

APLIKASI E-ARSIP PADA STMIK PALCOMTECH PALEMBANG

Atin Triwahyuni

Prodi Teknik Elektro Sekolah Pascasarjana Universitas Gadjah Mada Yogyakarta
Jln. Balap Sepeda Lr. Muhajirin IV No. 1594 Rt. 28 Palembang 30137 Telp. (0711) 373833
Email : ukthi_atr@yahoo.com

Abstract

Student administration archiving system of STMIK PalComTech Palembang, mostly still concentrated in the BAAK (Administrative Bureau of Academic Affairs). And without realizing it, will affect to the performance of administrative staff which finally can cause the service to the academic community to be slow. With the information technology, expected part of the burden of administration staff in providing information services contributed. So that activities that should be the duties and responsibilities of administrative staff can be carried out optimally. For this reason, student administration archiving system created digitally (e-archive), using OCR system (Optical Character Recognition) and string search methods (String Matching). For analysis and design, we use PHP as the programming language and MySQL as the database. This study aims to develop a digital archiving system (e-archive) with the help of the shelter database as the data, thus making the performance of administrative staff, especially BAAK getting better and professional in service quality.

Keywords : E-Archive, Digital, OCR, String Matching

1. PENDAHULUAN

Sistem pengarsipan administrasi mahasiswa yang saat ini berjalan di STMIK PalComTech Palembang dilakukan dengan menyimpan arsip dalam bentuk *hard copy* dan disimpan dalam lemari arsip. Hal ini menjadikan *space* penyimpanan ruangan memakan tempat. Di sisi lain pencarian arsip tersebut membutuhkan waktu yang relatif lama.

Sistem pengarsipan administrasi mahasiswa STMIK PalComTech sebagian besar masih terkonsentrasi di bagian BAAK (Biro Administrasi Akademik Kemahasiswaan) sehingga secara tidak langsung akan berpengaruh kepada kinerja tenaga administrasi dan muaranya adalah pelayanan kepada civitas akademika menjadi lambat, sehingga dengan bantuan teknologi informasi diharapkan dapat memudahkan sebagian beban tenaga administrasi dalam memberikan pelayanan informasi dalam hal ini sistem pengarsipan administrasi mahasiswa yang dibuat digital (*e-archive*), sehingga aktifitas yang seharusnya menjadi tugas dan tanggung jawab tenaga administrasi dapat dilaksanakan dengan optimal.

Penelitian ini membuat pengarsipan administrasi mahasiswa secara digital (*e-archive*) pada STMIK PalComTech Palembang dengan menggunakan sistem OCR (*Optical Character Recognition*) dengan metode pencarian string (*String Matching*) dimana sebagai *tool* dalam analisis dan desain mempergunakan PHP sebagai bahasa pemrogramannya dan MySQL sebagai basis datanya.

Penelitian ini diharapkan dapat memberikan solusi bagi perguruan tinggi untuk mengolah sistem pengarsipan administrasi mahasiswa yang dibuat digital (*e-archive*) dengan baik sebagai layanan informasi akademik, diharapkan juga dapat memberikan masukan kepada praktisi dalam mengolah sistem pengarsipan administrasi mahasiswa yang dibuat digital (*e-archive*).

Penelitian ini juga bertujuan untuk membangun sebuah sistem pengarsipan digital (*e-archive*) dengan bantuan basis data sebagai tempat penampungan data, sehingga membuat unjuk kerja *staff* administrasi khususnya BAAK semakin membaik dan profesional dalam pelayanannya.

2. TINJAUAN PUSTAKA

2.1. Business Process Reengineering (BPR)

BPR mengusahakan terobosan-terobosan, tidak dengan meningkatkan proses yang ada, tetapi dengan membuang proses yang sudah tidak digunakan dan menggantinya dengan yang baru sama sekali. Masalah yang dihadapi perusahaan bukanlah akibat dari struktur organisasi, tetapi struktur proses.

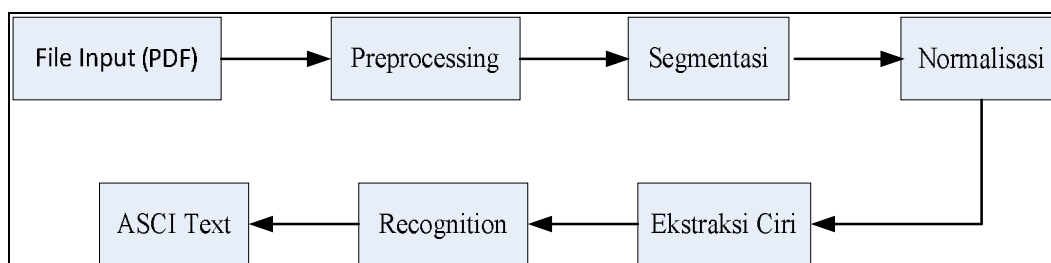
Hammer dan Champy (1990) mendefinisikan proses bisnis sebagai sekumpulan aktivitas yang meliputi satu atau lebih *output* dan menciptakan sebuah *output* yang bernilai bagi pelanggan. Beberapa karakteristik BPR antara lain:

- 1) Fundamental: dalam melaksanakan rekayasa ulang harus ditanyakan hal-hal yang paling mendasar tentang operasi perusahaan. Karena menanyakan hal yang mendasar ini akan memaksa orang untuk melihat aturan-aturan tak tertulis dan asumsi yang mendasari cara penyelenggaraan bisnis yang kadang sudah terasa asing dan tidak sesuai.
- 2) Radikal: merancang ulang secara radikal berarti memulai dari akar permasalahan dan tidak hanya mengadakan perubahan pada proses yang sudah ada.
- 3) Dramatis: rekayasa ulang bukanlah usaha untuk mencapai peningkatan secara marjinal/*incremental*, tetapi untuk pencapaian suatu lompatan besar mengenai kinerja perusahaan.
- 4) Proses: sebagian besar kalangan bisnis tidak berorientasi pada proses tetapi mereka memusatkan perhatian pada tugas-tugas, pekerjaan, orang-orang, dan struktur.

2.2. OCR (*Optical Character Recognition*)

Sistem OCR adalah sebuah sistem komputer yang dapat membaca huruf, baik yang berasal dari sebuah pencetak (*printer* atau mesin ketik) maupun yang berasal dari tulisan tangan.

OCR dapat dipandang sebagai bagian dari pengenalan otomatis yang lebih luas yakni pengenalan pola otomatis (*automatic pattern recognition*). Cara kerja sistem OCR dapat dilihat pada Gambar 2.1.



Gambar 2.1 Cara Kerja Sistem OCR

2.3. Algoritma Pencarian *String* (*String Matching*)

Permasalahan pencocokan *string* (*string matching*) merupakan permasalahan yang sangat terkenal dalam dunia informatika. Contoh implementasi dari permasalahan pencocokan *string* adalah pada pencocokan sebuah *string* pada *Microsoft Word* atau editor, atau dalam kasus yang lebih besar lagi, yaitu pencocokan *website* dengan memasukkan kata-kata kunci sebagaimana yang telah diimplementasikan pada *search engine*, seperti *Yahoo* atau *Google*.

Masalah Utama dalam pencarian *string* adalah untuk mencari sebuah *string* yang terdiri dari beberapa karakter (yang biasa disebut pola) dalam sejumlah besar teks.

3. METODE PENELITIAN

3.1. Jalan Penelitian

Penelitian ini merupakan penelitian lapangan dan bersifat eksperimen yaitu membangun aplikasi perangkat lunak sistem manajemen *e-archieve* (*electronic archieve*) dengan objek penelitian pada STMIK PalComTech khususnya pada Divisi BAAK.

Untuk menyelesaikan penelitian tersebut, penulis akan membagi jalannya penelitian kedalam tahapan-tahapan tertentu. Tahapan penelitian yang digunakan dalam penelitian ini adalah sebagai berikut: 1) pengumpulan data, 2) tahap perancangan, 3) tahap implementasi sistem, dan 4) tahap pengujian sistem. Gambar 3.1 menggambarkan tahapan-tahapan yang dilakukan pada keseluruhan penelitian ini.

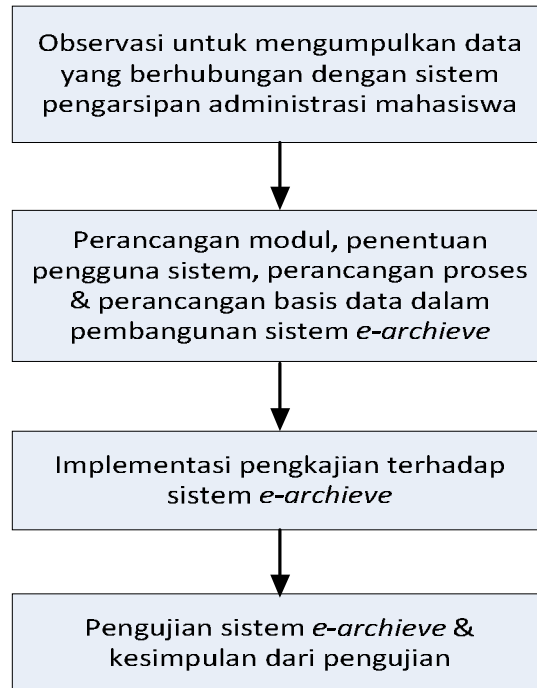
3.2. Pengujian Sistem

Pada tahap ini dilakukan pengujian sistem yang telah dibuat dengan memasukkan dokumen kertas yang sudah di *scan* menjadi dokumen PDF kemudian diproses dengan menggunakan OCR menjadi dokumen teks dan di *konversi* (*ekstraksi*) dengan menggunakan algoritma pencarian *string* (*string matching*) suatu *input* berupa data-data akademik dan memeriksa apakah sistem mampu secara aktif mengontrol proses *input-output* data

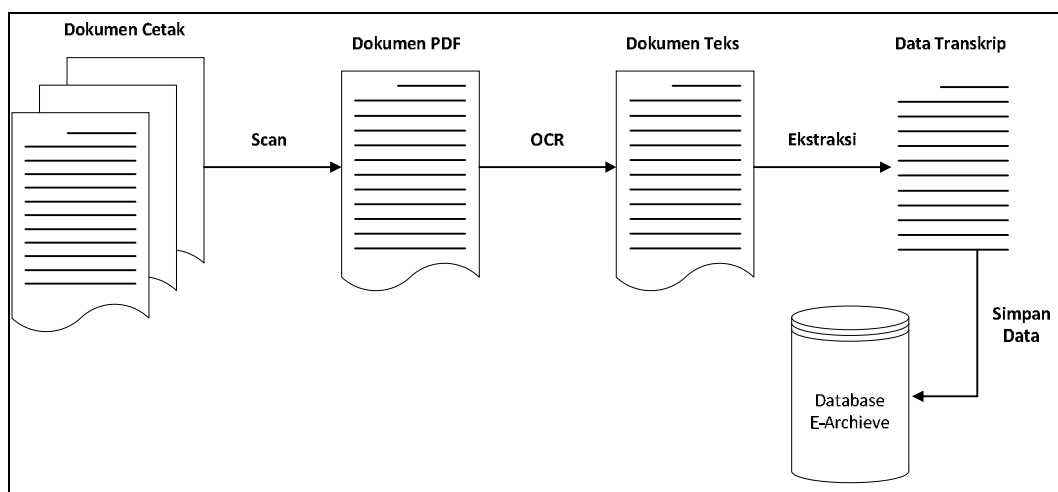
dalam *e-archie* dan kemudian data disimpan di dalam database. Proses tahap pengujian ini dapat dilihat pada Gambar 3.2.

Hasil pengujian tersebut dibandingkan dengan mekanisme pengarsipan dengan menggunakan sistem yang ada dengan pola yang lama diantaranya adalah:

- a) Menghemat kertas.
- b) Menghemat tempat penyimpanan arsip.
- c) Menghemat waktu untuk pengetikan ulang dokumen yang panjang dan dalam pencarian data arsip yang dibutuhkan.



Gambar 3.1 Tahapan Pelaksanaan Penelitian



Gambar 3.2 Tahapan Proses Pengujian

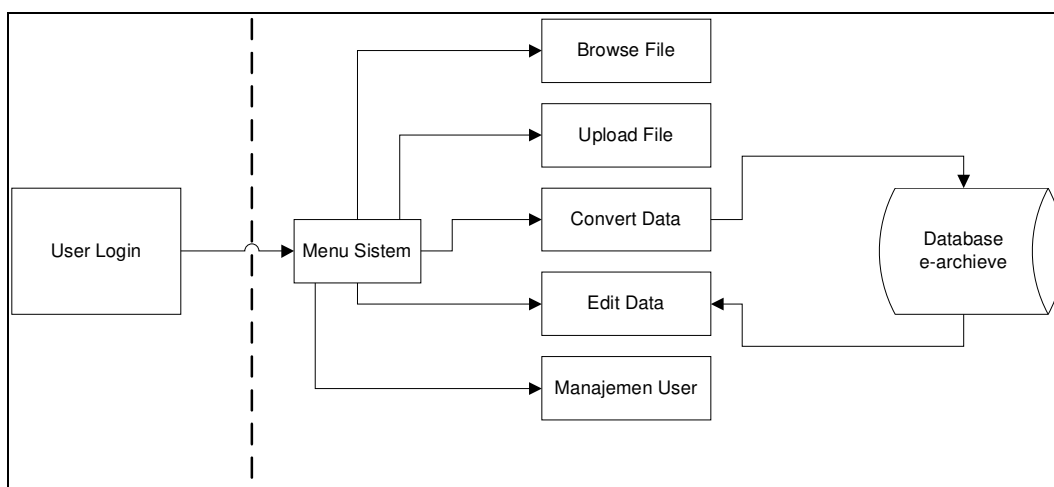
4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil dari penelitian ini berupa sebuah sistem pengarsipan elektronik (*e-archieve*). Sistem ini berbasis *web* dengan bantuan sebuah *Database Management System* (DBMS) yang diharapkan dapat memudahkan pencatatan dalam sistem dan proses pencarian berkas. Sistem juga dilengkapi dengan *user login* yang membatasi otoritas *user* yang berhak menjalankan sistem.

Dari sisi arsitektur aplikasi sistem pengarsipan elektronik dapat digambarkan dalam diagram seperti Gambar 4.1.

Proses untuk melakukan pengujian terhadap sistem yang telah dikembangkan dilakukan dengan cara mengekstrak file transkrip yang telah di-*scan*. *File* hasil *scan* selanjutnya dibaca dengan program bantu untuk menghasilkan *file* teks. *File* transkrip yang diolah sistem ini menggunakan 2 model desain yang berbeda dengan harapan dari perbedaan desain ini apakah sistem tetap mampu menghasilkan data ekstraksi sesuai dengan yang diinginkan.



Gambar 4.1 Arsitektur Aplikasi

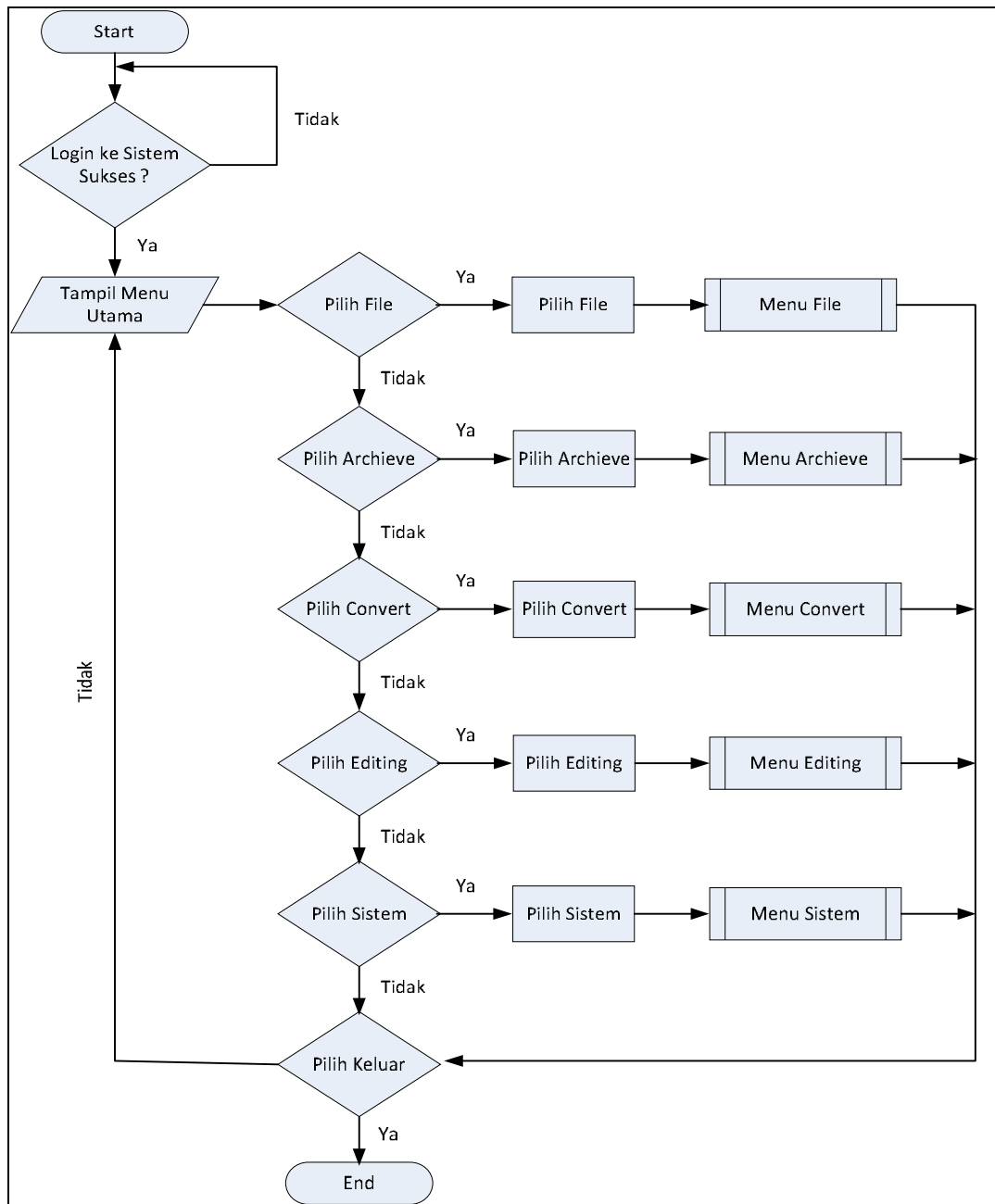
Flowchart program menu utama dapat dilihat pada Gambar 4.2 yang menggambarkan alur untuk menjalankan sistem. Pada saat aplikasi dijalankan sistem menampilkan menu *login* untuk melakukan *verifikasi user* yang akan menjalankan aplikasi. Jika proses *verifikasi user* berhasil, maka akan dimunculkan menu utama. Selanjutnya *user* dapat mengklik menu yang ada di dalam menu utama untuk mengolah data *e-archieve*.

4.2. Pembahasan

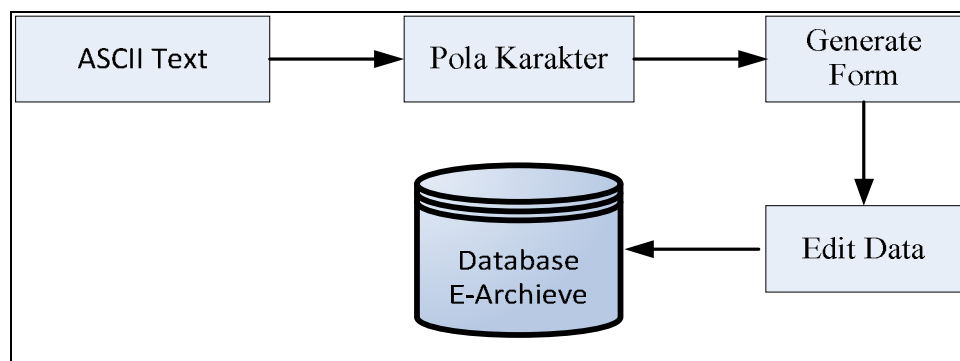
Proses ekstraksi teks dari ASCII teks diambil pola karakter untuk data yang diambil, misalnya: segmen nama diambil, segmen nilai diambil, dan seterusnya. Kemudian pola karakter dari hasil ekstraksi di *generate form* dan dimasukkan ke dalam *form*, kemudian data yang kurang cocok di *edit* lalu disimpan ke dalam database, seperti terlihat pada Gambar 4.3.

Proses ekstraksi teks ini dengan cara mengambil segmen-segmen tertentu dari sekumpulan teks berupa ASCII Teks. Untuk menemukan pola karakter dari setiap segmen menggunakan algoritma *string matching*. Setelah sekumpulan segmen diperoleh, selanjutnya dengan bantuan objek *form* dari halaman web ditampilkan segmen-segmen hasil proses ekstraksi teks.

Pengambilan segmen didasarkan pada pola yang telah didefinisikan sebelumnya. Jadi misalnya pola yang dicari adalah kata "saya", maka hanya kata yang mengandung "saya" saja yang terambil. Karena di dalam pengambilan segmen setiap karakter di dalam pola dipertimbangkan. Jadi hanya benar-benar kata yang berisi pola yang telah didefinisikan saja yang terambil sebagai segmen.



Gambar 4.2 Flowchart Menu Utama



Gambar 4.3 Diagram Ekstraksi Teks ASCII

4.3. Analisis

1) Proses *Scanning*

Ditampilkan proses yang bersih dan yang kotor dimana tidak ada karakter yang menjadi bermasalah, misalnya: PKL jadi PKU. Untuk mengatasi data yang kotor dapat diperbaiki melalui *form* yang berhasil di *generate*. Proses pembersihan data dilakukan secara manual dalam arti memperbaiki langsung dengan mengubah data yang salah melalui *form*.

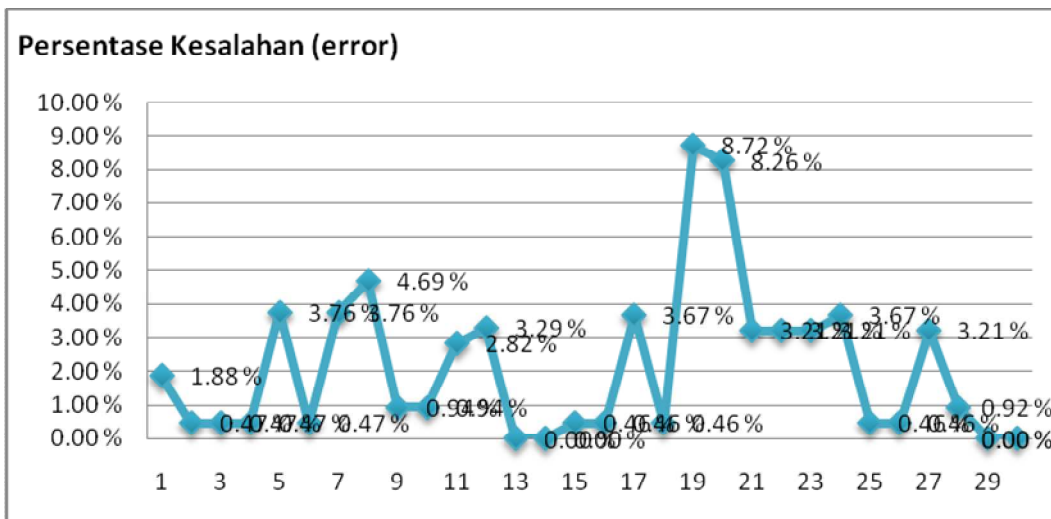
2) Proses Ekstraksi *File*

Pengambilan data atau segmen yang akan digunakan pada saat akan ditampilkan ke dalam *form*, misalnya: pada saat mengambil segmen nama maka data segmen nama yang muncul, dan seterusnya. Ketelitian hasil *file* teks yang dihasilkan dari proses ekstraksi ini diperoleh dengan mengambil segmen setiap karakter dari pola yang telah ditetapkan. Sehingga memungkinkan adanya kata yang tidak *valid* atau tidak benar secara *semantic*. Kekurangan ini dapat diminimalisir dengan menyediakan vasilitas koreksi *error* dalam sistem, sehingga data yang disimpan di dalam database nantinya merupakan data yang benar-benar *valid*. Kesalahan hasil ekstraksi *file* teks terutama karena data dari *file* teks memang terdapat *noise* sebelum proses ekstraksi dilakukan. Pengujian pada proses ekstraksi *file* ini dapat dilihat pada Gambar 4.4 dengan data uji sebanyak 30 dan rata-rata tingkat kesalahannya (*error*) adalah 2,14 %. Dengan perhitungan rumus persentase kesalahan (*error*) dapat dilihat pada persamaan (4.1).

3) Proses Penyimpanan

Menyimpan data yang dari *form* ke database, prosesnya dapat berjalan dengan baik.

$$\text{Persentase Kesalahan (Error)} = \frac{\sum \text{Segmen Salah}}{\sum \text{Segmen Total}} \times 100 \% \dots (4.1)$$



Gambar 4.4 Grafik Persentase Kesalahan (*Error*)

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Proses ekstraksi *file* teks masih memiliki kekurangan untuk menghasilkan pemrosesan *error* secara otomatis. Karena ekstraksi *file* hanya membaca setiap karakter yang ada dalam *file* teks tanpa mempertimbangkan kebenaran yang dikandung dalam setiap kata. Dari hasil pengujian diperoleh, persentase kesalahan (*error*) rata-rata 2,14 %.

5.2. Saran

Untuk penelitian selanjutnya dapat dikembangkan teknik konversi data dengan bentuk format yang berbeda, agar bisa digunakan di perguruan tinggi manapun, serta dapat lebih menekan persentase kesalahan (*error*) pada hasil dokumen yang di *scan*.

DAFTAR PUSTAKA

- Hammer, M. and Champy, J. (1990), *Reengineering the Corporation*, Massachussetts, Hammer and Co.
- Budi (2008), *Sistem OCR*, <http://standardisasi.wordpress.com/tag/sistem-ocr/>. Diakses tanggal 3 Juni 2010
- Fathansyah (2004), *Buku Teks Komputer Sistem Basis Data Lanjutan Buku Basis Data*, Penerbit Informatika, Bandung.
- Kendall, Kenneth E., Julie E. Kendall, (2003), *Analisis dan Perancangan Sistem*, Alih Bahasa Thamir Abdul Hafedh, Jakarta, Gramedia.
- Piazza, P., (2008), *The Basics of OCR*,
http://printscan.about.com/cs/scanners/a/about_ocr.htm. Diakses tanggal 3 Juni 2010
- Robles M., (2010), *The E-Office : What Exactly Is It?*,
[http://www.findarticles.com/p/articles/mi_m0FAU/is_6_18/ai_76295488]. Diakses tanggal 20 Juni 2010.
- Sedarmayanti (2003), *Tata Kearsipan dengan Memanfaatkan Teknologi Modern*, Bandung, CV. Mandar Maju.