

**EVALUASI PENGGUNAAN ISTILAH ASING BIDANG ILMU
KOMPUTER PADA KARYA ILMIAH AKADEMISI**

SKRIPSI

MUHAMMAD IMAM MUTHAQIN

121402114



**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA**

2017

EVALUASI PENGGUNAAN ISTILAH ASING BIDANG KOMPUTER

PADA KARYA ILMIAH AKADEMISI

SKRIPSI

Dijaukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah

Sarjana Teknologi Informasi

MUHAMMAD IMAM MUTHAQIN

121402114



PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA

MEDAN

2017

PERSETUJUAN

Judul : EVALUASI PENGGUNAAN ISTILAH ASING BIDANG ILMU
KOMPUTER PADA KARYA ILMIAH AKADEMISI

Kategori : SKRIPSI

Nama : MUHAMMAD IMAM MUTHAQIN

Nomor Induk Mahasiswa : 121402114

Program Studi : S1 TEKNOLOGI INFORMASI

Departemen : TEKNOLOGI INFORMASI

Fakultas : FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI
INFORMASI

Komisi Pembimbing :

Pembimbing 2

Pembimbing 1

Dr. Maya Silvi Lydia B.Sc., Msc
NIP. 19740127 200212 2 001

Dani Gunawan, ST., M.T
NIP. 19820915 201212 1 002

Diketahui/disetujui oleh
Program Studi S1 Teknologi Informasi
Ketua,

Romi Fadillah Rahmat B.Comp.Sc., M.Sc
NIP. 19860303 201012 2 003

PERNYATAAN

**EVALUASI PENGGUNAAN ISTILAH ASING BIDANG ILMU KOMPUTER
PADA KARYA ILMIAH AKADEMISI**

SKRIPSI

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, 20 Juli 2017

Muhammad Imam Muthaqin
121402114

UCAPAN TERIMA KASIH

Puji dan syukur penulis sampaikan ke hadirat Allah SWT yang telah memberikan rahmat serta restu-Nya sehingga penulis dapat menyelesaikan skripsi ini sebagai syarat untuk memperoleh gelar Sarjana Teknologi Informasi. Adapun beberapa nama yang penulis ingin ucapkan terima kasih antara lain.

1. Dekan Fakultas Ilmu Komputer dan Teknologi Informasi USU Bapak Prof. Dr. Opim Salim Sitompul M.Sc.
2. Ketua Jurusan Teknologi Informasi USU Bapak Romi Fadillah Rahmat B.Comp.Sc., M.Sc.
3. Sekretaris Jurusan Teknologi Informasi USU Ibu Sarah Purnamawati, ST., M.Sc.
4. Pembimbing I Bapak Dani Gunawan ST., MT. Pembimbing II Ibu Dr. Maya Silvy Lydia B.Sc., M.Sc.
5. Dosen – Dosen Teknologi Informasi yang telah memberikan ilmu selama perkuliahan.
6. Orang Tua Penulis yaitu Bapak Solli Barkan B.Sc dan Ibunda Dra. Iriani Jali SE., M.Sc.

Penulis tentu tidak lupa untuk mengucapkan terima kasih banyak atas semua dukungan teman – teman dalam penyelesaian tugas akhir dan kehidupan kampus antara lain, Indra Charisma S.Kom, Atras Najwan S.Kom, Maliki Khoirul Ilman S.Kom, Daniel Bonoffi S.Kom, Muhammad Wardana S.Kom, Reza Ramadiansyah S.Kom, Novira Indryani Ginting, Siti Fatimah, Anggi Saragih, Siti Hazizah Hrp, Yana Trisha Andini, Theresia Aruan S.Kom, Harysa Octafine, Grace Lumanauw, Rona I Dona, Franco Baggio, Tito D. Pandiangan dan teman – teman lain yang tidak dapat disebutkan satu per satu. Tidak lupa terimakasih kepada adik – adik junior, Afzalur Rahmah, Annisa Adindra Daulay, Mhd Faris Pratama, Cindy Pakpahan atas bantuannya selama kegiatan kampus.

ABSTRAK

Berdasarkan Instruksi Presiden No.2 tahun 2001 mengenai pembakuan istilah asing bidang komputer ke dalam Bahasa Indonesia, dimana Menteri Negara Riset dan Teknologi dan Menteri Pendidikan Nasional wajib ikut serta dalam regulasi, untuk melihat keberhasilan dari regulasi ini perlu dilakukan mekanisme untuk mengecek karya ilmiah akademisi yang bernaung dibawah kementerian yang dimaksud. Pada penelitian ini metode yang digunakan terbagi menjadi empat tahapan. Tahapan pertama adalah tahap pengumpulan *metadata(xml)* melalui *web-apps jOAI* pada OAI-PMH, lalu dilakukan *parsing xml* untuk mengambil tag judul, tahun terbit dan alamat unduh dokumen. Tahapan kedua adalah tahap pengkonversian pdf menjadi bentuk text. Tahapan ketiga adalah *text pre-processing* yang terdiri dari proses *tokenizing*, *stop-word* dan *case-folding*. Tahapan terakhir adalah tahapan pencarian kata baku dan kata asing istilah bidang ilmu komputer berdasarkan 629 kata SPI dengan menggunakan algoritma *boyer-moore*. Pada penelitian ini menghasilkan 1156 pdf dari enam universitas negeri di Indonesia dan dengan hasil *string-matching* yang cukup baik dan menghasilkan rata-rata waktu 1.4 detik untuk setiap dokumen. Dengan kesimpulan masih banyak penggunaan istilah asing bidang ilmu komputer pada karya ilmiah akademisi pada universitas yang terkait.

Kata Kunci : *Text Mining, Harvesting, OAI-PMH, Boyer-Moore*

**EVALUATION OF THE USE OF FOREIGN TERMS IN
THE FIELD OF COMPUTER SCIENCE IN
ACADEMIC SCHOLARLY WORK**

ABSTRACT

Based on the Presidential Decree No.2 of 2001 concerning the introduction of foreign terms of computer into Indonesian, where the Minister of State for Research and Technology and Minister of National Education must participate in the regulation, to see the success of this regulation a mechanism is needed to check the academic work of academics who under of the ministry. In this study the method used is divided into four stages. The first step is to collect metadata (xml) through web-apps jOAI on OAI-PMH, then parsing xml to retrieve title tag, year of publication and address of download document. The second stage is the conversion phase of pdf into text form. The third stage is text pre-processing which consists of tokenizing process, stop-word and case-folding. The last stage is the stages of searching the word raw and foreign words term computer science field based on 629 words SPI using boyer-moore algorithm. In this study yielded 1156 documents from six state universities in Indonesia and with a good string-matching result and generating an average of 1.4 seconds per document. With the conclusion there are 240.781 words uses of foreign words of computer field on academic scholarly work in related universities.

Kata Kunci : *Text Mining, Harvesting, OAI-PMH, Boyer-Moore.*

DAFTAR ISI

	Hal.
BAB 1_PENDAHULUAN	12
1.1 Latar Belakang	12
1.2 Rumusan Masalah	14
1.3 Batasan Masalah	14
1.4 Tujuan Penelitian	15
1.5 Manfaat Penelitian	15
1.6 Metode Penelitian	15
1.7 Sistematika Penulisan	16
BAB 2_LANDASAN TEORI	18
2.1 Intruksi Presiden No.2 Tahun 2001	18
2.2 Senarai Padanan Istilah (SPI)	18
2.3 Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)	19
2.4 Dublin Core	20
2.5 Text Preprocessing	21
2.6 String Matching	22
2.6.1 Pengertian String Matching	22
2.6.2 Klasifikasi Algoritma String Matching	22
2.6.3 Teknik Algoritma String Matching	23
2.7 Algoritma Boyer-Moore	24
2.7.1 Deskripsi kerja algoritma Boyer-Moore	24
2.7.2 Good-suffix Shift	25
2.7.3 Bad-Character Shift	25
2.8 Penelitian Terdahulu	26

BAB 3 PERANCANGAN SISTEM	28
3.1 Perancangan Sistem	28
3.1.1. Tahap Harvesting Metadata.	29
3.1.2. Proses PDF Conversion to Text.	31
3.1.3. Proses Text Pre-processing	31
3.1.4. String Matching	33
3.1.5 Perancangan Sistem Bagian Depan (Front End)	34
BAB 4 IMPLEMENTASI DAN PENGUJIAN	36
4.1 Hasil Pencarian Data Provider	36
4.2 Tahapan Data Harvesting	37
4.3 Hasil Data Harvesting	38
4.3.1 Hasil Request Metadata	38
4.3.2 Tahapan Pengambilan Elemen – Elemen pada XML	40
4.3.3 Hasil Penngunduhan Dokumen	41
4.4 Proses PDF dikonversi menjadi Text.	43
4.5 Hasil Text-Preprocessing	43
4.6 Proses Pencarian Kata dengan Algoritma Boyer-Moore.	46
4.7 Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer	47
4.7.1 Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Andalas	47
4.7.2 Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Airlangga	48
4.7.3 Hasil Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Diponegoro	50
4.7.4 Hasil Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Sriwijaya	51
4.7.5 Hasil Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Negeri Sultan Syarif Kasim	52

4.7.6 Hasil Hasil Pencarian Kata Baku dan Kata Istilah Asing Bidang Ilmu Komputer Universitas Sumatera Utara	53
4.7.7 Hasil Perbandingan Data Enam Universitas	55
4.7.8. Hasil Perhitungan Waktu (detik) proses Perhitungan Kata dengan Algoritma Boyer-Moore	58
4.8 Implementasi Sistem Bagian Depan	58
BAB 5_KESIMPULAN DAN SARAN	60
5.1 Kesimpulan	60
5.2 Saran	60
DAFTAR PUSTAKA	61
LAMPIRAN	64

DAFTAR GAMBAR

Gambar 2. 1 Struktur Umum OAI-PMH (digiacademy, 2016)	20
Gambar 2. 2 Contoh format metadata pada Dublin core (dc)	21
Gambar 2. 3 Good-suffix shift, u terjadi lagi didahului karakter c berbeda dari a	25
Gambar 2. 4 Good-suffix shift, hanya suffix dari u yang terjadi lagi di pattern x	25
Gambar 2. 5 Bad-character shift, b terdapat di pattern x	26
Gambar 2. 6 Bad-character shift, b tidak ada di pattern x	26
Gambar 3. 1 Arsitektur Umum	28
Gambar 3. 2Tampilan Form jOAI Harvester	29
Gambar 3. 3 Contoh format XML dengan MetadataFormat oai-dc.....	30
Gambar 3. 4 Contoh Hasil Tokenizing	31
Gambar 3. 5 Contoh Hasil Stop-Word.....	32
Gambar 3. 6 Contoh Hasil Case Folding	32
Gambar 3. 7 Contoh Hasil Stemming	33
Gambar 3. 8 Flowchart Kinerja Algoritma Boyer-Moore	33
Gambar 3. 9 GUI Menu Utama.....	34
Gambar 3. 10 GUI Lihat Data.....	35
Gambar 4. 1 From Harvester pada web apps jOAI (terisi)	37
Gambar 4. 2 Salah Satu XML Universitas Sumatera Utara	40
Gambar 4. 3 Jumlah Dokumen Karya Ilmiah per Tahun Enam Universitas Negeri	42

Gambar 4. 4 Abstrak Dokumen dengan Judul Algoritma Ant Colony System Dalam Penjadwalan Kegiatan Belajar Mengajar Di Sekolah Dasar	43
Gambar 4. 5 Kalimat Kedua dari Abstrak dengan Judul Algoritma Ant Colony System Dalam Penjadwalan Kegiatan Belajar Mengajar Di Sekolah Dasar	43
Gambar 4. 6 Hasil Proses Tokenizing	44
Gambar 4. 7 Hasil Proses Stop-Word Filtering	44
Gambar 4. 8 Hasil Proses Case Folding	45
Gambar 4. 9 Hasil Proses Stemming	45
Gambar 4. 10 Abstrak Dokumen dengan Judul Algoritma Ant Colony System Dalam Penjadwalan Kegiatan Belajar Mengajar Di Sekolah Dasar	46
Gambar 4. 11 Grafik Perkembangan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer pertahun Universitas Andalas	48
Gambar 4. 12 Grafik Perkembangan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer pertahun Universitas Airlangga	49
Gambar 4. 13 Grafik Perkembangan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer pertahun Universitas Diponegoro	50
Gambar 4. 14 Grafik Perkembangan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing istilah Bidang Ilmu Komputer pertahun Universitas Sriwijaya	52
Gambar 4. 15 Grafik Perkembangan Jumlah Penggunaan Kata Baku dan Kata Asing istilah Bidang Komputer pertahun Universitas Negeri Sultan Syarif Kasim	53
Gambar 4. 16 Grafik Perkembangan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer pertahun Universitas Sumatera Utara	54
Gambar 4. 17 Grafik Perbandingan Keseluruhan Jumlah Penggunaan Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer Enam Universitas	55
Gambar 4. 18 Sepuluh Kata Baku Sesuai SPI Terbanyak	56
Gambar 4. 19 Sepuluh Kata Asing Istilah Bidang Ilmu Komputer Terbanyak	56
Gambar 4. 20 Menu Utama Antarmuka	58
Gambar 4. 21 Halaman Grafik Penggunaan Kata Baku dan Kata Non Baku	59

DAFTAR TABEL

Tabel 2. 1 10 Contoh Istilah Asing yang telah di Indonesiakan untuk Bidang Komputer	19
Tabel 4. 1 Hasil Pencarian Data Provider beserta URL dan URI	36
Tabel 4. 2 Hasil Request Metadata (xml)	39
Tabel 4. 3 Hasil Jumlah Pengunduhan Dokumen	41
Tabel 4. 4 Hasil Perhitungan Kata Universitas Andalas pertahun	47
Tabel 4. 5 Hasil Perhitungan Kata Universitas Airlangga pertahun	48
Tabel 4. 5 Hasil Perhitungan Kata Universitas Airlangga pertahun (lanjutan).....	49
Tabel 4. 6 Hasil Perhitungan Kata Universitas Diponegoro pertahun	50
Tabel 4. 7 Hasil Perhitungan Kata Universitas Diponegoro pertahun	51
Tabel 4. 7 Hasil Perhitungan Kata Universitas Diponegoro pertahun (lanjutan).....	51
Tabel 4. 8 Hasil Perhitungan Kata Universitas Negeri Sultan Syarif Kasim pertahun	52
Tabel 4. 9 Hasil Perhitungan Kata Universitas Sumatera Utara pertahun	54
Tabel 4. 10 Hasil Keseluruhan Data Penggunaan Kata Baku dan Kata Asing Enam Universitas Negeri..	55
Tabel 4. 11 Persentase antara Kata Baku Sesuai SPI dan Kata Asing Istilah Bidang Ilmu Komputer	57