URL adalah sebagai berikut :

Protokol_transfer :// nama_host / path / nama_file

Contoh : http :// www.amazon.com/ buku / index.html

Internet yang sangat besar merupakan interkoneksi, terdistribusi, tempat yang tidak seragam dan URL menstandarkan dari keseragaman ini.

2.3.5 Domain Name System (DNS)

Dalam dunia internet, kita bisa masuk ke host – host apapun dengan 2 cara. Cara pertama dan paling efisien adalah dengan mengetik alamat internet protokol atau IP address dari host yang ingin kita tuju. Walaupun ini merupakan cara yang paling efisien tetapi bukan cara yang paling praktis.

Cara yang kedua yaitu yang paling praktis adalah mengakses ke host dengan mengetik nama host yang kita tuju, misalnya www.hotmail.com.

Kebanyakan host IP akan mempunyai cara kedua baik IP address berbentuk numeric maupun nama untuk tetap menjaga kestabilan peningkatan dari nama – nama baru yang semakin bertambah di internet maka dibuatlah DNS (Domain Name System).

DNS merupakan database yang terdistribusi yang mengandung nama host dan informasi IP address serta nama semua domain yang ada di internet. Sebuah nama yang merupakan host dari sebuah server ada pada setiap domain. Misalnya .com yang mengandung semua informasi yang berhubungan DNS tentang domain tersebut. Nama – nama domain yang mempunyai level tinggi (top level domain) dapat di lihat pada tabel 2.1:

Tabel 2.1 Macam-macam Domain Name Server

Top level domain	Deskripsi	Contoh
.com	commercial	Microsoft.com Compaq.com
.gov	government	Whitehouse.gov Senate.gov
.mil	military	Army.mil Navy.mil
.edu	education	Umich.edu UMN.edu
.net	network service	InterNIC.net Earthlink.net

2.3.6 Hypertext Markup Language (HTML)

HTML adalah suatu sistem yang digunakan untuk menandai dokumen dengan pembatas informasional yang mengindikasikan bagaimana teks pada dokumen harus direpresentasikan dan bagaimana dokumen dihubungkan satu sama lain. HTML sendiri termasuk turunan dari SGML (Standard Generalized Markup Language) yang merupakan bahasa standar untuk markup.

Dokumen HTML disebut sebagai markup language karena mengandung tanda tertentu yang digunakan untuk menentukan tampilan suatu teks dan tingkat kepentingan dari teks tersebut pada suatu dokumen. HTML juga mendukung multimedia secara penuh, karena dapat menampilkan seluruh komponen multimedia (text, hypertext, gambar, animasi, audio, video).

2.4 Pemrograman Internet

Agar website yang kita tampilkan dapat bersifat dinamis dan informasi yang akan di tampilkan pada internet dapat di tampilkan dengan baik, maka diperlukan

suatu program. Adapun program yang diperlukan untuk pembuatan website dinamis adalah :

2.4.1 PHP

PHP adalah bahasa server-side scripting yang menyatu dengan HTML untuk membuat halaman web yang dinamis. Maksud dari server-side scripting adalah sintaks dan perintah – perintah yang diberikan akan sepenuhnya dijalankan di server tetapi disertakan pada dokumen HTML. Pembuatan web ini merupakan kombinasi antara PHP sendiri sebagai bahasa pemrograman dan HTML sebagai pembangun halaman web.

PHP merupakan software open source (gratis) dan mampu lintas platform, yaitu dapat digunakan dengan sistem operasi dan web server apapun. PHP mampu berjalan di Windows dan beberapa versi Linux. PHP juga dapat di bangun sebagai modul pada web server Apache dan sebagai binary yang dapat berjalan sebagai CGI.

Keunggulan dari server-side antara lain: (Sutarman,2003:109)

- Tidak di perlukan kompabilitas browser atau harus menggunakan browser tertentu, karena serverlah yang akan mengerjakan script PHP. Hasil yang di kirim kembali ke browser umumnya berupa teks atau gambar saja.
- Dapat memanfaatkan sumber aplikasi yang dimiliki oleh server, misalnya koneksi ke database.
- Script tidak dapat dilihat dengan fasilitas view HTML source.

2.4.2 Web Browser

Web browser merupakan aplikasi yang memungkinkan pengguna untuk menjelajahi world wide web untuk mendapatkan informasi dan berkomunikasi. Pengguna hanya mengetahui alamat halaman web yang dimaksud. Kemudian web browser menunggu informasi yang diminta dikirimkan kembali oleh web server, sehingga pengguna dapat melihat informasi tersebut dari web browser. Contoh web

browser : Netscape Communicator, Microsoft Internet Explorer, Opera, dan lain – lain.

Fungsi utama browser adalah :

- Memungkinkan untuk mengambil dan melihat informasi dari komputer server www, gopher, dan FTP di internet, atau media disk yang berisi dokumen HTML.
- Berinteraksi dengan sistem berbasis server.
- Merupakan alat untuk melihat dokumen elektronik
- Untuk melakukan download / upload informasi digital.
- Untuk mengirim dan menerima e-mail.

2.4.3 ApacheWeb Server

Web server adalah suatu program yang terletak pada komputer dengan akses internet, yang merespon permintaan browser untuk suatu URL. Web server memenuhi kebutuhan pengguna dengan mensuplai atau melayani permintaan halaman web.

Jadi, halaman web harus diletakkan dalam web server agar dapat dilihat dari internet. Idealnya, web server harus memiliki koneksi internet yang tidak bisa terputus, sehingga halaman – halaman yang ditangani dapat selalu tersedia.

Apache merupakan pengembangan dari server yang dikeluarkan oleh NSCA yaitu NSCA HTTP pada tahun 1995 dan saat ini merupakan tulang punggung dari World Wide Web (WWW). Apache berfungsi memenuhi permintaan dari client dengan browser seperti Internet Explorer, Mozilla.

2.5 WIKIPEDIA

Wikipedia adalah proyek ensiklopedia multibahasa dalam jaringan yang bebas dan terbuka, yang dijalankan oleh Wikimedia Foundation, sebuah organisasi nirlaba yang berbasis di Amerika Serikat. Nama Wikipedia berasal dari gabungan kata wiki

dan encyclopedia. Wikipedia dirilis pada tahun 2001 oleh Jimmy Wales dan Larry Sanger dan kini merupakan karya referensi paling besar, cepat berkembang, dan populer di Internet. Proyek Wikipedia bertujuan untuk mengumpulkan seluruh ilmu pengetahuan manusia.

Keistimewaan Wikipedia adalah selain menyajikan informasi yang biasa ditemui di dalam sebuah ensiklopedia, Wikipedia juga memuat artikel-artikel yang biasanya ditemukan di dalam almanak, majalah spesialis, dan topik-topik berita yang masih hangat. dan banyak orang yang memakai wikipedia ini untuk menyelesaikan tugas dan pekerjaan rumah.

Wikipedia berawal sebagai proyek sampingan Nupedia, ensiklopedia bebas daring yang artikelnya ditulis oleh para ahli. Larry Sanger, yang mendirikan Nupedia bersama Jimmy Wales, melontarkan ide mengenai ensiklopedia berbasis wiki pada 10 Januari 2001 di milis Nupedia. Kemudian pada 15 Januari 2001, Wikipedia secara resmi diluncurkan.

Berikut ini beberapa karakteristik penting di dalam proyek Wikipedia.

1. Wikipedia adalah ensiklopedia utama.
2. Wikipedia adalah sebuah wiki yang dapat diubah oleh *setiap orang* (kecuali untuk pengguna yang diblok dan pengecualian di halaman terlindung)
3. Wikipedia mengandung isi bebas dan menggunakan copyleft Lisensi dokumentasi bebas GNU

Sejak resmi diluncurkan pada tanggal 15 Januari 2001, Wikipedia Bahasa Inggris telah mengalami perkembangan jumlah artikel yang luar biasa. Tanggal 21 Januari 2003, jumlah artikel yang dimiliki telah mencapai 100.000 buah, kemudian pada tanggal 01 Maret 2006, telah memiliki 1.000.000 artikel. Angka ini terus bertambah menjadi 2.000.000 artikel pada tanggal 09 September 2007 dan menembus angka 3.000.000 artikel pada tanggal 17 Agustus 2009.

Hingga bulan September 2010, proyek Wikipedia yang telah memiliki jumlah artikel lebih dari 1.000.000 hanya ada 3 bahasa, yaitu Bahasa Inggris, Bahasa Jerman

yang telah menembus angka ini pada tanggal 27 Desember 2009 dan Bahasa Perancis yang telah melampauinya pada tanggal 21 September 2010.

Sedangkan Wikipedia Bahasa Indonesia, pertama kali memiliki artikel pada tanggal 30 Mei 2003. Angka ini menjadi 1.000 artikel pada tanggal 16 Maret 2004, kemudian menembus angka 10.000 artikel pada tanggal 31 Mei 2005. Perolehan 100.000 artikel didapat pada tanggal 21 Februari 2009. Pada tanggal 24 September 2010, Wikipedia Bahasa Indonesia telah memiliki lebih dari 132.450 artikel.

2.6 Wordpress

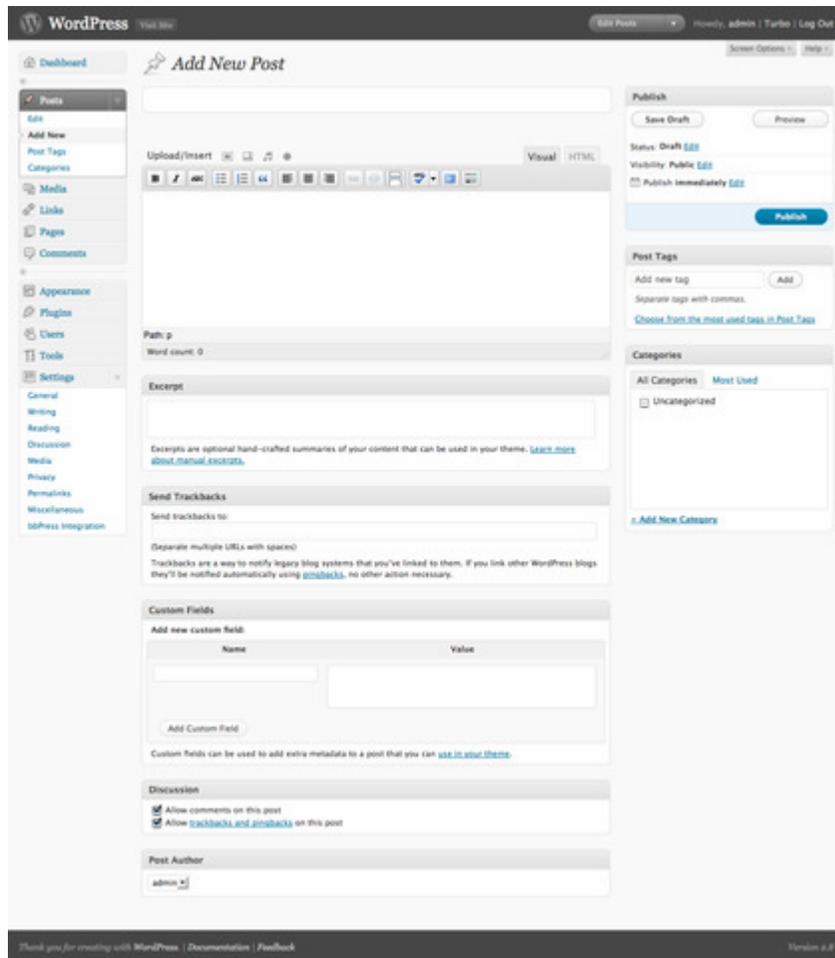
WordPress adalah sebuah aplikasi sumber terbuka (*open source*) yang sangat populer digunakan sebagai mesin blog (*blog engine*). WordPress dibangun dengan bahasa pemrograman PHP dan basis data (*database*) MySQL. PHP dan MySQL, keduanya merupakan perangkat lunak sumber terbuka (*open source software*). Selain sebagai blog, WordPress juga mulai digunakan sebagai sebuah CMS (*Content Management System*) karena kemampuannya untuk dimodifikasi dan disesuaikan dengan kebutuhan penggunanya. WordPress adalah penerus resmi dari b2/cafelog yang dikembangkan oleh Michel Valdrighi. Nama WordPress diusulkan oleh Christine Selleck.

WordPress dimulai saat Matt Mullenweg yang merupakan pengguna aktif dari b2 mengetahui bahwa proses pengembangan b2 dihentikan oleh pemrogramnya (programmer) yang bernama Michel Valdrighi, Matt Mullenweg merasa sayang dan mulai melanjutkan pengembangan b2.

WordPress muncul pertama kali di tahun 2003 hasil kerja keras Matt Mullenweg dengan Mike Little. Yang membuat WordPress makin terkenal, selain karena banyaknya fitur dan tampilan yang menarik, adalah juga karena dukungan komunitas terhadap perangkat lunak sumber terbu Keunggulan dan Fitur WordPress

WordPress memiliki banyak keunggulan dan fitur untuk dunia blog, antara lain :

- Gratis. Untuk mendapatkan perangkat lunak WordPress hanya perlu mengunduh dari situsnya (www.wordpress.org) tanpa dipungut biaya, bahkan untuk blog komersial sekalipun.



Gambar 2.3 Halaman Administrasi WordPress

- Berbasis kode sumber terbuka (*Open Source*). Pengguna dapat melihat dan memperoleh barisan kode-kode penyusun perangkat lunak WordPress tersebut secara bebas, sehingga pengguna tingkat lanjut yang memiliki kemampuan pemrograman dapat bebas melakukan modifikasi, bahkan dapat

mengembangkan sendiri program WordPress tersebut lebih lanjut sesuai keinginan.

- Template atau desain tampilannya mudah dimodifikasi sesuai keinginan pengguna. Sehingga apabila pengguna memiliki pengetahuan HTML yang memadai, maka pengguna tersebut dapat berkreasi membuat template sendiri.
- Pengguna yang tidak mengerti HTML, tentu saja masih dapat memilih ribuan template yang tersedia di internet secara bebas, yang tentu saja gratis (<http://wordpress.org/extend/themes/>).
- Pengoperasiannya mudah. Bagi pemula, tentu hal ini amat membantu.
- Satu blog WordPress, dapat digunakan untuk banyak pengguna (*multi user*). Sehingga WordPress juga sering digunakan untuk blog komunitas. Anggota komunitas tersebut dapat berperan sebagai kontributor.
- Jika pengguna sebelumnya telah mempunyai blog tidak berbayar, misalnya di alamat Blogger / Blogspot, LiveJournal, TypePad, dan beberapa mesin blog lain, pengguna dapat mengimpor isi blog-blog tersebut ke alamat hosting blog pribadi yang menggunakan perangkat lunak WordPress. Dengan demikian pengguna tidak perlu khawatir isi blog yang lama akan menjadi sia-sia setelah menggunakan perangkat lunak WordPress.
- Selain pengguna yang banyak, banyak pula dukungan komunitas (*community support*) untuk WordPress. Banyak forum yang mendukung pengembangan WordPress ini. Contohnya : <http://wordpress.org/support/>, <http://id.forums.wordpress.com/>, dan <http://id.forums.wordpress.com/>.
- Tersedia banyak plugin yang selalu berkembang (<http://wordpress.org/extend/plugins/>), karena para pengembang web di Indonesia juga memberikan kontribusinya kepada WordPress. Plugin WordPress sendiri yaitu sebuah program tambahan yang bisa diintegrasikan dengan WordPress untuk memberikan fungsi-fungsi lain yang belum tersedia pada instalasi standar. Misalnya plugin anti-spam, plugin web counter, dan sebagainya.

- Kemampuan untuk dapat memunculkan XML, XHTML, dan CSS standar.
- Tersedianya struktur permalink yang memungkinkan mesin pencari mengenali struktur blog dengan baik.
- Kemungkinan untuk meningkatkan performa blog dengan ekstensi (<http://wordpress.org/extend/>).
- Mampu mendukung banyak kategori untuk satu artikel. Satu artikel dalam WordPress dapat dikategorisasikan ke dalam beberapa kategori. Misalnya jika kita mempunyai artikel tentang “DKI Jakarta” kita dapat memasukkan artikel tersebut dalam kategori “Tentang Indonesia”, “Propinsi di Indonesia”, atau “Wilayah di Pulau Jawa”. Dengan multi kategorisasi ini pencarian dan pengaksesan informasi menjadi lebih mudah.
- Fasilitas Trackback dan Pingback.
- Fasilitas format teks dan gaya teks. WordPress menyediakan fitur pengelolaan teks yang cukup lengkap. Fitur – fitur format dan gaya teks pada kebanyakan perangkat lunak pengolah kata seperti cetak tebal, cetak miring, rata kanan, rata kiri, tautan tersedia di WordPress.
- Halaman statis (Halaman khusus yang terpisah dari kumpulan tulisan pada blog).
- Mendukung LaTeX.

2.7 Web Scraping

Scraping Web (juga disebut panen Web atau Web ekstraksi data) adalah sebuah perangkat lunak komputer teknik penggalian informasi dari situs web. Biasanya, program perangkat lunak tersebut mensimulasikan eksplorasi manusia dari Web oleh salah satu rendah menerapkan-Hypertext Transfer Protocol (HTTP), atau embedding Web browser tertentu penuh, seperti Internet Explorer (IE) dan Mozilla Web browser.

Generator Konten berkaitan erat dengan pengindeksan Web, yang indeks konten Web menggunakan bot dan merupakan teknik universal yang diadopsi oleh kebanyakan mesin pencari. Sebaliknya, menggores Web lebih memfokuskan pada

transformasi konten Web yang tidak terstruktur, biasanya dalam format HTML, menjadi data terstruktur yang dapat disimpan dan dianalisa dalam database lokal pusat atau spreadsheet. Generator Konten juga terkait dengan otomasi Web, yang mensimulasikan browsing Web manusia menggunakan perangkat lunak komputer. Penggunaan Generator Konten termasuk perbandingan harga online, cuaca data monitoring, deteksi situs berubah, penelitian Web, Mashup konten Web dan Web integrasi data.

Scraping Web adalah proses otomatis mengumpulkan informasi Web. Scraping Web adalah lapangan dengan perkembangan aktif berbagi tujuan bersama dengan visi Web semantik, sebuah inisiatif ambisius yang masih memerlukan terobosan dalam pengolahan teks, pengertian semantik, kecerdasan buatan dan interaksi manusia-komputer. Generator Konten, sebaliknya, nikmat solusi praktis berdasarkan ada teknologi yang sering sekali ad hoc. Oleh karena itu, ada berbagai tingkat otomatisasi yang teknologi web-Scraping yang ada dapat memberikan:

- Manusia copy-dan-paste: Kadang-kadang bahkan teknologi Web-Scraping terbaik tidak dapat menggantikan pemeriksaan manual manusia dan copy-dan-paste, dan kadang-kadang ini mungkin satu-satunya solusi dikerjakan ketika website untuk menggores eksplisit hambatan setup untuk mencegah otomatisasi mesin .
- Teks grepping dan mencocokkan ekspresi reguler: Sebuah pendekatan sederhana namun kuat untuk mengekstrak informasi dari halaman Web dapat didasarkan pada perintah grep UNIX atau pencocokan regular expression fasilitas bahasa pemrograman (misalnya Perl atau Python).
- HTTP pemrograman: halaman web statis dan dinamis dapat diambil dengan mengirimkan permintaan HTTP ke server Web remote menggunakan pemrograman socket.
- parsing DOM: Dengan memasukkan browser web penuh, seperti Internet Explorer atau Mozilla kontrol web browser, program dapat mengambil isi yang dinamis yang dihasilkan oleh script sisi klien. Web browser ini kontrol

juga mengurai halaman Web ke dalam pohon DOM, berdasarkan program yang dapat mengambil bagian dari halaman Web.

- HTML parsing: Beberapa data semi-terstruktur bahasa query, seperti bahasa query XML (XQL) dan bahasa query hiper-teks (HTQL), dapat digunakan untuk mengurai halaman HTML dan untuk mengambil dan mengubah konten Web.

Ada banyak software Web-Scraping tersedia yang dapat digunakan untuk menyesuaikan solusi Web-Scraping. Perangkat lunak ini mungkin mencoba untuk secara otomatis mengenali struktur data dari sebuah halaman atau menyediakan antarmuka rekaman Web yang menghilangkan kebutuhan menulis kode secara manual Web-Scraping, atau beberapa fungsi script yang dapat digunakan untuk mengekstrak dan mengubah konten Web, dan interface database yang dapat menyimpan data tergores dalam database lokal.

Generator Konten mungkin terhadap persyaratan penggunaan beberapa website. Penegakan ketentuan ini tidak jelas. Sementara duplikasi langsung ekspresi asli akan dalam banyak kasus tidak sah, di Amerika Serikat, pengadilan memerintah di Feist Publikasi ay Telepon Pedesaan Layanan yang duplikasi fakta-fakta yang diijinkan. pengadilan AS telah mengakui bahwa pengguna "pencakar" atau "robot" bisa bertanggung jawab karena melakukan pelanggaran untuk barang bergerak, yang melibatkan sistem komputer itu sendiri yang dianggap sebagai harta pribadi yang di atasnya pengguna scraper adalah masuk tanpa izin.

2.8 Web Konten

Elemen-elemen seperti teks, forms, images (GIFs, JPEGs, Portable Network Graphics) dan video dapat diletakkan didalam halaman menggunakan tag-tag HTML/XHTML/XML. Browser terkadang juga memerlukan Plug-ins seperti Adobe Flash, QuickTime, Java, dan sebagainya untuk menampilkan beberapa media yang diletakkan didalam halaman web menggunakan tag-tag HTML/XHTML.

Halaman web dan situs web dapat berupa halaman statis, atau dapat diprogram secara dinamis sehingga menghasilkan halaman web dengan konten atau tampilan visual yang diinginkan, tergantung pada berbagai faktor, seperti masukan dari pengguna akhir, masukan dari Webmaster, atau perubahan dalam lingkungan komputasi (seperti situs yang terkait dengan database yang telah diubah).

2.9 XML

XML (Extensible Markup Language) adalah bahasa markup untuk keperluan umum yang disarankan oleh W3C untuk membuat dokumen markup keperluan pertukaran data antar sistem yang beraneka ragam. XML merupakan kelanjutan dari HTML (HyperText Markup Language) yang merupakan bahasa standar untuk melacak Internet.

XML didesain untuk mampu menyimpan data secara ringkas dan mudah diatur. Kata kunci utama XML adalah data (jamak dari datum) yang jika diolah bisa memberikan informasi.

XML menyediakan suatu cara terstandarisasi namun bisa dimodifikasi untuk menggambarkan isi dari dokumen. Dengan sendirinya, XML dapat digunakan untuk menggambarkan sembarang view database, tetapi dengan suatu cara yang standar. XML memiliki tiga tipe file :

- XML, merupakan standar format dari struktur berkas (file).
- XSL, merupakan standar untuk memodifikasi data yang diimpor atau diekspor.
- XSD, merupakan standar yang mendefinisikan struktur database dalam XML.

2.10 RSS

RSS adalah sebuah file berformat XML untuk sindikasi yang telah digunakan (diantaranya dan kebanyakan) situs web berita dan weblog. Singkatan ini biasanya mengarah ke beberapa :

- Rich Site Summary (RSS 0.91)

- RDF Site Summary (RSS 0.9 and 1.0)
- Really Simple Syndication (RSS 2.0)

Teknologi yang dibangun dengan RSS memungkinkan kita untuk berlangganan kepada situs web yang menyediakan umpan web (feed) RSS, biasanya situs web yang isinya selalu diganti secara reguler. Untuk memanfaatkan teknologi ini kita membutuhkan layanan pengumpul. Pengumpul bisa dibayangkan sebagai kotak surat pribadi. Kita kemudian dapat mendaftar ke situs yang ingin kita tahu perubahannya. Namun, berbeda dengan langganan koran atau majalah, untuk berlangganan RSS tidak diperlukan biaya, gratis. Tapi, kita biasanya hanya mendapatkan satu baris atau sebuah pengantar dari isi situs berikut alamat terkait untuk membaca isi lengkap artikelnya.

RSS digunakan secara luas oleh komunitas weblog untuk menyebar ringkasan tulisan terbaru di jurnal, kadang-kadang juga menyertakan artikel lengkap dan bahkan gambar dan suara. Sekitar 2000, penggunaan RSS meluas di berbagai penerbitan berita, termasuk Reuters, CNN, dan BBC. RSS digunakan pada hampir semua situs berita atau weblog, dengan berbagai tujuan termasuk: pemasaran, press release, laporan reguler produk, atau aktivitas lain yang membutuhkan pemberitahuan periodik dan tentunya publikasi.

Sebuah program komputer yang dikenal sebagai pembaca umpan (feed reader) bertindak sebagai pengumpul. Program ini mengecek situs yang menyediakan RSS dan menampilkan berbagai artikel baru yang ditemukan. Tenggang waktu dan siklus pengumpulan RSS biasanya dapat diatur oleh penggunanya. Program pengumpul dapat berupa program komputer atau sebuah layanan web yang tersedia online.

Program pengumpul RSS di komputer biasanya berupa aplikasi (software) sendiri yang harus dipasang di komputer sebelum dapat digunakan. Program ini tersedia untuk berbagai jenis sistem operasi dengan harga bervariasi. Ada juga program pengumpul RSS yang gratis.

Program pengumpul di web tidak memerlukan pemasangan dan pengaturan, kita dapat melihat dan mengecek RSS kita di mana saja, asal ada browser dan koneksi internet. Beberapa layanan pengumpul RSS juga menyediakan penggabungan dan juga pencarian.

Sebuah web yang menyediakan umpan RSS biasanya ada link dengan sebuah tombol bertuliskan XML (XML iconic button) atau RSS (RSS iconic button). Namun disarankan menggunakan RSS daripada XML agar tidak membuat kerancuan (sebab setiap RSS adalah XML, tapi tidak tiap XML itu RSS).

Seperti ditulis di atas, ada beberapa versi RSS, dan bahkan ada kelompok membuat format baru Atom yang juga telah diadopsi oleh banyak situs. Tapi pemakai RSS tidak perlu bingung, para pembuat program pengumpul sindikasi biasanya bisa mengenali kedua format tersebut, bahkan masih menyediakan pembaca untuk format yang lama.¹

BAB III

PERANCANGAN SISTEM

3.1 Spesifikasi Kebutuhan Perangkat Lunak Aplikasi.

3.1.1 Ruang lingkup produk

Sistem ini adalah Rekayasa Perangkat Lunak Komputer berbasis web yang bertujuan untuk menghasilkan konten secara otomatis dari suatu kata kunci. Hal-hal yang diharapkan oleh pengguna agar dapat diwujudkan dalam sistem ini diantaranya adalah hal-hal sebagai berikut :

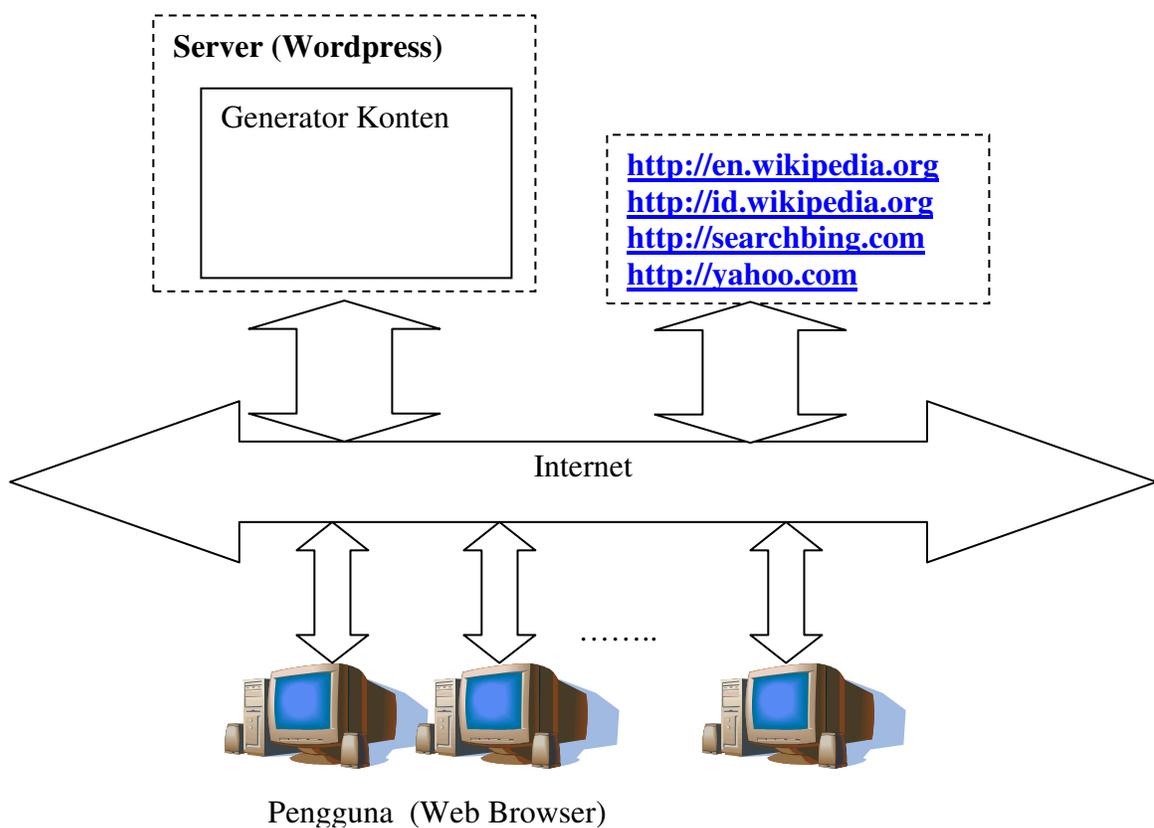
- Sistem dapat secara otomatis menghasilkan suatu konten yang nantinya akan ditampilkan di halaman web, berdasarkan pada suatu kata kunci.
- Untuk mempersingkat penelitian sehingga fokus pada penelitian generator konten maka digunakan CMS wordpress untuk menangani proses manajemen website.
- Sistem dapat secara otomatis menampilkan konten hasil generator dalam bentuk halaman web site atau memasukan hasil tersebut ke dalam artikel di wordpress, baik untuk judul, isi, rangkuman maupun kata kunci / tag.
- Sistem ini dapat diintegrasikan dengan plugin wordpress yang lain sehingga mempermudah instalasi.
- Sistem dapat di install di web server manapun yang mendukung instalasi CMS wordpress.

Dalam pengembangan aplikasi ini diharapkan dapat memberikan manfaat sbb :

- Sistem dapat mengenerasi konten utama dengan bantuan mesin pencari.
- Sistem dapat membuat konten secara otomatis sesuai kebutuhan pada website dengan basis wordpress.

3.1.2 Perspektif produk

Aplikasi yang dibangun ditanam pada web server yang terkoneksi dengan jaringan internet. Aplikasi berjalan menggunakan service http dengan format transaksi data html, sehingga dapat dibuka menggunakan terminal yang terkoneksi ke jaringan komputer dan mampu / mempunyai browser WEB. Pengguna dapat melihat konten yang telah dihasilkan dalam bentuk artikel dalam wordpress.



Gambar 3.1 Perspektif produk

Pada gambar 3.1 diperlihatkan Komputer Server berfungsi sebagai web server yang terinstall wordpress. Web server akan mencoba mengambil dokumen dengan mencari kata kunci di situs <http://en.wikipedia.org> jika di temukan halaman akan di scrap, jika tidak cari kata kunci di situs <http://id.wikipedia.org> demikian seterusnya

dan terakhir jika tidak ditemukan dokumen dengan kata kunci tersebut akan di lakukan pencarian di mesin pencari <http://searchbing.com> dan <http://yahoo.com>

3.1.3 Fungsi-fungsi Produk

Produk Aplikasi dibangun dengan antarmuka web, sehingga semua fungsi dapat langsung diakses dari halaman aktif manapun. Walaupun demikian tidak ada menu atau fungsi yang secara eksplisit merujuk ke sistem generator konten

3.2 Aturan Bisnis Aplikasi

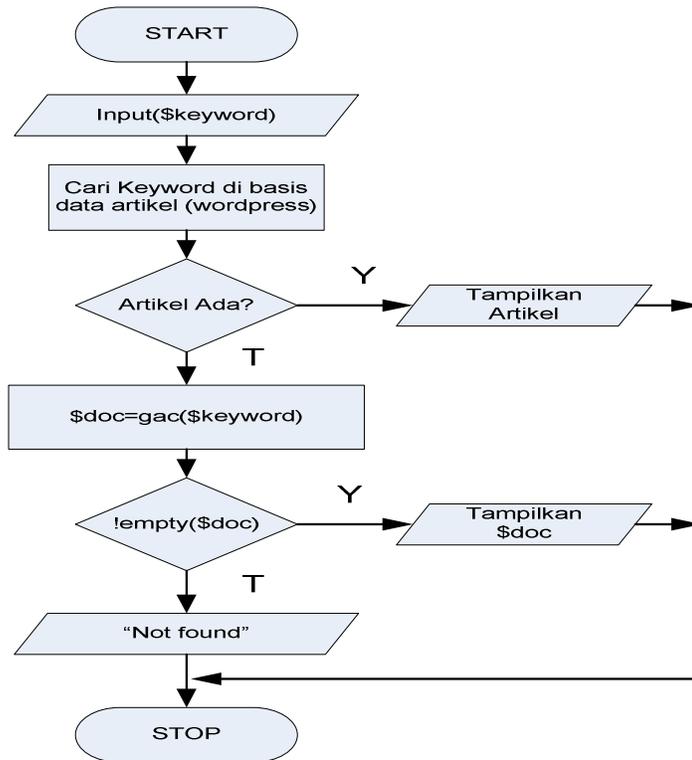
Aturan bisnis digunakan sebagai acuan kemampuan dari aplikasi yang akan dibuat. aturan bisnis untuk generator konten adalah sbb:

- Sistem berupa fungsi menggunakan aplikasi berbasis web digunakan untuk melakukan proses generator konten.
- Menggunakan CMS wordpress sebagai manajemen kontennya
- Terimplementasi di dalam wordpress sebagai plugin.

3.3 Diagram Alir Aplikasi

3.3.1 Diagram Alir Pencarian Artikel pada Wordpress

Aliran Proses pencarian artikel pada website berbasis wordpress dengan Generator Konten terinstall diperlihatkan pada gambar 3.2. Variabel keyword dimasukan sebagai dasar pencarian artikel, jika artikel di temukan di dalam basis data maka artikel akan di tampilkan.



Gambar 3.2 Diagram aliran proses menu utama aplikasi

Jika artikel tidak ditemukan maka dilakukan proses generasi konten untuk kata kunci yang dimasukkan. Jika generator konten berhasil menemukan artikel maka akan ditampilkan ke halaman web apabila tidak maka akan ditampilkan error Not Found.

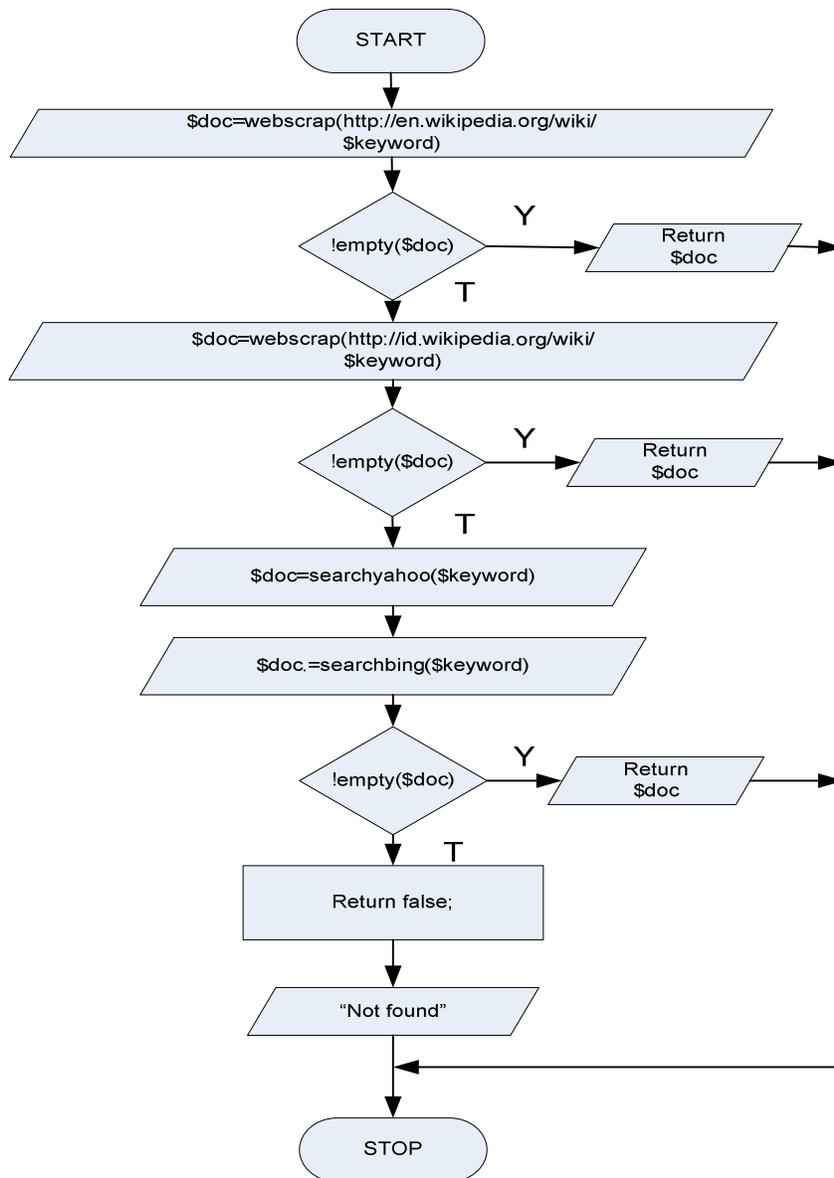
3.3.2 Diagram Alir Fungsi Generator Konten

Aliran proses generator konten ditampilkan pada gambar 3.3. Fungsi akan memulai dengan mengambil halaman web pada url [http://en.wikipedia.org/wiki/\\$keyword](http://en.wikipedia.org/wiki/$keyword). Kemudian dilakukan proses scraping pada halaman tersebut jika di temukan sistem berhasil menghasilkan dokumen yang berhubungan dengan keyword dan mengembalikannya dalam bentuk hasil fungsi.

Jika gagal menemukan dokumen dari [http://en.wikipedia.org/wiki/\\$keyword](http://en.wikipedia.org/wiki/$keyword) maka dilakukan pengambilan dokumen di situs selanjutnya yaitu [http://id.wikipedia.org/wiki/\\$keyword](http://id.wikipedia.org/wiki/$keyword). Kemudian dilakukan proses scraping pada

halaman tersebut jika di temukan sistem berhasil menghasilkan dokumen yang berhubungan dengan keyword dan mengembalikannya dalam bentuk hasil fungsi.

Jika gagal menemukan dokumen dari [http://id.wikipedia.org/wiki/\\$keyword](http://id.wikipedia.org/wiki/$keyword) maka dilakukan pengambilan dokumen dari mesin pencari dengan hasil berupa dokumen berformat xml / rss.



Gambar 3.3 Diagram Aliran Proses Fungsi Generator Konten

BAB IV

IMPLEMENTASI SISTEM

4.1 Implementasi Fungsi Generator Konten

4.1.1 en.wikipedia.org scraping

Pada gambar 3.3 diperlihatkan alur proses dari fungsi Generator Konten, dimana fungsi terlebih dahulu melakukan pengambilan halaman web di website en.wikipedia.org. Proses pengambilan halaman atau biasa disebut dengan fetching dapat dilakukan dengan perintah PHP `file_get_content`, tetapi pada kasus en.wikipedia.org metode `file_get_content` tidak dapat dilakukan dikarenakan adanya leech detection di webserver wikipedia.

Metode lainnya untuk mengambil konten web adalah CURL, metode ini dapat mengirim informasi lengkap dan detil layaknya sebuah web browser sehingga web server menganggap permintaan dilakukan oleh seorang pengguna dengan menggunakan web browser. Pada penelitian ini tidak secara langsung menggunakan perintah CURL untuk mengambil halaman web, tetapi menggunakan pustaka snooply. Pustaka snooply berisi fungsi-fungsi PHP untuk melakukan fetch di suatu halaman web dengan perintah dasar CURL.

Pada program 4.1 diperlihatkan potongan kode php yang melakukan fetch, sebelum melakukan fetch terlebih dahulu dilakukan deklarasi obyek `$snoopy` dari kelas Snooply, kemudian user agent string di set opera sehingga web server akan mengenali fungsi sebagai browser opera. Sebelum keyword di gunakan terlebih dahulu di filter menjadi huruf kecil semua dan mengganti karakter `+` menjadi karakter spasi.

Program 4.1 Implementasi Operasi Fetch Halaman WEB en.wikipedia.org

```
$snoopy = new Snoopy;  
$snoopy->agent='opera';  
$keyword=strtolower($keyword);  
$keyw=str_replace(" ", " ", $keyword);
```

```

$keyw=str_replace(" ", "+", $keyword);
$uristr="http://en.wikipedia.org/w/index.php?title=Special%3ASearch&
search=$keyword";
$snoopy->fetch($uristr);
$striklan= $snoopy->results;

```

Pada program 4.2 . Konten utama dibatasi dari header oleh teks ”<!-- bodytext -->” dan dibatasi oleh footer oleh ”<!-- /bodytext -->”, sehingga pada program 4.3 digunakan perintah regular ekspresi dengan mendeteksi pola <!-- bodytext -->(.*<!-- /bodytext -->.

Program 4.2 Pola Konten utama halaman dokumen wikipedia

```

...bagian header
<!-- bodytext -->
.
.
.
Konten Utama
.
.
.
<!-- /bodytext -->
....bagian footer

```

Pada program 4.3 diperlihatkan operasi Generator Konten dengan menggunakan metode regular ekspresi. Konten utama halaman web hasil en.wikipedia.com diapit oleh bagian header dan footer seperti terlihat

Program 4.3 Perintah ekspresi regular untuk memisahkan konten utama wikipedia

```

$str= $snoopy->results;
ereg("<!-- bodytext -->(.*<!-- /bodytext -->", $str , $match);
$str=$match[0];

```

4.1.2 id.wikipedia.org scraping

Pada program 4.4 diperlihatkan potongan kode php yang melakukan fetch, sebelum melakukan fetch terlebih dahulu dilakukan deklarasi obyek \$snoopy dari kelas Snoopy, kemudian user agent string di set opera sehingga web server akan mengenali fungsi sebagai browser opera. Sebelum keyword di gunakan terlebih dahulu di filter menjadi huruf kecil semua dan mengganti karakter + menjadi karakter spasi.

Program 4.4 Implementasi Operasi Fetch Halaman WEB id.wikipedia.org

```
$snoopy = new Snoopy;
$snoopy->agent='opera';
$keyword=strtolower($keyword);
$keyw=str_replace(" ", " ", $keyword);
$keyw=str_replace(" ", "+", $keyword);
$uristr="http://id.wikipedia.org/w/index.php?title=Special%3ASearch&
search=$keyword";
$snoopy->fetch($uristr);
$striklan= $snoopy->results;
```

Pada program 4.5 . Konten utama dibatasi dari header oleh teks ”<!-- bodytext -->” dan dibatasi oleh footer oleh ”<!-- /bodytext -->”, sehingga pada program 4.6 digunakan perintah regular ekspresi dengan mendeteksi pola <!-- bodytext -->(.*?)<!-- /bodytext -->.

Program 4.5 Pola Konten utama halaman dokumen wikipedia

```
...bagian header
<!-- bodytext -->
.
.
.
Konten Utama
.
.
.

<!-- /bodytext -->
....bagian footer
```

Pada program 4.6 diperlihatkan operasi Generator Konten dengan menggunakan metode regular ekspresi. Konten utama halaman web hasil en.wikipedia.com diapit oleh bagian header dan footer seperti terlihat

Program 4.6 Perintah ekspresi regular untuk memisahkan konten utama wikipedia

```
$str= $snoopy->results;  
ereg("<!-- bodytext -->(.*?)<!-- /bodytext -->", $str , $match);  
$str=$match[0];
```

4.1.3 Search Engine scraping

Pada program 4.7 diperlihatkan potongan kode php yang melakukan fetch kata kunci dari API yahoo search. Yahoo menyediakan API untuk hasil pencariannya pada alamat url <http://api.search.yahoo.com/WebSearchService/rss/webSearch.xml>, hasil dari API Yahoo search adalah dokumen berupa XML dengan standar rss feed.

Pada program 4.8 diperlihatkan potongan kode php yang melakukan fetch kata kunci dari API bing search. Bing menyediakan API untuk hasil pencariannya pada alamat <http://www.bing.com/search?go=&form=QBLH&filt=all&format=rss>, hasil dari API bing search adalah dokumen berupa XML dengan standar rss feed.

Program 4.7 program untuk mengambil feed dari yahoo search API

```
$snoopy = new Snoopy;  
$snoopy->agent='opera';  
$keyword=strtolower($keyword);  
$keyw=str_replace(" ", "$keyword");  
$keyw=str_replace(" ", "+", $keyw);  
$uristr="http://api.search.yahoo.com/WebSearchService/rss/webSearch.xml?appid=yahoosearchwebrss&query=$keyword";  
$snoopy->fetch($uristr);  
$str= $snoopy->results;
```

Program 4.8 program untuk mengambil feed dari bing search API

```
$snoopy = new Snoopy;
```

```

$snoopy->agent='opera';
$keyword=strtolower($keyword);
$keyw=str_replace(" ", "$keyword");
$keyw=str_replace(" ", "+", $keyw);
$uristr="http://www.bing.com/search?q=$keyword&go=&form=QBLH&filt=all&format=rss";
$snoopy->fetch($uristr);
$str= $snoopy->results;

```

4.1.4 Implementasi Wordpress Plugin

Setelah fungsi selesai ditulis maka untuk mempermudah penggunaan dan integrasi dengan wordpress maka struktur program fungsi Generator Konten di rubah menjadi struktur plugin pada wordpress.

Struktur program plugin pada wordpress mengharuskan ditambahkan header remark yang berfungsi untuk memuat informasi seputar plugin tersebut, seperti terlihat pada potongan program 4.4.

Program 4.4 Header plugin wordpress Generator Konten

```

/*
Plugin Name: Generator Konten
Plugin URI: http://unisbank.ac.id
Description: menghasilkan konten secara otomatis
Version: 1.0
Author: FTI
Author URI: http://unisbank.ac.id
*/

```

4.2 Instalasi CMS

4.2.1 Instalasi Wordpress

Proses instalasi wordpress dapat dilakukan dengan bantuan utilitas fantastico pada website dengan dukungan Cpanel, tetapi apabila tidak terdapat utilitas fantastico source code wordpress dapat di download pada url

<http://wordpress.org/latest.zip>. Source code diekstrak pada direktori sesuai dengan kebutuhan.

Instalasi dilanjutkan dengan menjalankan script [http://\[namadomain\]/wp-admin/install.php](http://[namadomain]/wp-admin/install.php) sehingga akan didapat tampilan seperti pada gambar 4.1. Setelah semua proses instalasi di selesaikan maka situs telah terinstal worpress dan siap digunakan.



The image shows the initial database configuration screen for WordPress. At the top, there is the WordPress logo and the text "WORDPRESS". Below this, a message reads: "Below you should enter your database connection details. If you're not sure about these, contact your host." The form contains five input fields, each with a label and a description:

Database Name	<input type="text" value="wordpress"/>	The name of the database you want to run WP in.
User Name	<input type="text" value="username"/>	Your MySQL username
Password	<input type="text" value="password"/>	...and MySQL password.
Database Host	<input type="text" value="localhost"/>	You should be able to get this info from your web host, if localhost does not work.
Table Prefix	<input type="text" value="wp_"/>	If you want to run multiple WordPress installations in a single database, change this.

At the bottom left of the form is a "Submit" button.

Gambar 4.1 Tampilan awal instalasi wordpress

4.2.2 Instalasi Plugin Generator Konten

Plugin pada wordpress diinstall melalui menu administrator di url [http://\[namadomain\]/wp-admin/](http://[namadomain]/wp-admin/) setelah terlebih dahulu memasukan username dan password untuk administrator. Plugin dipasang pada menu plugins->add new , plugin dapat di upload ataupun langsung di download dari repository wordpress.

Pada penelitian ini plugin tidak terdapat di repository wordpress, sehingga digunakan menu upload untuk menambahkan plugin. Plugin akan aktif setelah

diaktifkan dengan meng klik url activate dibawah nama plugin. Pada gambar 4.2 diperlihatkan tampilan layar daftar plugin yang terpasang di wordpress.



Gambar 4.2 Tampilan layar daftar plugin wordpress

4.2.3 Aktifasi Fungsi Generator Konten

Setelah plugin terpasang tidak serta merta fungsi Generator Konten telah digunakan pada wordpress. Untuk dapat berfungsi seperti proses yang digambarkan pada gambar 3.2 maka dibutuhkan modifikasi pada bagian template / theme yang digunakan. Pada gambar 3.2 diperlihatkan bahwa proses yang di sisipi fungsi Generator Konten adalah proses pencarian.

Pada wordpress proses pencarian melibatkan file search.php pada theme yang aktif untuk menghasilkan halaman konten. File search.php pada bagian tampilan "Not Found" di ganti dengan pemanggilan fungsi Generator Konten sehingga sebelum menampilkan not found terlebih dahulu dilakukan pencarian pada situs en.wikipedia.org, seperti terlihat pada program 4.5

Program 4.5 Perubahan pada search.php

```
<?php if ( have_posts() ) { ?>
<h1 class="page-title"><?php printf( __( 'Search Results for: %s',
'twentyten' ), '<span>' . get_search_query() . '</span>' ); ?></h1>
<?php } else {
global $s;
include "wp-includes/class-snoopy.php";
if((function_exists('generatorkonten') and (!empty($s))) $doc=
generatorkonten ($s);
if(strlen($doc)>500) echo "<div id='post-0' class='post no-results
not-found'>
<h2 class='entry-title'>$s</h2>
<div class='entry-content'>
<p>$doc</p>
</div><!-- .entry-content -->
</div><!-- #post-0 -->";
```

```

else { ?>
  <div id="post-0" class="post no-results not-found">
  <h2 class="entry-title"><?php _e( 'Nothing Found', 'twentyten' );
?></h2>
      <div class="entry-content">
        <p><?php _e( 'Sorry, but nothing
matched your search criteria. Please try again with some different
keywords.', 'twentyten' ); ?></p>
          <?php get_search_form(); ?>
        </div><!-- .entry-content -->
      </div><!-- #post-0 -->

<?php } } ?>

```

4.3 Pengujian

4.3.1 Google.com trend

Pengujian pertama dilakukan dengan memasukan kata kunci-kata kunci yang terdapat dalam trend pencarian google.com. Seperti pada tabel 4.1 adalah daftar kata-kata kunci pencarian yang populer pada tanggal 8 Februari 2011.

Tabel 4.1 Hasil pengujian dengan kata kunci dari trend google.com

No	Kata Kunci	Hasil
1	aol huffington post	Search Result
2	sting wwe	Search Result
3	sean payton	Dokumen
4	gizmodo	Dokumen
5	mike munchak	Dokumen
6	chicago code	Dokumen
7	reyes	Search Result
8	groupon	Dokumen
9	aguilera	Search Result
10	national anthem	Dokumen

Kata-kata kunci dari google trend tersebut sebagai masukan pada sistem Generator Konten. Kemudian pada kolom hasil di perlihatkan apakah Generator Konten berhasil memberikan keluaran berupa halaman web. Pada tabel juga diperlihatkan apakah hasil dari Generator Konten menghasilkan dokumen atau halaman hasil cari. Dari tabel 4.1 diperlihatkan 6 dari 10 kata kunci populer pada tanggal 8 february 2011 dapat diproduksi halaman situsnya dengan menggunakan teknik Generator Konten.

4.3.2 Google.co.uk trend

Pengujian kedua dilakukan dengan memasukan kata kunci-kata kunci yang terdapat dalam trend pencarian google.co.uk. Seperti pada tabel 4.2 adalah daftar kata-kata kunci pencarian yang populer pada tanggal 8 Februari 2011.

Tabel 4.2 Hasil pengujian dengan kata kunci dari trend google.com

No	Kata Kunci	Hasil
1	paul getty	Dokumen
2	sting wwe	Search Result
3	gma news	Dokumen
4	sean payton	Dokumen
5	huffington post aol	Search Result
6	darth vader commercial	Search Result
7	mike munchak	Dokumen
8	reyes	Search Result
9	chicago code	Dokumen
10	groupon	Dokumen

Kata-kata kunci dari google trend tersebut sebagai masukan pada sistem Generator Konten. Kemudian pada kolom hasil di perlihatkan apakah Generator

Konten berhasil memberikan keluaran berupa halaman web. Pada tabel juga diperlihatkan apakah hasil dari Generator Konten menghasilkan dokumen atau halaman hasil cari. Dari tabel 4.2 diperlihatkan 6 dari 10 kata kunci populer pada tanggal 8 februari 2011 dapat diproduksi halaman situsnya dengan menggunakan teknik Generator Konten.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian dari bab sebelumnya maka dapat disimpulkan beberapa hal sebagai berikut :

- Teknik Generator Konten dapat di implementasikan on the fly sehingga hanya dilakukan saat seorang pengguna membutuhkan dan dokumen yang dimaksud tidak terdapat didalam website
- Dari tren pencarian google.com dan google.co.uk pada tanggal 8 februari 2011, dihasilkan 60% halaman dokumen yang berhubungan dengan kata-kata kunci pada trend pencarian google.

5.2 Saran

Berdasarkan hasil penelitian yang diperoleh dapat disarankan beberapa hal sbb:

- Untuk dapat di rangking dengan bagus oleh mesin pencari dibutuhkan konten unik, karena dengan teknik Generator Konten, konten yang dihasilkan tidak unik maka teknik ini website tidak dapat menjadi urutan awal pada hasil mesin pencari.
- Untuk menghasilkan artikel yang unik perlu digabungkan dengan teknik text summarization ataupun random sentences producer dan perbaiki di meta tag.
- Masalah hukum dan HAKI menjadi masalah utama pada situs yang menggunakan teknik Generator Konten ini. Tetapi dapat di tambahkan pada footer sumber asli dari dokumen tersebut.

DAFTAR PUSTAKA

<http://bing.com/> diakses tanggal 2 januari 2011

<http://en.wikipedia.org/wiki/wikipedia> diakses tanggal 2 januari 2011

<http://en.wikipedia.org/> diakses tanggal 2 januari 2011

<http://google.com/trend> diakses tanggal 8 Februari 2011

<http://google.co.uk/trend> diakses tanggal 8 Februari 2011

<http://wordpress.org> diakses tanggal 2 januari 2011

<http://id.wikipedia.org/> diakses tanggal 2 januari 2011

<http://yahoo.com/> diakses tanggal 2 januari 2011

JISC Briefing Paper, 2006, *Text mining*, JISC, Inggris

Murhadin, Endy, 2003, *PHP Programming Fundamental dan MySQL Fundamental*,
<http://ikc.cbn.net.id/umum/andy-php.php>

Nugroho, Bunafit, 2004, *PHP & MySQL Dengan Editor Dreamweaver MX*, Andi,
Yogyakarta

Pressman R, 1997, *Software Engineering*, Mc Graw Hill, USA

Prothelon's, 2005, *Web Desain, PHP Programming, Language Learning*,
<http://prothelon.com/mambo/tutorial>

LAMPIRAN

Kode Sumber Plugin Generator Konten

```
<?php
/*
Plugin Name: Generator Konten
Plugin URI: http://unisbank.ac.id
Description: menghasilkan konten secara otomatis
Version: 1.0
Author: FTI
Author URI: http://unisbank.ac.id
*/

function enwiki($keyw){
$snoopy = new Snoopy;
$snoopy ->agent='opera';
$keyw=strtolower($keyw);
$keyw=str_replace(" ", " ", $keyw);
$keyw=str_replace(" ", " ", $keyw);
$keyw=str_replace(" ", "+", $keyw);

$uristr="http://en.wikipedia.org/w/index.php?title=Special%3ASearch&search=$keyw";
$snoopy->fetch($uristr);
$str= $snoopy->results;
ereg("<!-- bodytext -->(.*?)<!-- /bodytext -->", $str, $match);
$str= $match[0];
ereg("<table class=\"infobox vcard\"(.*)<div class=\"printfooter\">", $stri, $match);
$str= $match[0];
$str=rmtag( $str, 'input');
$str=rmtag( $str, 'sup');
$str=rmtag( $str, 'a');
$str=str_replace("<span class=\"editsection\">[edit ]</span> ", "", $str);
$str=str_replace("//<![CDATA[", "", $str);
$str=str_replace("if (window.showTocToggle) { var tocShowText = \"show\"; var tocHideText = \"hide\"; showTocToggle(); } ", "", $str);
$str=str_replace("//]]>", "", $str);
$str=str_replace("]]", "", $str);
return "<p text-align='justify'> $str</p>";
```

```

}

function idwiki($keyw){
$snoopy = new Snoopy;
  $snoopy ->agent='opera';
  $keyw=strtolower($keyw);
  $keyw=str_replace(" ", " ", $keyw);
  $keyw=str_replace(" ", " ", $keyw);
  $keyw=str_replace(" ", "+", $keyw);

$uristr="http://id.wikipedia.org/w/index.php?title=Special%3ASearch&search=$keyw";
  $snoopy->fetch($uristr);
  $str= $snoopy->results;
  ereg("<!-- bodytext -->(.*?)<!-- /bodytext -->", $str, $match);
  $str= $match[0];
  ereg("<table class=\"infobox vcard\"(.*)<div class=\"printfooter\">", $stri, $match);
  $str= $match[0];
  $str=rmtag( $str, 'input');
  $str=rmtag( $str, 'sup');
  $str=rmtag( $str, 'a');
  $str=str_replace("<span class=\"editsection\">[edit ]</span> ", "", $str);
  $str=str_replace("//<![CDATA[", "", $str);
  $str=str_replace("if (window.showTocToggle) { var tocShowText = \"show\"; var tocHideText = \"hide\"; showTocToggle(); } ", "", $str);
  $str=str_replace("//]]>", "", $str);
  $str=str_replace("]]", "", $str);
  return "<p text-align='justify'> $str</p>";
}

function generatorkonten($keyw){

  $doc=enwiki($keyw);
  $pos=strpos($doc, ":Search");
  if (!$pos === false) $doc="";
  if (empty($doc)){
    $doc=idwiki($keyw);
    $pos=strpos($doc, ":Search");
  }
}

```

```

    if (!( $pos === false )) $doc="";
}
if (empty($doc)){
    $doc=getrss('http://www.bing.com/search?q='
clearQuery($keyw)
'&go=&form=QBLH&filt=all&format=rss');

$doc.=getrss('http://api.search.yahoo.com/WebSearchService/rss/webSearch.xml?appid=yahoosearchwebrss&query='
clearQuery($keyw) . '&adult_ok=1');
}
return $doc;
}

function clearQuery($query) {
    return str_replace(' ', '+',
wp_specialchars(stripslashes($query)));
}

function getrss($url){
    $out="";
    include_once(ABSPATH . WPINC . '/rss.php');
    $feed = fetch_rss($url);
    $items = array_slice($feed->items, 0, 5);
    if (!empty($items)){
        foreach ($items as $item) {
            $out.= "<h2><a href='$item[link]' rel='nofollow'
target='_blank'>$item[title]</a></h2>";
            $out.= "<p>$item[description]</p>";
        }
    }
    return $out;
}
?>

```

PERSONALIA PENELITIAN

1. Ketua Peneliti

- a. Nama : Dwi Budi Santoso, S.Kom
- b. Golongan/Pangkat : IIIA/ Ahli Madya
- c. Fungsional : Asisten Ahli
- d. Bidang Keahlian : Ilmu Komputer
- e. Pekerjaan : Tenaga Edukatif
- f. Institusi : Universitas Stikubank Semarang
- g. Waktu yang tersedia : 10 Jam per minggu

2. Anggota Peneliti I

- a. Nama : Mardi Siswo Utomo, S.Kom, M.Cs
- b. Golongan/Pangkat : IIIA / Penata Muda
- c. Fungsional : Lektor
- d. Bidang Keahlian : Ilmu Komputer
- e. Pekerjaan : Tenaga Edukatif
- f. Institusi : Universitas Stikubank Semarang
- g. Waktu yang tersedia : 10 Jam per minggu

3. Anggota Peneliti II

- a. Nama : Herny Februariyanti , ST , M.Cs
- b. Golongan/Pangkat : IIIA / Penata Muda
- c. Fungsional : Lektor
- d. Bidang Keahlian : Ilmu Komputer
- e. Pekerjaan : Tenaga Edukatif
- f. Institusi : Universitas Stikubank Semarang
- g. Waktu yang tersedia : 10 Jam per minggu

4. Anggota Peneliti III

- a. Nama : Fitri Hudaini
- b. Bidang Keahlian : Ilmu Komputer
- c. Pekerjaan : Mahasiswa
- d. Institusi : Universitas Stikubank Semarang
- e. Waktu yang tersedia : 10 Jam per minggu

5. Anggota Peneliti IV

- a. Nama : Andry Harmanto
- b. Bidang Keahlian : Ilmu Komputer
- c. Pekerjaan : Mahasiswa
- d. Institusi : Universitas Stikubank Semarang
- e. Waktu yang tersedia : 10 Jam per minggu