

# PEMETAAN AKSES HALAMAN SITUS WEB BERBASIS LOG-ACCESS (*Log-Access Based Web Site Access Mapping Page*)

Dison Librado<sup>1)</sup>, dan Wagito<sup>2)</sup>

<sup>1, 2)</sup> Program Studi Teknik Informatika STMIK AKAKOM

Jl. Raya Janti no. 143, Yogyakarta

e-mail: disonlibrado@gmail.com<sup>1)</sup>, wagito@akakom.ac.id<sup>2)</sup>

## ABSTRAK

*STMIK AKAKOM sebagai perguruan tinggi swasta telah lama memanfaatkan teknologi internet sebagai media informasi dan komunikasinya. Berbagai informasi yang disediakan dalam situs resmi kampus ini adalah Beranda, Profil Akakom, Program Studi/Jurusan, Info Kampus, Berita, Arsip Umum, dan lainnya. Namun demikian, belum ada kajian tentang bagaimana situs tersebut dikunjungi pengguna. Kajian tentang halaman yang dikunjungi berguna untuk bahan evaluasi terhadap situs. Penelitian bertujuan untuk melihat jumlah serta pemetaan kunjungan tiap halaman situs.*

*Tahapan yang dikerjakan adalah studi pustaka dari sumber referensi yang relevan, mengonfigurasi server web Apache, mengumpulkan data log access selama waktu tertentu, kemudian mengolah data log access, membuat pembahasan hasil pengolahan data, dan membuat kesimpulan. Pengolahan log access dilakukan dengan cara mengubah format log access menjadi format csv. Selanjutnya format csv diubah menjadi format basisdata MySQL. Pengolahan selanjutnya dilakukan menggunakan perintah SQL. Hasil disajikan dalam bentuk informasi jumlah pengunjung pada halaman situs, pola kunjungan dan identifikasi halaman situs yang sering dikunjungi.*

*Hasil penelitian menunjukkan bahwa kategori yang paling diminati oleh pengunjung adalah Berita, disusul dengan kategori Info Kampus, selanjutnya Profil Akakom, Program Studi/Jurusan, Lainnya, dan kategori terakhir adalah Arsip Umum. Hasil penelitian diharapkan dapat digunakan untuk membantu pihak pengelola dan pengembang agar dapat meningkatkan kualitas layanan informasinya.*

**Kata Kunci:** pemetaan, log, situs, web, mining.

## ABSTRACT

*STMIK AKAKOM as a private university has long utilized internet technology as a medium of information and communications. Various information provided on the official website of this campus is Home, Akakom Profile, Study Program / Department, Campus Info, News, General Archives, and more. However, there is no study of how the site is visited by users. A review of the pages covered is useful for the evaluation of the site. The study aims to see the number and mapping of visits per page of the site.*

*The steps taken are literature studies from relevant reference sources, configuring the Apache web server, collecting log data access for a certain time, then processing log access data, making discussion of the results of data processing, and making conclusions. Processing log access is done by changing the format of log access into csv format. Furthermore csv format changed to MySQL database format. The next processing is done using SQL commands. The results are presented in the form of information on the number of visitors on the site page, the pattern of visits and identification of frequently visited sites pages.*

*The result of the research shows that the category most interested by visitors is News, followed by Campus Info category, then Akakom Profile, Study Program / Department, Others, and the last category is General Archive. The results of the research are expected to be used to assist the managers and developers in order to improve the quality of information services.*

**Keywords:** mapping, log, site, web, mining.

## I. PENDAHULUAN

Internet adalah teknologi yang menghubungkan komputer dengan jaringan yang sangat populer pada zaman ini. Tercatat informasi dari web Asosiasi Penyelenggara Jasa Internet Indonesia (<http://www.apjii.or.id/> diakses tanggal 9 Februari 2017) bahwa pengguna internet pada tahun 2014 sebanyak 88,1 juta orang, dan jumlah ini meningkat menjadi 132,7 juta orang pada tahun 2016. Ini menunjukkan adanya peningkatan yang signifikan dari jumlah pengguna internet. Artinya kebutuhan dan penggunaan teknologi oleh masyarakat sudah sangat meningkat. Hal ini yang mendorong perubahan bentuk berbagai proses dalam aktivitas sehari-hari dari cara yang manual bergeser menjadi cara yang otomatis. Pergeseran ini terjadi karena masyarakat merasakan kemudahan dalam berbagai hal dari penggunaan internet.

Berdasarkan informasi dari web Asosiasi Penyelenggara Jasa Internet Indonesia pula dapat diketahui komposisi pengguna internet di Indonesia berdasarkan usia adalah sebanyak 24,4 juta (18,4%) adalah berusia 10 – 24 tahun, 32,3 juta (24,4%) berusia 25 – 34 tahun, 38,7 (29,2%) berusia 35 – 44 tahun, 23,8 juta (18%) berusia 45 – 54 tahun, dan sebanyak 13,2 juta orang (10%) pengguna yang berusia 55 tahun ke atas. Artinya pengguna internet terbanyak berdasarkan usia adalah pada kisaran 35 – 44 tahun. Sementara berdasarkan informasi tersebut, penetrasi pengguna

internet Indonesia berdasarkan pekerjaan adalah sebanyak 89,7% mahasiswa, 69,8% pelajar, 58,4% pekerja, 25,3% ibu rumah tangga, dan sebanyak 6,7% adalah pengguna dengan pekerjaan lainnya.

STMIK AKAKOM sebagai salah satu perguruan tinggi swasta di wilayah KOPERTIS 5, telah lama memanfaatkan teknologi internet sebagai media informasi dan komunikasinya. Berbagai informasi yang disediakan dalam situs resmi kampus ini adalah Beranda, Profil Akakom, Program Studi/Jurusan, Info Kampus, Berita, Arsip Umum, dan Lainnya. Di dalam setiap menu ini masih terdapat beberapa sub menu yang berhubungan dengan menu utamanya. Selain dari menu-menu utama yang berada pada baris navigasi ini juga terdapat menu lain yang terletak di panel bagian kanan yaitu Visi dan Misi, Kumpulan Pedoman dan Peraturan, Sikeu Portal, dan Tautan.

Selanjutnya dapat dirumuskan masalahnya yaitu berapa jumlah pengunjung dalam kurun waktu enam bulan pada tahun 2016 dan halaman apa yang sering dikunjungi oleh pengunjung situs. Studi ini didasarkan log dari kunjungan ke situs web STMIK AKAKOM. Untuk membatasi jumlah pengamatan data yang demikian banyaknya maka penelitian ini membatasi cakupan datanya pada tahun 2016 dengan rincian pengamatan yang tidak sampai pada identifikasi lokasi asal pengguna melakukan akses.

Informasi hasil pengumpulan dan pengamatan data kunjungan para pengguna web bertujuan untuk menyajikan informasi jumlah pengunjung berdasarkan kunjungan dan keragaman pengunjung pada situs, melihat pola kunjungan pada web dan identifikasi halaman yang sering dikunjungi. Pengolahan data didasarkan pada *log access* server web yang diolah menjadi bentuk format MySQL. Dengan melakukan studi dan mengamati pada data statistik kunjungan web ini diharapkan dapat digunakan sebagai bahan evaluasi terhadap perkembangan situs web STMIK AKAKOM oleh pihak yang berkompeten.

Penelitian tentang *web usage mining* pernah dilakukan dengan konsentrasi dan objek yang berbeda-beda. [1] dalam tulisannya menyampaikan materi tentang penggalian data di situs [www.faperta.unja.ac.id](http://www.faperta.unja.ac.id) dengan mengembangkan sistem atau program yang mengaplikasikan metode *browser sniffing* (metode tanpa *submitform*) untuk menangkap alamat log yang dihasilkan oleh user secara otomatis. Pengembangan program tersebut menggunakan php dan mysql yang disisipkan di file dalam situs tersebut. Data kunjungan ke situs tersebut, yang tersimpan dalam database mysql, selanjutnya diproses secara manual menggunakan program bantu Microsoft Excel dan Matlab 2008a. Dari log file yang telah digali dapat diketahui jumlah pengunjung situs yang ditampilkan melalui *web traffic* serta *pattern discovery*-nya yang dapat menjadi rekomendasi bagi perancang web untuk mengatur penempatan item atau modul yang ada di situs tersebut.

Berkaitan dengan perkembangan internet yang telah memicu peningkatan aktivitas belanja online, [2] membuat penelitian yang menerapkan *web usage mining* untuk membantu perusahaan dalam meningkatkan penjualan. Metode *web usage mining* yang akan diimplementasikan ke dalam sistem *ecommerce* itu dapat memberikan satu solusi baru untuk menganalisis transaksi yang terjadi di toko online maupun pangsa pasar suatu produk berdasarkan pembelian konsumen, dan kunjungan ke web. Simpulan yang dibuat dalam tulisan *web usage mining* itu menjadi data yang dapat digali untuk menunjang analisis penjualan agar dapat menjadi alat bantu bagi para pemilik perusahaan untuk menganalisis perkembangan alur distribusi perdagangan di perusahaannya.

Dalam referensi yang lain, [3] dilakukan penelitian tentang pengelolaan sistem informasi akademik (SIA) perguruan tinggi berbasis teknologi informasi dan komunikasi (TIK). Efektivitas aplikasi TIK dalam proses manajemen kelembagaan sering terhambat oleh banyak faktor non-teknis yang tidak dipersiapkan oleh lembaga. Faktor-faktor non-teknis itu seperti penyiapan orang, budaya, mekanisme organisasi, sampai dengan masalah teknis pemeliharaan. Tidak selamanya SIA yang berbasis TIK dapat meningkatkan kinerja pengelolaan administrasi akademik manakala lembaga hanya menganggap bahwa implementasi TIK untuk SIA hanya sekadar menyiapkan perangkat keras TIK. Selanjutnya dengan memperhatikan hal-hal yang telah disebutkan, penelitian ini mencoba untuk memperoleh data empirik yang cukup lengkap dan dapat dipercaya untuk menggambarkan tentang keadaan faktor-faktor yang terkait dengan pengelolaan Sistem Informasi Akademik berbasis Teknologi Informasi dan Komunikasi terhadap kinerja perguruan tinggi dan dampaknya terhadap prestasi akademik mahasiswa. Data yang telah diperoleh juga dapat digunakan untuk mengembangkan alternatif model Sistem Informasi Manajemen Akademik yang dapat memberi dukungan pada setiap proses pelayanan akademik maupun pengambilan keputusan baik di lingkungan internal maupun eksternal yang terkait dengan stakeholders.

Agar dapat mengelola server web secara efektif, perlu mendapat umpan balik tentang aktivitas dan kinerja server serta masalah yang mungkin terjadi. Apache HTTP Server menyediakan kemampuan *logging* yang sangat komprehensif dan fleksibel. Dokumen ini menjelaskan cara mengonfigurasi kemampuan *logging*-nya, dan bagaimana memahami isi log. Apache HTTP Server menyediakan berbagai mekanisme yang berbeda untuk mencatat semua yang terjadi di server, dari permintaan awal, melalui proses pemetaan URL, hingga penyelesaian akhir dari koneksi, termasuk kesalahan yang mungkin terjadi dalam prosesnya. Selain itu, modul pihak ketiga mungkin menyediakan kemampuan *logging*, atau memasukkan masukan ke dalam file log yang ada, dan aplikasi

seperti program CGI, atau skrip PHP, atau penangan lainnya, dapat mengirim pesan ke log kesalahan server. Log akses server mencatat semua permintaan yang diproses oleh server. Lokasi dan isi log akses dikontrol oleh *direktif CustomLog*. Petunjuk LogFormat dapat digunakan untuk menyederhanakan pemilihan isi log. Bagian ini menjelaskan cara mengonfigurasi server untuk mencatat informasi di log akses. Tentu saja penyimpanan informasi di access log hanyalah awal dari pengelolaan log. Langkah selanjutnya adalah menganalisis informasi ini untuk menghasilkan statistik yang bermanfaat. Analisis log secara umum berada di luar cakupan dokumen ini, dan tidak benar-benar bagian dari pekerjaan server web itu sendiri. Untuk informasi lebih lanjut tentang topik ini, dan untuk aplikasi yang melakukan analisis log, periksalah Direktori Open. Berbagai versi Apache httpd telah menggunakan modul dan arahan lain untuk mengendalikan akses logging, termasuk `mod_log_referer`, `mod_log_agent`, dan `TransferLogDirective`. Petunjuk CustomLog sekarang mencakup fungsionalitas dari semua arahan yang lebih tua. Format log akses sangat dapat dikonfigurasi. Formatnya ditentukan menggunakan format string yang terlihat seperti string format Cf style (1). Beberapa contoh disajikan pada bagian berikutnya. Untuk daftar lengkap isi format string yang mungkin, lihat format `mod_log_config` string.

Konfigurasi khas untuk log akses mungkin terlihat seperti berikut:

```
LogFormat "%h %l %u %t \"%r\" %>s %b" common
CustomLog logs/access_log common
```

String ini mendefinisikan julukan yang umum dan mengaitkannya dengan format string log tertentu. Format string terdiri dari persen arahan, masing-masing memberi tahu server untuk mencatat informasi tertentu. Karakter literal juga dapat ditempatkan dalam format string dan akan disalin langsung ke output log. Karakter kutipan (") harus diloloskan dengan menempatkan garis miring terbalik sebelum mencegahnya ditafsirkan sebagai akhir dari format string. String format juga berisi karakter kontrol khusus " \n "untuk baris baru dan" \T "untuk tab.

Petunjuk CustomLog menyiapkan file log baru menggunakan julukan yang ditentukan. Nama file untuk log akses relatif terhadap ServerRoot kecuali jika dimulai dengan garis miring. Konfigurasi di atas akan menuliskan entri log dalam format yang dikenal dengan Common Log Format (CLF). Format standar ini dapat diproduksi oleh banyak server web yang berbeda dan dibaca oleh banyak program analisis log. Entri file log yang diproduksi di CLF akan terlihat seperti ini:

```
127.0.0.1 - frank [10/Oct/2000:13:55:36 -0700] "GET /apache_pb.gif
HTTP/1.0" 200 2326
```

Setiap bagian dari entri log ini dijelaskan di bawah ini.

- 1) Nilai 127.0.0.1 (%h) adalah alamat IP dari client (remote host) yang melakukan request ke server. Jika `HostnameLookups` disetel ke On, maka server akan mencoba menentukan nama host dan mencatatnya di tempat alamat IP. Namun, konfigurasi ini tidak disarankan karena bisa memperlambat server secara signifikan.
- 2) Nilai - (%l). "Tanda hubung" pada output menunjukkan bahwa bagian informasi yang diminta tidak tersedia. Dalam hal ini, informasi yang tidak tersedia adalah identitas RFC 1413 dari klien yang ditentukan oleh identd pada mesin klien.
- 3) Nilai frank (%u) adalah userid dari orang yang meminta dokumen sebagaimana ditentukan oleh otentikasi HTTP. Nilai yang sama biasanya diberikan ke skrip CGI di variabel lingkungan `REMOTE_USER`.
- 4) Nilai [10/Oct/2000:13:55:36 -0700] (%t) adalah waktu permintaan itu diterima. Formatnya adalah: [day/month/year:hour:minute:second zone]. Nilai day = 2\*digit, month = 3\*letter, year = 4\*digit, hour = 2\*digit, minute = 2\*digit, second = 2\*digit dan zone = (^+'|^-) 4\*digit
- 5) Nilai "GET /apache\_pb.gif HTTP/1.0" ("%r") adalah baris permintaan dari klien diberikan dengan tanda kutip ganda. Baris permintaan berisi banyak informasi bermanfaat. Pertama, metode yang digunakan oleh klien adalah GET. Kedua, klien meminta sumber daya /apache\_pb.gif, dan ketiga, klien menggunakan protokol HTTP / 1.0.
- 6) Nilai 200 (%>s) adalah kode status yang dikirim server kembali ke klien. Informasi ini sangat berharga, karena mengungkapkan apakah permintaan tersebut menghasilkan respons yang berhasil (kode diawali dengan 2), pengalihan (kode dimulai dari 3), kesalahan yang disebabkan oleh klien (kode diawali dengan 4), atau kesalahan dalam Server (kode diawali dengan 5).
- 7) Nilai 2326 (%b) adalah bagian terakhir menunjukkan ukuran objek yang dikembalikan ke klien, tidak termasuk header respon. Jika tidak ada konten yang dikembalikan ke klien, nilai ini akan "-".

## II. METODE

Penelitian terdiri dari beberapa tahap yaitu: melakukan studi pustaka bersumber dari referensi yang relevan dengan topik yang akan diteliti, melakukan konfigurasi pada server web Apache, mengumpulkan *data log access* selama waktu tertentu, mengolah data log access, membuat pembahasan pada hasil pengolahan data, dan membuat

kesimpulan. Objek yang digunakan adalah situs [www.akakom.ac.id](http://www.akakom.ac.id) yang merupakan situs resmi dari kampus STMIK AKAKOM. Waktu pengamatan selama kurun waktu di tahun 2016. Pengamatan dilakukan dalam satu semester. Model yang digunakan dalam penelitian ini adalah Web Usage Mining. Penelitian ini dipersiapkan dengan rancangan yang mengikuti tahapan penelitian sebagai berikut.

- 1) Menentukan objek penelitian. Dalam penelitian ini telah dipilih objek penelitian pada Web kampus STMIK AKAKOM dengan url [www.akakom.ac.id](http://www.akakom.ac.id),
- 2) Melakukan konfigurasi pada server web Apache
- 3) Mengumpulkan data log access selama kurun waktu tertentu, kemudian data tersebut dikonversi ke dalam format MySQL
- 4) Mengolah data log access
- 5) Membuat kesimpulan

Pengerjaan penelitian ini melibatkan kegiatan-kegiatan dengan rincian

- 1) Mengaktifkan mode log konfigurasi web server Apache
- 2) Mengamati log akses server web selama waktu tertentu
- 3) Mengonversi log akses server web dalam database MySQL
- 4) Mengelompokkan data berdasarkan halaman situs
- 5) Menghitung jumlah akses untuk tiap halaman situs
- 6) Memetakan distribusi akses halaman situs

Konfigurasi server web Apache secara umum dituliskan dalam file `/etc/httpd/conf/http.conf`. Dalam distri-distro Linux tertentu, letak file konfigurasi ditempatkan pada direktori yang bervariasi. Namun demikian, isi file konfigurasi dapat dikatakan sama. Isi konfigurasi sangat banyak mulai dari nama server, alamat IP kerja, port kerja, letak dokumen HTML, pengaturan log baik log error maupun log akses, virtual host, pengaturan izin akses direktori dan sebagainya. Pengaturan konfigurasi yang berkaitan dengan penelitian ini adalah pengaturan log server khusus berkaitan dengan log akses pada situs.

Supaya server web Apache dapat menghasilkan log akses, maka pada file konfigurasi perlu dituliskan pengarah yang mengatur pencatatan log. Pengaturan terutama berupa format log serta file pencatatan log. Pengaturan yang dilakukan berbentuk sebagai berikut.

```
# The following directives define some format nicknames for use with
# a CustomLog directive (see below).
#
LogFormat "%h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-Agent}i\"" combined

# For a single logfile with access, agent, and referer information
# (Combined Logfile Format), use the following directive:
#
CustomLog logs/access_log combined
```

Konfigurasi tersebut, terutama mengatur bentuk catatan log yang dihasilkan dan file tempat menyimpan catatan log tersebut. Format catatan log diatur pada pengarah LogFormat. Letak file tempat menyimpan catatan log diatur pada pengarah CustomLog. Pada pengaturan tersebut, catatan log disimpan dalam file teks `access_log`.

Data log akses dikumpulkan selama waktu tertentu yang dalam penelitian ini dilakukan dalam kurun waktu lebih dari enam bulan (kira-kira satu semester). Selama kurun waktu tersebut web Akakom akan diakses secara acak oleh pengguna baik dari lingkungan kampus maupun dari luar kampus. Dalam penelitian tidak diperhatikan dari mana dan oleh siapa web Akakom diakses, namun hanya halaman mana yang diakses.

Log akses Apache mengandung beberapa informasi seperti alamat IP user yang akses, waktu akses, halaman yang diakses, browser untuk akses dan sebagainya. Informasi yang diperlukan dalam penelitian ini terutama adalah halaman akses. Sebagai informasi tambahan dapat juga dilibatkan alamat IP dan waktu akses. Informasi itu dapat diperoleh pada kolom satu untuk IP, kolom empat untuk waktu dan kolom tujuh untuk halaman akses.

Untuk mengolah catatan log server digunakan beberapa langkah. Langkah pertama adalah mengkonversi format file `access_log` menjadi format csv dan mengambil informasi yang diperlukan saja. Informasi yang paling diperlukan untuk penelitian ini adalah halaman akses. Beberapa *tool* Linux dilibatkan untuk keperluan ini yaitu `echo`, `cat`, `grep` dan `awk`. `Echo` untuk menampilkan suatu string, `cat` untuk menampilkan isi file teks, `grep` untuk menjalankan filter, sedangkan `awk` untuk pemrosesan pola string.

```
echo "ip;timestamp;urlakses">access_log.csv;
cat access_log | grep \"GET | awk '{printf(\"%s;%s;\",$1,$4)}{print $7}'>>access_log.csv
```

Baris pertama pada perintah adalah untuk membuat *header* file csv. Isi header harus sesuai dengan nama *field* yang nantinya digunakan untuk menyimpan data pada database MySQL. Dalam hal ini nama *field*-nya adalah *ip*, *timestamp*, *urlakses* yang masing-masing mewujudkan alamat IP asal mesin yang digunakan untuk akses, waktu akses dan halaman yang di akses. Hasil keluaran perintah diarahkan pada file *access\_log.csv* dengan mode *timepa*.

Baris kedua digunakan untuk menampilkan isi kolom pertama (\$1), kolom keempat (\$4) dan kolom ketujuh (\$7) yang pada catatan *access\_log* berisi informasi tentang alamat IP asal, waktu akses dan halaman yang diakses. Inti pada perintah adalah *awk* dikombinasikan dengan perintah filter *grep* untuk menyaring baris yang ada kata kunci *"GET*". Pada perintah digunakan *cat* untuk menampilkan seluruh isi file *access\_log* secara baris demi baris. Hasil keluaran diarahkan pada file *access\_log.csv* dengan mode tambah (*add*) supaya hasil pengolahan ditambahkan pada baris paling bawah.

*Awk* adalah bahasa pengolah dan pemindai pola. *Awk* yang dipakai pada perintah adalah bagian dari perangkat lunak *Gawk*. *Gawk* adalah implementasi GNU Project dari bahasa pemrograman *AWK*. Ini Sesuai dengan definisi bahasa di standar POSIX 1003.1. Versi ini pada gilirannya berdasarkan uraian pada bahasa pemrograman *AWK*, oleh Aho, Kernighan, dan Weinberger. *Gawk* menyediakan fitur tambahan yang ditemukan dalam versi saat ini dari *awk* Brian Kernighan dan sejumlah ekstensi GNU spesifik. [7]

Pengolahan file *access\_log* akan menghasilkan file teks berformat csv yaitu *access\_log.csv*. File csv sebetulnya dapat langsung diolah menggunakan pengolah spreadsheet seperti excel atau libre office calc. Namun karena ukuran file log Apache yang cukup besar, pengolah spreadsheet tidak cukup mampu untuk menolah data *access\_log*. Untuk mengolahnya, data tersebut perlu disimpan dalam sistem database MySQL. Sistem database ini mampu mengolah data dalam jumlah yang sangat besar. Untuk memasukkan data csv pada database MySQL digunakan bantuan *tool* Adminer. Adminer adalah *tool* untuk pengolahan database yang mendukung banyak sistem database termasuk MySQL. Banyak fitur yang disediakan oleh Adminer untuk pengolahan database mulai dari mengelola database, mengelola tabel, menjalankan SQL serta import dan ekspor data. Import dan ekspor data dapat berasal atau menuju format SQL ataupun csv. Adminer adalah alat manajemen database fitur lengkap yang ditulis dalam PHP. Sebaliknya untuk *phpMyAdmin*, itu terdiri dari satu file siap untuk menyebarkan ke server target. Adminer yang tersedia untuk MySQL, PostgreSQL, SQLite, MS SQL, Oracle, Firebird, SimpleDB, Elasticsearch dan mongodb [8]. Beberapa langkah perlu dilakukan untuk menyimpan *data access\_log* pada sistem database MySQL. Langkah pertama yang dilakukan adalah membuat database baru misalnya *logdb* atau memanfaatkan database yang sudah ada misalnya *test*. Jika dibuat database baru *logdb*, maka pada kolom SQL command Adminer diberi perintah sebagai berikut.

```
CREATE DATABASE `logdb`
```

Selanjutnya, dalam database dibuat tabel misalnya *logakses* yang strukturnya dibuat sepadan dengan file csv yang sudah disiapkan. Tabel *logakses* dibuat dengan field-field yaitu: *id* bertipe *int* auto increment, *ip* bertipe *varchar*, *timestamp* bertipe *varchar*, *urlakses* bertipe *varchar*. Perintah SQL untuk membentuk struktur tabel *logakses* dijalankan pada kolom SQL command. Perintah SQL yang digunakan adalah sebagai berikut.

```
CREATE TABLE `logakses` (
  `id` int(11) NOT NULL AUTO_INCREMENT,
  `ip` varchar(16) NOT NULL,
  `timestamp` varchar(25) NOT NULL,
  `urlakses` varchar(200) NOT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
```

Setelah data berhasil masuk pada database MySQL, pengolahan selanjutnya dilakukan menggunakan perintah SQL. Perintah SQL dapat dijalankan melalui Adminer pada kolom SQL command. Informasi yang diinginkan adalah jumlah akses pada suatu halaman web selama kurun waktu tertentu. Dengan demikian harus dihitung cacah akses setiap halaman yang sama. Hasilnya harus berupa halaman akses versus jumlah akses. Untuk mendapatkan informasi tersebut, dapat digunakan perintah SQL sebagai berikut.

```
select urlakses, count(*) as jumlah from logakses group by urlakses order by jumlah desc limit 100;
```

Maksud perintah SQL tersebut adalah menampilkan field *urlakses* dan cacah *urlakses* sama sebagai jumlah. Hasilnya dikelompokkan oleh *urlakses*. Selanjutnya hasil diurutkan berdasarkan hasil jumlah akses secara urut turun (*descending*). Hasilnya difilter lagi supaya hanya ditampilkan 100 urutan jumlah akses terbesar. Hasilnya dapat diekspor dalam file csv yang dapat diolah lebih mudah, karena dibatasi sampai 100 data terbesar.

### III. HASIL

Analisis Statistik dilakukan dengan menggunakan data yang diperoleh dari hasil pengonversian logakses server web dalam bentuk database. Kemudian database tersebut telah dikelompokkan berdasar halaman situs yang kemudian disebut sebagai kategori. Untuk kepentingan pembatasan jumlah data dilakukan penyaringan data untuk seratus (100) log akses terbanyak dalam periode waktu itu. Dalam kurun waktu enam (6) bulan lebih di tahun 2016 dari data tersaring tercatat sebanyak 3.194.427 pengakses yang mengakses ke halaman situs www.akakom.ac.id. Dari akses log tersebut dapat diidentifikasi kateori-kategori yang diakses beserta jumlah pengunjungnya.

Rekapitulasi data yang diperoleh kemudian disusun ke dalam tabel berdasarkan kategori menu dalam web. Informasi tersebut dapat dilihat dalam Tabel I. jumlah pengakses dapat disebutkan dengan urutan lima (5) terbanyak yaitu akses ke kategori Berita, disusul dengan kategori Info Kampus, kemudian kategori Profil Akakom, kategori Program Studi/Jurusan, dan kategori Arsip Umum, dan Lainnya..yang memiliki jumlah pengakses kurang dari 8000 dalam kurun waktu pengumpulan data.

Masing-masing kategori dalam web memiliki sub kategori yang terdiri dari beberapa informasi yang disebutkan dalam tabel dengan “tentang”. Dimulai dari Profil Akakom, kategori ini terdiri dari beberapa sub kategori yaitu Visi dan Misi, Profile STMIK AKAKOM, Kontak Kami, Pejabat di STMIK AKAKOM, dan Kontak Dosen STMIK AKAKOM. Dari lima sub kategori Profil Akakom ini informasi yang paling banyak diakses adalah tentang Kontak Dosen STMIK AKAKOM sebanyak 9851 akses, artinya informasi dosen di STMIK AKAKOM dianggap paling penting dari kategori Profil Akakom ini sehingga informasi ini perlu dijaga kualitasnya. Informasi yang banyak diakses selanjutnya adalah tentang Pejabat di STMIK AKAKOM dengan jumlah pengakses sebanyak 7273. Hal ini tentunya berkaitan dengan kepentingan pengguna terhadap keberadaan pejabat di kampus ini yang berhubungan dengan proses akademik yang dijalankan. Informasi jumlah pengakses kategori Profil Akakom selengkapnya dapat dilihat pada Tabel II.

TABEL I  
DAFTAR KATERGORI MENU WEB DAN JUMLAH PENGAKSES

KATEGORI	TOTAL per KATEGORI
Profil Akakom	35.174
Program Studi/Jurusan	25.142
Info Kampus	35.794
Berita	37.802
Arsip Umum	11.244
Lainnya	14.053

TABEL II  
DAFTAR JUMLAH PENGAKSES PADA KATEGORI PROFIL AKAKOM

SUB KATEGORI	TENTANG	JUMLAH
Profil Akakom	Visi dan Misi	6.597
	Profile STMIK AKAKOM	6.445
	Kontak Kami	5.008
	Pejabat di STMIK AKAKOM	7.273
	Kontak Dosen STMIK AKAKOM	9.851

Informasi jumlah pengakses kategori Program Studi ini dapat dilihat pada Tabel III. Dari kelima sub kategori program studi tersebut jumlah pengakses yang paling banyak adalah ke sub kategori program studi Teknik Informatika (S1) dengan jumlah pengakses 8.219 dalam kurun waktu yang ditentukan dalam penelitian. Selanjutnya pada urutan kedua adalah pengaksesan ke sub kategori program studi Sistem Informasi (S1) dengan jumlah pengakses 6587, diikuti program studi Komputerisasi (D3) dengan jumlah pengakses 3.638, program studi Manajemen Informatika (D3) dengan jumlah pengakses 3434, dan terakhir jumlah pengakses ke program studi Teknik Komputer (D3) terdapat sebanyak 3.264.

TABEL III  
DAFTAR JUMLAH PENGAKSES PADA KATEGORI PROGRAM STUDI/JURUSAN

SUB KATEGORI	TENTANG	JUMLAH
Program Studi/Jurusan	Komputerisasi Akuntansi (D3)	3.638
	Teknik Komputer (D3)	3.264
	Manajemen Informatika (D3)	3.434
	Sistem Informasi (S1)	6.587
	Teknik Informatika (S1)	8.219

TABEL IV  
DAFTAR JUMLAH PENGAKSES PADA KATEGORI ARSIP UMUM

SUB KATEGORI	TENTANG	JUMLAH
Arsip Umum	Kalender Akademik	7.497
	Blog STMIK AKAKOM Yogyakarta	2.147
	menu arsip-umum	1.600

Sub Kategori dari kategori Info Kampus ini sebenarnya akan selalu berubah mengikuti perubahan informasi kegiatan yang terjadi di kampus. Selama ada kegiatan atau jadwal acara yang terselenggara di kampus tentunya akan mengisi sub kategori info kampus ini. Pada kurun waktu analisis dilakukan maka sub kategori yang paling banyak diakses oleh pengguna adalah Pengumuman PKL TA 2015/2016 dengan jumlah pengakses sebanyak 2143. Hal ini mungkin terjadi karena pada saat itu informasi kegiatan PKL sangat diperlukan sehingga jumlah pengaksesnya meningkat.

Kategori lainnya dalam web kampus ini adalah Berita. Jika kategori Info Kampus berisikan informasi-informasi terkait dengan kegiatan akademik maka kategori Berita ini berisikan informasi tentang kegiatan-kegiatan yang berhubungan dengan pemegang kepentingan (stakeholders) dari luar kampus, seperti perbankan, perguruan tinggi

lain, atau instansi yang terkait dengan kegiatan kampus.

Kategori selanjutnya adalah Arsip Umum yang hanya memiliki 2 sub kategori yaitu Kalender Akademik dan Blog STMIK AKAKOM Yogyakarta. Jumlah pengakses untuk sub kategori Kalender Akademik adalah sebanyak 7.497, sedangkan jumlah pengakses sub kategori Blog STMIK AKAKOM Yogyakarta adalah sebanyak 2.147. Namun demikian terdapat sejumlah 1.600 pengakses yang hanya menuju pada menu Arsip Umum saja. Informasi jumlah pengakses kategori Arsip Umum ini dapat dilihat pada Tabel IV.

Kategori terakhir dalam web kampus STMIK AKAKOM ini adalah Lainnya. Sub kategori ini berisi empat (4) sub kategori yaitu Unduh Arsip, Framework Yii, Tutorial MySQL, dan Linux. Sub kategori Framework Yii dapat berisikan informasi dinamis disesuaikan dengan topik pembahasan yang sedang diminati. Informasi jumlah pengakses kategori Lainnya ini dapat dilihat pada Tabel V.

TABEL V  
DAFTAR JUMLAH PENGAKSES PADA KATEGORI LAINNYA..

SUB KATEGORI	TENTANG	JUMLAH
Unduh Arsip	Sumber Unduh	7.669
Framework Yii	Akses SQL Menggunakan CArrayDataProvider(...) Dengan Parameter Input	1.385
	Studi Kasus Menghitung Gaji Sederhana Dengan Yii menu Framework Yii	1.120
		2.499
Tutorial MySQL	Perhitungan Tanggal Dan Waktu Dengan MySQL	1.666
	Stored Procedure Dan Function	1.071
	menu Tutorial MySQL	1.873
Linux	Linux	1.143

#### IV. PEMBAHASAN

Dengan jumlah tersebut dapat dikatakan bahwa informasi yang paling dibutuhkan dari web kampus adalah kategori Berita. Pengguna akan merasa puas jika kebutuhan informasinya dapat terpenuhi. Pemenuhan kebutuhan ini artinya jika informasi dapat diperoleh pengguna. Dengan demikian untuk menjaga loyalitas para pengguna web kampus ini, pengelola web harus menjaga kualitas berita yang ditampilkannya dalam halaman web. Kualitas berita dapat terpenuhi dengan akurasi, tepat waktu, dan relevan [10]. Hal ini sebenarnya berlaku juga untuk kategori Info Kampus, kemudian kategori Profil Akakom, kategori Program Studi/Jurusan, dan kategori Arsip Umum.

Dari lima sub kategori Profil Akakom seperti dalam Tabel II, informasi yang paling banyak diakses adalah tentang Kontak Dosen STMIK AKAKOM, yaitu sebanyak 9.851 akses, artinya informasi dosen di STMIK AKAKOM dianggap paling penting dari kategori Profil Akakom ini sehingga informasi ini perlu dijaga kualitasnya. Informasi yang banyak diakses selanjutnya adalah tentang Pejabat di STMIK AKAKOM dengan jumlah pengakses sebanyak 7.273. Hal ini tentunya berkaitan dengan kepentingan pengguna terhadap keberadaan pejabat di kampus ini yang berhubungan dengan proses akademik yang dijalankan. Informasi jumlah pengakses kategori Profil Akakom selengkapnya dapat dilihat pada Tabel II. Informasi jumlah pengakses kategori Program Studi yang dapat dilihat pada Tabel III mendeskripsikan bahwa jumlah pengakses yang paling banyak adalah ke sub kategori program studi Teknik Informatika (S1) dengan jumlah pengakses 8.219 dalam kurun waktu yang ditentukan dalam penelitian. Selanjutnya pada urutan kedua adalah pengaksesan ke sub kategori program studi Sistem Informasi (S1) dengan jumlah pengakses 6587, diikuti program studi Komputerisasi (D3) dengan jumlah pengakses 3.638, program studi Manajemen Informatika (D3) dengan jumlah pengakses 3434, dan terakhir jumlah pengakses ke program studi Teknik Komputer (D3) terdapat sebanyak 3.264. Kategori berikutnya adalah Info Kampus yang memiliki 23 sub kategori. Lima kategori terbanyak yang diakses oleh pengguna adalah Pengumuman PKL TA 2015/2016, Informasi KRS Jurusan TI Tanggal 25 Februari 2015, Buku Praktik Mahasiswa Jenjang S1, Pengumuman Sosialisasi Mekanisme Skripsi, dan Seminar Bioenergi.

Sub Kategori dari kategori Info Kampus akan selalu berubah mengikuti perubahan informasi kegiatan yang terjadi di kampus. Selama ada kegiatan atau jadwal acara yang terselenggara di kampus tentunya akan mengisi sub kategori info kampus ini. Pada kurun waktu analisis dilakukan maka sub kategori yang paling banyak diakses oleh pengguna adalah Pengumuman PKL TA 2015/2016 dengan jumlah pengakses sebanyak 2.143. Hal ini mungkin terjadi karena pada saat itu informasi kegiatan PKL sangat diperlukan sehingga jumlah pengaksesnya.

Sama halnya dengan kategori Info Kampus, kategori Berita ini akan selalu berubah mengikuti perubahan informasi kegiatan yang diberitakan oleh kampus. Selama ada kegiatan atau jadwal acara yang terselenggara di kampus akan diisikan pada sub kategori berita ini. Pada kurun waktu analisis dilakukan, sub kategori yang paling banyak diakses oleh pengguna adalah “Akakom Berlakukan Student Payment” dengan pengakses sebanyak 2.287. Sub kategori Berita selanjutnya yang banyak diakses adalah berita tentang Kuliah Perdana Mahasiswa STMIK AKAKOM dengan jumlah pengakses sebanyak 1.897, diikuti dengan sub kategori lainnya.

Banyaknya jumlah pengakses Kalender Akademik menunjukkan pentingnya informasi ini bagi sivitas akademika sebagai pedoman pelaksanaan kegiatannya. Sementara jumlah pengakses sub kategori Blog STMIK AKAKOM Yogyakarta hanya sebanyak 2.147 dalam kurun waktu tersebut meningkat.

Pada saat data diambil, informasi yang diakses dalam sub kategori Framework Yii ini adalah “Akses SQL Menggunakan CArrayDataProvider(...) Dengan Parameter Input” dengan jumlah pengakses sebanyak 1.385, “Studi Kasus Menghitung Gaji Sederhana Dengan Yii” dengan jumlah pengakses sebanyak 1.120, dan pengguna yang hanya mengakses menu Framework Yii saja yang berjumlah 2.499. Untuk sub kategori MySQL memiliki dua (2) informasi di dalamnya yaitu “Perhitungan Tanggal dan Waktu dengan MySQL” yang memiliki jumlah pengakses sebanyak 1666 dan informasi “Stored Procedure dan Function” yang diakses oleh 1.071 pengakses. Sementara terdapat sejumlah 1.873 pengguna yang hanya mengakses menu Tutorial MySQL saja. Dari sub kategori Linux terdapat 1.143 pengakses.

## V. SIMPULAN DAN SARAN

Kesimpulan yang dapat diambil dari penelitian yang dilakukan adalah bahwa sebagai berikut.

- 1) Pemetaan jumlah kunjungan pada setiap halaman situs [www.akakom.ac.id](http://www.akakom.ac.id) dapat dilakukan berdasarkan akses log server.
- 2) Kategori yang paling diminati oleh pengunjung pada halaman situs [www.akakom.ac.id](http://www.akakom.ac.id) adalah Berita, disusul dengan kategori Info Kampus, selanjutnya Profil Akakom, Program Studi/Jurusan, Lainnya, dan kategori terakhir adalah Arsip Umum.

Untuk kepentingan pengembangan di masa yang akan datang dapat dilakukan analisis statistik yang lebih detail berkaitan dengan hal sebagai berikut.

- 1) pengidentifikasian IP pengakses untuk dapat mengenali asal pengguna, sehingga dapat menganalisis perilaku pengguna web
- 2) pengidentifikasian waktu akses, sehingga dapat dianalisis waktu-waktu yang sering digunakan untuk pengelolaan jaringan
- 3) pengidentifikasian lama waktu akses, sehingga dapat menunjukkan retensi pengguna pada informasi-informasi yang tersedia.

## REFERENSI

- [1] Suratno, T., Prahasto, T., Rochim, A.F., 2011, Web Usage Mining, Pattern Discovery dan Log File, Jurnal Sistem Informasi Bisnis 02, Universitas Diponegoro, Semarang
- [2] Hermanto, A., Sartono, Sandra E., \_, Pemanfaatan Web Usage Mining pada e-commerce untuk Peningkatan Penjualan pada Retailer Online, Universitas 17 Agustus Surabaya 1945, Surabaya
- [3] Indrayani, E., 2011, Pengelolaan Sistem Informasi Akademik Perguruan Tinggi Berbasis Teknologi Informasi dan Komunikasi (TIK), Jurnal Penelitian Pendidikan, vol. 12 no.1, Universitas Pendidikan Indonesia, Bandung
- [4] Tyagi, N. Kumar; Solanki; Wadhwa, M., 2010. Analysis of server log by web usage mining for website improvement. International Journal of Computer Science, Vol. 7, Issue 4, No 8
- [5] Srivastava, J., Cooley, R., Deshpande, M., Tan, P., 2000. Web usage mining: discovery and applications of usage patterns from web data. ACM SIGKDD, 1(2):1–12.
- [6] The Apache Software Foundation, 2017, Apache HTTP Server Version 2.4, <https://httpd.apache.org/docs/2.4/logs.html>
- [7] Kernighan, Brian, 2009, gawk - pattern scanning and processing language, Free Software Foundation, Inc, Dokumentasi Linux
- [8] Vrána, J., 2017, Adminer, Database management in a single PHP file, <https://www.adminer.org/en/>.
- [9] Davis, Gordon B., 1993, Kerangka Dasar Sistem Informasi Manajemen Bagian I Pengantar, PT. Pustaka Binaman Pressindo, Jakarta.