

***QUERY EXPANSION* UNTUK MENINGKATKAN AKURASI  
JAWABAN PADA SISTEM TANYA JAWAB TERJEMAHAN  
AL-QUR'AN BAHASA INDONESIA**

*Diajukan Sebagai Syarat Untuk Menyelesaikan  
Pendidikan Program Strata-1 Pada  
Jurusan Teknik Informatika*



Oleh :

**ANNISA NURIKA  
NIM : 09021181419127**

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2018**

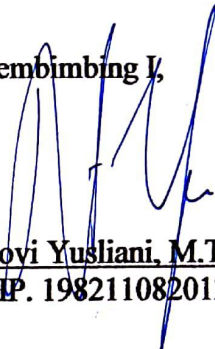
LEMBAR PENGESAHAN SKRIPSI

*QUERY EXPANSION* UNTUK MENINGKATKAN AKURASI  
JAWABAN PADA SISTEM TANYA JAWAB TERJEMAHAN AL-  
QUR'AN BAHASA INDONESIA

Oleh :

ANNISA NURIKA  
NIM : 09021181419127

Pembimbing I,

  
Novi Yusliani, M.T  
NIP. 198211082012122001

Palembang,

2018

Pembimbing II,

  
Kanda Januar Miraswan, M.T  
NIPUS. 1671080901900006

Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T  
NIP.197706012009121004

## TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Jum'at, 26 Oktober 2018 telah dilaksanakan ujian sidang skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : ANNISA NURIKA  
NIM : 09021181419127  
Judul : *Query Expansion* untuk Meningkatkan Akurasi Jawaban pada Sistem Tanya Jawab Terjemahan Al-Qur'an Bahasa Indonesia

### 1. Ketua Penguji

Novi Yusliani, M.T  
NIP. 198211082012122001



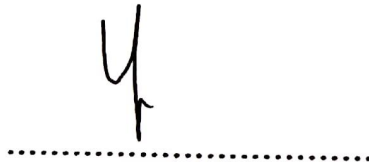
### 2. Sekretaris Penguji

Kanda Januar Miraswan, M.T  
NIPUS. 1671080901900006



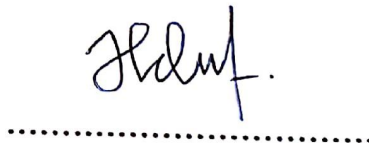
### 3. Penguji I

Yunita, M.Cs  
NIP. 198306062015042002



### 4. Penguji II

Hardini Novianti, M.T  
NIP. 197911012014042002



Mengetahui,  
Ketua Jurusan Teknik Informatika



Rifkie Primartha, M.T  
NIP.197706012009121004

## HALAMAN PERNYATAAN BEBAS PLAGIAT

Yang bertanda tangan dibawah ini :

Nama : Annisa Nurika  
NIM : 09021181419127  
Program Studi : Teknik Informatika  
Judul Skripsi : *Query Expansion* untuk Meningkatkan Akurasi Jawaban pada Sistem Tanya Jawab Terjemahan Al-Qur'an Bahasa Indonesia  
Hasil Pengecekan Software *iThenticate/Turnitin* : 16 %

Menyatakan bahwa Laporan Projek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan projek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.



Palembang, November 2018



Annisa Nurika  
NIM. 09021181419127

***Motto :***

- ❖ *Hidup adalah proses belajar dan berjuang*
- ❖ *Growing Strong!*
- ❖ *“...Janganlah bersedih, Sesungguhnya Allah bersama kita...”*  
*(Q.S At-Taubah : 40)*

*Kupersembahkan karya tulis ini kepada :*

- *Kedua Orang tua, dan Kakakku*
- *Semua Keluarga besar dan sahabatku*
- *Almamaterku*

# QUERY EXPANSION TO IMPROVE THE ANSWER ACCURACY IN THE QUESTION ANSWERING SYSTEMS OF INDONESIA QUR'AN TRANSLATION

By:  
Annisa Nurika  
09021181419127

## ABSTRACT

This study compares the accuracy of question answering system of Indonesia Qur'an translation using the Query Expansion and without the Query Expansion. The application of the Query Expansion aims to increase the level of accuracy of the system. In this study, the Expansion Query used is the Thesaurus based on co-occurrence analysis in the expansion of query queries. Tests were carried out on 90 questions with categories of questions of person / name, location and time. The comparison results show that the level of accuracy of the question and answer system using the Thesaurus has increased by 4.45%. Accuracy improvement occurs in the name / person and time question categories but cannot improve accuracy in the location question category. On average, the level of accuracy that can be achieved by the Question Answering System for Indonesian Language Translation using the Thesaurus is 90%. While the level of accuracy without using Thesaurus is 85.55%. Accuracy improvement shows that the Expansion Query method using the Thesaurus can be used to expand query queries in the Indonesian question and answer system of Al-Qur'an translations.

**Keywords:** Query Expansion, Thesaurus, Question Answering System, Indonesian Translation Qur'an.

*QUERY EXPANSION* UNTUK MENINGKATKAN AKURASI JAWABAN  
PADA SISTEM TANYA JAWAB TERJEMAHAN AL-QUR'AN  
BAHASA INDONESIA

Oleh:  
Annisa Nurika  
09021181419127

ABSTRAK

Penelitian ini membandingkan hasil akurasi sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia menggunakan *Query Expansion* dan tanpa *Query Expansion*. Penerapan *Query Expansion* bertujuan untuk meningkatkan tingkat akurasi sistem. Dalam penelitian ini, *Query Expansion* yang digunakan yaitu *Thesaurus* berdasarkan analisis *co-occurrence* dalam perluasan *query* pertanyaan. Pengujian dilakukan terhadap 90 pertanyaan dengan kategori pertanyaan orang/nama, lokasi dan waktu. Hasil perbandingan menunjukkan bahwa tingkat akurasi sistem tanya jawab menggunakan *Thesaurus* mengalami peningkatan sebesar 4,45%. Peningkatan akurasi terjadi pada kategori pertanyaan nama/orang dan waktu namun tidak dapat meningkatkan akurasi pada kategori pertanyaan lokasi. Secara rata-rata, tingkat akurasi yang mampu diraih Sistem Tanya Jawab Terjemahan Al-Qur'an Bahasa Indonesia menggunakan *Thesaurus* sebesar 90%. Sedangkan tingkat akurasi tanpa menggunakan *Thesaurus* sebesar 85,55%. Peningkatan akurasi menunjukkan bahwa metode *Query Expansion* menggunakan *Thesaurus* dapat digunakan untuk perluasan *query* pertanyaan pada sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia.

**Kata Kunci** : *Query Expansion*, *Thesaurus*, Sistem Tanya Jawab, Terjemahan Al-Quran bahasa Indonesia.

## KATA PENGANTAR



Puji dan syukur kehadiran Allah SWT karena atas rahmat-Nya penulis dapat menyelesaikan Tugas Akhir ini. Tugas akhir yang berjudul “*Query Expansion* untuk Meningkatkan Akurasi Jawaban pada Sistem Tanya Jawab Terjemahan Al-Qur’an Bahasa Indonesia” disusun untuk memenuhi salah satu syarat guna menyelesaikan pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada pihak-pihak telah memberikan dukungan, bimbingan, motivasi dan kemauan kepada penulis untuk menyelesaikan tugas akhir ini, yaitu kepada:

1. Kedua orang tuaku, Muis dan Darwati (Almarhummah), Abang Riche Fahriawan, Mba Asna Wachida, Abang Romi Ahmad Badri, Oom Zulkifli dan seluruh keluarga besarku.
2. Bapak Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Bapak Rifkie Primartha, M.T selaku Ketua Jurusan Teknik Informatika.
4. Ibu Novi Yusliani, M.T selaku dosen pembimbing I Tugas Akhir maupun dosen pembimbing akademik dan Kakak Kanda Januar Miraswan, M.T



selaku dosen pembimbing II yang telah memberikan arahan serta dukungan dalam proses pengerjaan Tugas Akhir.

5. Ibu Yunita, M.Cs selaku dosen penguji I dan Ibu Hardini Novianti, M.T selaku dosen penguji II yang telah memberikan masukan dan dorongan dalam proses pengerjaan Tugas Akhir.
6. Mba Winda selaku admin Jurusan Teknik Informatika.
7. Seluruh dosen Program Studi Teknik Informatika yang selama ini telah melimpahkan ilmunya kepada penulis selama proses belajar mengajar dan Staf dan karyawan Fakultas Ilmu Komputer atas bantuannya dalam memperlancar kegiatan akademik dan administasi selama masa perkuliahan.
8. Teman-teman Teknik Informatika Reguler (TIA) angkatan 2014, khususnya Rachma, Rosita, Azwina, Yessi, Nova, Hakim, Achmad Dwi dan lain-lain yang tidak bisa disebutkan satu persatu untuk support, bantuan, persahabatan dan masa-masa perkuliahan yang menyenangkan dan tak terlupakan.
9. Seluruh keluarga dan sahabatku di IMSAK-Sumsel (Himpunan Mahasiswa Sakti Alam Kerinci) Khususnya Bang Ari, Hilna, Elan, dan Indah yang telah memotivasi, mendukung, mewarnai hidup Penulis.
10. Seluruh Keluarga Rusunawa Unsri, khususnya Kak Ijal, Dessi, Ari, Novita, Uli, dan Vetu untuk support dan pengalaman tak terlupakan suka-duka di Lantai 5.

11. Keluarga Evo Kost, Nabila, Feni, Tila, Rahma, Sumi, dan Sutri yang selalu memberi kobaran semangat “gawe ke skripsi” dan menghibur penulis melalui candaannya.
12. Kawan nan Jauh, Mas Nur dan Yogi yang telah menyediakan waktu untuk *sharing* mengenai bidang Informatika serta memotivasi penulis.
13. Untuk semua pihak yang telah membantu dan berperan dalam penyelesaian tugas akhir ini yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa Tugas Akhir ini masih terdapat banyak kekurangan disebabkan keterbatasan pengetahuan dan pengalaman, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk penelitian selanjutnya. Akhir kata dengan selaga kerendahan hati, semoga Tugas Akhir ini dapat bermanfaat bagi kita semua

Palembang, November 2018

Penulis

## DAFTAR ISI

Halaman

HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN KOMISI PENGUJI.....	iii
HALAMAN PERNYATAAN BEBAS PLAGIAT.....	iv
HALAMAN MOTTO DAN PERSEMBAHAN.....	v
ABSTRACT.....	vi
ABSTRAKSI.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xvi
BAB I PENDAHULUAN	
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang.....	I-1
1.3 Rumusan Masalah.....	I-4
1.4 Tujuan Penelitian.....	I-5
1.5 Manfaat Penelitian.....	I-5
1.6 Batasan Masalah.....	I-5
BAB II KAJIAN LITERATUR	
2.1 Pendahuluan.....	II-1
2.2 Sistem Tanya Jawab.....	II-1
2.3 <i>Query Expansion</i> .....	II-2
2.4 <i>Thesaurus</i> .....	II-3
2.5 Analisis <i>term co-occurrence</i> .....	II-4
2.6 <i>Term Frequency/Inverse Document Frequency (tf/idf)</i> .....	II-6
2.7 Ukuran Kemiripan ( <i>Cosine Similarity</i> ).....	II-7
2.8 Al-Qur'an.....	II-8
2.8.1 Shalat.....	II-9
2.8.2 Zakat.....	II-9
2.8.3 Puasa.....	II-10
2.8.4 Haji.....	II-10
2.9 Penelitian Terkait.....	II-11
BAB III METODE PENELITIAN	
3.1 Pendahuluan.....	III-1
3.2 Data.....	III-1

3.2.1 Jenis dan Sumber Data.....	III-1
3.2.2 Metode Pengumpulan Data.....	III-1
3.3 Tahapan Penelitian.....	III-2
3.3.1 Menetapkan Kerangka Kerja / Framework .....	III-2
3.3.2 Menetapkan Kriteria Pengujian.....	III-6
3.3.3 Menetapkan Format Data Pengujian.....	III-7
3.3.4 Menentukan Alat yang digunakan dalam Pelaksanaan Penelitian.....	III-8
3.3.5 Melakukan Pengujian Penelitian.....	III-9
3.3.6 Melakukan Analisa Hasil Pengujian dan Membuat Kesimpulan.....	III-10
3.4 Metode Pengembangan Perangkat Lunak.....	III-11
3.4.1 Fase Insepsi.....	III-11
3.4.2 Fase Elaborasi.....	III-12
3.4.3 Fase Konstruksi.....	III-12
3.4.4 Fase Transisi.....	III-13
3.5 Manajemen Proyek Penelitian.....	III-13

#### BAB IV PENGEMBANGAN PERANGKAT LUNAK

4.1 Pendahuluan .....	IV-1
4.2 Fase Insepsi.....	IV-1
4.2.1 Pemodelan Bisnis.....	IV-1
4.2.2 Kebutuhan Sistem.....	IV-3
a. Fitur Praproses .....	IV-3
b. Fitur <i>Question Analyzer</i> .....	IV-3
c. Fitur <i>Query Expansion</i> dan Pembentukan <i>Thesaurus</i> .....	IV-4
d. Fitur <i>Document Retriever</i> .....	IV-4
e. Fitur <i>Answer Finder</i> .....	IV-4
4.2.3 Analisis dan Desain.....	IV-6
4.2.3.1 Analisis Perangkat Lunak.....	IV-6
a. Analisis Data.....	IV-6
b. Analisis Arsitektur Perangkat Lunak.....	IV-8
c. Analisis Masukan.....	IV-10
d. Analisis Praproses.....	IV-11
e. Analisis Pembentukan <i>Thesaurus</i> .....	IV-13
f. Analisis Proses Analisa Pertanyaan.....	IV-17
g. Analisis Perluasan <i>Query</i> menggunakan <i>Thesaurus</i> .....	IV-19
h. Analisis Proses Pengambilan Dokumen.....	IV-21
i. Analisis Proses Pencarian Jawaban.....	IV-26
4.2.3.2 Desain Perangkat Lunak.....	IV-28
1) Model <i>Use Case</i> .....	IV-28
2) Diagram Aktifitas.....	IV-33
4.3 Fase Elaborasi.....	IV-37
4.3.1 Pemodelan Bisnis.....	IV-37
4.3.1.1 Perancangan Antar Muka.....	IV-37
4.3.2 Kebutuhan Sistem.....	IV-38

4.3.3 Diagram Sequence.....	IV-39
4.3.3.1 Diagram <i>Sequence</i> Pembentukan <i>Thesaurus</i> .....	IV-40
4.3.3.2 Diagram <i>Sequence</i> Cari Jawaban.....	IV-41
4.3.3.3 Diagram <i>Sequence</i> Melakukan Praproses.....	IV-42
4.4 Fase Konstruksi.....	IV-43
4.4.2 Diagram Kelas.....	IV-43
4.4.3 Implementasi.....	IV-46
4.4.3.1 Implementasi Kelas.....	IV-46
4.4.3.2 Implementasi Antarmuka.....	IV-48
4.5 Fase Transisi.....	IV-50
4.5.1 Pemodelan Bisnis.....	IV-50
4.5.2 Kebutuhan Sistem.....	IV-50
4.5.3 Rencana Pengujian.....	IV-51
4.5.3.1 Rencana Pengujian <i>Use Case</i> Pembentukan <i>Thesaurus</i> .....	IV-51
4.5.3.2 Rencana Pengujian <i>Use Case</i> Cari Jawaban.....	IV-51
4.5.4 Implementasi.....	IV-52
4.5.4.1 Pengujian <i>Use Case</i> Pembentukan <i>Thesaurus</i> .....	IV-53
4.5.4.2 Pengujian <i>Use Case</i> Cari Jawaban.....	IV-53

## BAB V ANALISIS PENELITIAN

5.1 Pendahuluan .....	V-1
5.2 Hasil Percobaan Penelitian.....	V-1
5.2.1 Hasil Percobaan Penelitian tanpa Menggunakan <i>Thesaurus</i> .....	V-2
5.2.1.1 Hasil Pengujian Jenis Pertanyaan Lokasi.....	V-2
5.2.1.2 Hasil Pengujian Jenis Pertanyaan Waktu.....	V-3
5.2.1.3 Hasil Pengujian Jenis Pertanyaan Nama/Orang.....	V-5
5.2.2 Hasil Percobaan Penelitian Menggunakan <i>Thesaurus</i> .....	V-6
5.2.2.1 Hasil Pengujian Jenis Pertanyaan Lokasi.....	V-6
5.2.2.2 Hasil Pengujian Jenis Pertanyaan Waktu.....	V-10
5.2.2.3 Hasil Pengujian Jenis Pertanyaan Nama/Orang.....	V-13
5.3 Analisis Penelitian.....	V-16
5.3.1 Analisis Pengambilan Jawaban Hasil Penelitian.....	V-19
5.3.2 Analisis Perbandingan Tingkat Akurasi Hasil Penelitian.....	V-27

## BAB VI KESIMPULAN DAN SARAN

6.1 Pendahuluan .....	VI-1
6.2 Kesimpulan.....	VI-1
6.3 Saran.....	VI-2

DAFTAR PUSTAKA.....	xvii
LAMPIRAN.....	xix

## DAFTAR TABEL

	Halaman
III-1. Tabel Rancangan Evaluasi Sistem Tanya Jawab Terjemahan Al-Qur'an Bahasa Indonesia.....	III-9
III-2. Tabel Manajemen Proyek Penelitian dalam bentuk <i>Work Breakdown Structure</i> .....	III-15
IV-1. Kebutuhan Fungsional.....	IV-5
IV-2. Kebutuhan Non-Fungsional.....	IV-5
IV-3. Kategori Pertanyaan dan Contoh Pertanyaan.....	IV-8
IV-4. Tabel Daftar <i>Term</i> Hasil Praproses Dokumen.....	IV-13
IV-5. Tabel Penentuan Bobot <i>Term</i> .....	IV-15
IV-6. Tabel Penentuan Bobot Pasangan <i>Term</i> .....	IV-15
IV-7. Aturan Klasifikasi Tipe Jawaban (Zulen dan Purwarianti, 2011).....	IV-18
IV-8. Contoh Hasil Proses Analisa Pertanyaan.....	IV-19
IV-9. Contoh Hasil Proses <i>Query Expansion</i> .....	IV-20
IV-10. Tabel Contoh <i>Query</i> dan Dokumen.....	IV-21
IV-11. Tabel Daftar <i>Term</i> Hasil Praproses.....	IV-22
IV-12. Tabel Penentuan Bobot <i>Term</i> .....	IV-23
IV-13. Tabel Perhitungan <i>Cosine Similarity</i> .....	IV-25
IV-14. Tabel hasil perhitungan <i>Cosine Similarity</i> .....	IV-26
IV-15. Contoh Hasil Komponen <i>Answer Finder</i> .....	IV-28
IV-16. Definisi Aktor.....	IV-29
IV-17. Definisi <i>Use Case</i> .....	IV-30
IV-18. Skenario Pembentukan <i>Thesaurus</i> .....	IV-31
IV-19. Skenario Cari Jawaban.....	IV-32
IV-20. Skenario Melakukan praproses.....	IV-33
IV-21. Implementasi Kelas.....	IV-46
IV-22. Rencana pengujian untuk <i>Use Case</i> Pembentukan <i>Thesaurus</i> .....	IV-51
IV-23. Rencana Pengujian <i>Use Case</i> Cari Jawaban.....	IV-51
IV-24. Pengujian <i>Use Case</i> Pembentukan <i>Thesaurus</i> .....	IV-53
IV-25. Pengujian <i>Use Case</i> Cari Jawaban.....	IV-54
V-1. Hasil Pengujian Jenis Pertanyaan Lokasi.....	V-2
V-2. Hasil Pengujian Jenis Pertanyaan Waktu.....	V-4
V-3. Hasil Pengujian Jenis Pertanyaan Nama/Orang.....	V-5
V-4. Hasil Pengujian Jenis Pertanyaan Lokasi.....	V-6
V-5. Hasil Pengujian Jenis Pertanyaan Waktu.....	V-10
V-6. Hasil Pengujian Jenis Pertanyaan Nama/Orang.....	V-13
V-7. Hasil Pengujian Sistem Tanya Jawaban Terjemahan Al-Qur'an Bahasa Indonesia Tanpa Menggunakan <i>Thesaurus</i> Berdasarkan Jawaban berupa Dokumen Ayat dan Kalimat Jawaban.....	V-16
V-8. Hasil Pengujian Sistem Tanya Jawaban Terjemahan Al-Qur'an Bahasa Indonesia Tanpa Menggunakan <i>Thesaurus</i> Berdasarkan Jawaban berupa Dokumen Ayat .....	V-17

V-9. Hasil Pengujian Sistem Tanya Jawaban Terjemahan Al-Qur'an Bahasa Indonesia Menggunakan <i>Thesaurus</i> Berdasarkan Jawaban berupa Dokumen Ayat dan Kalimat Jawaban.....	V-17
V-10. Hasil Pengujian Sistem Tanya Jawaban Terjemahan Al-Qur'an Bahasa Indonesia Menggunakan <i>Thesaurus</i> Berdasarkan Jawaban berupa Dokumen Ayat.....	V-18
V-11. Perbandingan Jawaban Pada Sistem Menggunakan <i>Query Expansion</i> .	V-27

## DAFTAR GAMBAR

	Halaman
III-1. Kerangka Kerja Sistem.....	III-3
III-2. Penjadwalan untuk Tahap Menentukan Ruang Lingkup dan Unit Penelitian.....	III-20
III-3. Penjadwalan untuk Tahap Menentukan Dasar Teori yang Berkaitan dengan Penelitian dan Menentukan Kriteria Pengujian.....	III-21
III-4. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Insepsi.....	III-21
III-5. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Elaborasi.....	III-22
III-6. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Konstruksi.....	III-22
III-7. Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Transisi.....	III-23
III-8. Penjadwalan untuk Tahap Melakukan Pengujian Penelitian, Analisa Hasil Pengujian Penelitian dan Membuat Kesimpulan.....	III-23
IV-1. Diagram proses pencarian jawaban sistem tanya jawab bahasa Indonesia yang mengimplementasikan <i>Query Expansion</i> menggunakan <i>Thesaurus</i> .....	IV-9
IV-2. Proses <i>Query Expansion</i> .....	IV-20
IV-3. Diagram <i>Use Case</i> .....	IV-29
IV-4. Diagram Aktivitas Pembentukan <i>Thesaurus</i> .....	IV-34
IV-5. Diagram Aktivitas Cari Jawaban.....	IV-35
IV-6. Diagram Aktivitas Melakukan Praproses.....	IV-36
IV-7. Rancangan Antarmuka Menu Utama.....	IV-37
IV-8. Rancangan Antarmuka Hasil Pencarian Jawaban.....	IV-38
IV-9. Diagram <i>Sequence</i> Pembentukan <i>Thesaurus</i> .....	IV-40
IV-10. Diagram <i>Sequence</i> Cari Jawaban.....	IV-41
IV-11. Diagram <i>Sequence</i> Melakukan Praproses.....	IV-42
IV-12. Diagram Kelas Perangkat Lunak ExpansionQAS.....	IV-44
IV-13. Diagram Kelas Perangkat Lunak Pembentukan <i>Thesaurus</i> .....	IV-45
IV-14. Antarmuka Menu Utama.....	IV-49
IV-15. Antarmuka Hasil Pencarian Jawaban.....	IV-49
V-1. Contoh Hasil Pengujian Pertanyaan Waktu tanpa Menggunakan <i>Thesaurus</i> .....	V-20
V-2. Kunci Jawaban.....	V-20
V-3. Contoh Hasil Pengujian Pertanyaan Waktu Menggunakan <i>Thesaurus</i> .....	V-22
V-4. Contoh Hasil Pengujian Pertanyaan Orang Menggunakan <i>Thesaurus</i> .....	V-24
V-5. Contoh Hasil Pengujian Pertanyaan Lokasi Menggunakan <i>Thesaurus</i> .....	V-26



# **BAB I**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Pada bab ini membahas latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian serta batasan masalah. Bab ini akan memberikan penjelasan umum mengenai keseluruhan penelitian.

Pendahuluan dimulai dengan penjelasan mengenai tantangan dan tujuan proses menemukan pengetahuan baru pada sistem tanya jawab. Serta penelitian yang berkaitan dengan menerapkan *Query Expansion* yang menjadi latar belakang dari penelitian ini.

### **1.2 Latar belakang**

Sistem tanya jawab adalah sebuah kajian dalam bidang *Natural Language Processing* yang dapat dijadikan sebagai media dalam memberikan informasi mengenai suatu topik yang diperlukan oleh pihak tertentu. Sistem tanya jawab tidak hanya mengembalikan dokumen sebagai informasi namun sistem secara otomatis dapat memberikan jawaban yang relevan berdasarkan *query* pertanyaan yang diajukan oleh penggunaanya menggunakan bahasa alami (Gusmita et al., 2014).

Penelitian mengenai sistem tanya jawab bahasa Indonesia topik bidang keagamaan telah dilakukan oleh Anggraeny (2007) dan Gusmita et al. (2014).

Mereka melakukan penelitian mengenai sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia, namun hasil dari penelitian tersebut belum mampu memberikan tingkat akurasi yang maksimal, sehingga penelitian untuk meningkatkan akurasi sistem tanya jawab terjemahan Al-Quran bahasa Indonesia masih dapat diteliti. Anggraeny (2007) menggunakan metode *rule-based* untuk menemukan jawaban. Tingkat akurasi yang dihasilkan yaitu 53,14%. Gusmita et al. (2014) melakukan penelitian menggunakan metode *rule-based* dan *relevant document* dalam proses ekstraksi jawaban pada sistem tanya jawab. Tingkat akurasi yang didapat tidak lebih baik dari penelitian sebelumnya yaitu sebesar 53,03%.

Salah satu faktor penting dalam sistem tanya jawab adalah penggunaan *query* yang tepat dapat memberikan jawaban yang akurat dan sesuai dengan keinginan pengguna. Menurut Wibowo (2014) hal yang menentukan apakah suatu sistem mampu memberikan tingkat akurasi dan relevansi hasil pencarian dokumen jawaban adalah berdasarkan kesesuaian (presisi) antara *query* yang diberikan dan dokumen yang diperoleh dalam pencarian, sehingga pengambilan dokumen yang berisi jawaban relevan dipengaruhi oleh *query* pengguna

Sistem tanya jawab memiliki kekurangan dimana sistem tidak mampu mengembalikan jawaban yang diminta oleh pengguna karena ketidaksesuaian kata-kata yang diajukan oleh pengguna dengan kata-kata pada dokumen penjawab pertanyaan yang terdapat di dalam sistem, padahal bisa jadi *query* memiliki kesamaan kata pada dokumen jawaban (Al-Chalabi et al., 2015). Adanya kecenderungan pengguna memasukkan *query* yang kurang spesifik, menyebabkan hasil yang dikeluarkan sistem tidak sesuai dengan keinginan

pengguna (Miao et al., 2010). Untuk menutupi kekurangan ini, maka diperlukannya suatu metode untuk meningkatkan keefektifan dari *query* masukan yaitu menerapkan teknik *Query Expansion*. *Query Expansion* merupakan suatu teknik peluasan *query* dengan menambahkan *keywords* baru ke dalam *query* awal yang dirasa memiliki keterkaitan untuk memperjelas *query* yang memungkinkan untuk terambilnya dokumen berisi informasi relevan sehingga meningkatkan performansi pencarian (Saneifar et al., 2014).

Hamed and Ab Aziz (2016) menerapkan *Query Expansion* menggunakan *WordNet* Bahasa Inggris pada penelitiannya tentang sistem tanya jawab terjemahan Al-Qur'an bahasa Inggris dan klasifikasi dokumen menggunakan jaringan syaraf tiruan mampu meningkatkan akurasi jawaban. Tingkat akurasi jawaban sistem ini diuji dengan menggunakan perhitungan nilai *Recall*, *Precision* dan *F-score*, dimana tanpa *Query Expansion* memberikan hasil rata-rata nilai *Recall* 1, *Precision* 0,2 dan *F-score* sebesar 0,3, sedangkan sistem yang menggunakan *Query Expansion* mengalami peningkatan hasil rata-rata nilai *Recall* 1, *Precision* 0,8 dan *F-score* yang didapat sebesar 0,9.

Rasyidi et al. (2013) dalam penelitiannya tentang sistem temu kembali informasi *hadits* menerapkan *Query Expansion* menggunakan *Thesaurus*. Pengujian tingkat akurasi dari hasil penelitian tersebut menggunakan MAP (*Mean Average Precision*) dan *Recall*. Dalam penelitian ini adanya peningkatan performansi sistem temu kembali informasi sebelum dan sesudah menggunakan *Thesaurus* dalam proses *Query Expansion* dengan peningkatan MAP sebesar 34% dan *Recall* sebesar 43%.

Dari penelitian diatas, sistem tanya jawab menerapkan teknik *Query Expansion* dapat memberikan hasil tingkat akurasi yang lebih baik dibandingkan dengan sistem tanya jawab tanpa menggunakan *Query Expansion*. Sehingga sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia dengan menerapkan *Query Expansion* menggunakan *Thesaurus* diharapkan mampu meningkatkan kinerja sistem tanya jawab dalam menyajikan informasi lebih cepat dan akurat.

### **1.3 Rumusan Masalah**

Berdasarkan latar belakang diatas, rumusan masalah pada penelitian ini adalah bagaimana tingkat akurasi jawaban pada sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia dengan menerapkan *Query Expansion* menggunakan *Thesaurus* dalam menyajikan informasi. Untuk menyelesaikan permasalahan diatas, maka penelitian ini dibagi ke dalam beberapa pertanyaan penelitian (*research question*) yaitu:

1. Bagaimana menerapkan *Query Expansion* menggunakan *Thesaurus* dalam sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia?
2. Apakah penerapan *Query Expansion* menggunakan *Thesaurus* dapat menjadi solusi untuk meningkatkan akurasi jawaban pada sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia?

#### **1.4 Tujuan penelitian**

Tujuan dari penelitian ini adalah :

1. Menghasilkan sistem tanya jawab dengan menerapkan *Query Expansion* menggunakan *Thesaurus* yang bertujuan agar dapat meningkatkan akurasi jawaban serta efisiensi waktu dalam menemukan jawaban.
2. Menganalisis persentase akurasi jawaban terhadap sistem sebelum dan sesudah menerapkan *Query Expansion* menggunakan *Thesaurus*.

#### **1.5 Manfaat penelitian**

Adapun manfaat yang diperoleh dalam penelitian ini adalah

1. Dapat menjadi rujukan bagi peneliti lain dalam mengembangkan sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia.
2. Kemudahan dalam mencari informasi bersumber pada Al-Qur'an melalui sistem tanya jawab terjemahan Al-Qur'an bahasa Indonesia.

#### **1.6 Batasan Masalah**

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Sistem hanya menerima masukan berupa pertanyaan dalam bahasa Indonesia sesuai dengan ejaan yang disempurnakan. Sistem tidak dapat menerima masukan berupa singkatan kata.
2. Jenis pertanyaan yang akan menjadi data uji adalah pertanyaan-pertanyaan yang menanyakan lokasi, tanggal/waktu, orang/nama. Pertanyaan diawali dengan kata tanya berupa "dimana" dan "dimanakah" untuk lokasi,

“kapan” dan “kapankah” untuk tanggal/waktu, “siapa” dan “siapakah” untuk orang/nama.

3. Dokumen yang digunakan adalah terjemahan Al-Qur'an dalam bahasa Indonesia versi Kementerian Agama Republik Indonesia yang diperoleh dari situs <http://tanzil.net/trans/> dimana pada situs tersebut berisi terjemahan berbagai bahasa salah satunya Bahasa Indonesia. Dokumen memiliki format teks (\*.txt) dan tidak menyertakan informasi yang terdapat pada catatan kaki.
4. Pertanyaan yang dapat diajukan mengenai 4 topik yaitu Sholat, Puasa, Zakat, Haji&Umrah.

## DAFTAR PUSTAKA

- Al-Chalabi, H., Ray, S. & Shaalan, K. Semantic Based Query Expansion For Arabic Question Answering Systems. *Arabic Computational Linguistics (Acling)*, 2015 First International Conference On, 2015. IEEE, 127-132.
- Al-Hanan, T. P. 2009. *Al- Qur'ān Dan Terjemahannya*, Surakarta, Pustaka Al-Hanan
- Anggraeny, M. D. 2007. Implementasi Question Answering System Dengan Metode Rule-Based Pada Terjemahan Al Qur'an Surat Al Baqarah.
- Gusmita, R. H., Durachman, Y., Harun, S., Firmansyah, A. F., Sukmana, H. T. & Suhaimi, A. A Rule-Based Question Answering System On Relevant Documents Of Indonesian Quran Translation. *Cyber And It Service Management (Citsm)*, 2014 International Conference On, 2014. IEEE, 104-107.
- Hafidhuiddin, D. 2002. *Zakat Dalam Perekonomian Modern*, Gema Insani.
- Hamed, S. K. & Ab Aziz, M. J. 2016. A Question Answering System On Holy Quran Translation Based On Question Expansion Technique And Neural Network Classification. *Journal Of Computer Science*, 12, 169-177.
- Jia, K., Sun, Y. & Li, Z. Answer Extraction Based On Query Expansion In Chinese Question Answering System. *Natural Language Processing And Knowledge Engineering*, 2008. Nlp-Ke'08. International Conference On, 2008. IEEE, 1-4.
- Khafajeh, A. P. D. H., Refai, A. P. D. M. & Yousef, A. P. D. N. 2013. Building Arabic Automatic Thesaurus Using Co-Occurance Technique. *Iccmtd-2013*, 28.
- Miao, Y., Su, X. & Li, C. Improving Question Answering Based On Query Expansion With Wikipedia. *Tools With Artificial Intelligence (Ictai)*, 2010 22nd Ieee International Conference On, 2010. IEEE, 233-240.
- Nazief, B. A. A. & Adriani, M. 1996. Confix Stripping: Approach To Stemming Algorithm For Bahasa Indonesia. *Internal Publication, Faculty Of Computer Science, University Of Indonesia, Depok, Jakarta*, 41.
- Pressman, R. S. 2005. *Software Engineering: A Practitioner's Approach*, Palgrave Macmillan.

- Rahman, N. A., Bakar, Z. A. & Sembok, T. M. T. Query Expansion Using Thesaurus In Improving Malay Hadith Retrieval System. *Information Technology (Itsim)*, 2010 International Symposium In, 2010. IEEE, 1404-1409.
- Rasyidi, I., Romadhony, A. & Wibowo, A. T. Indonesian Hadith Retrieval System Using Thesaurus. *Computer, Control, Informatics And Its Applications (Ic3ina)*, 2013 International Conference On, 2013. IEEE, 285-288.
- Riloff, E. & Thelen, M. A Rule-Based Question Answering System For Reading Comprehension Tests. 2000 2000. *Association For Computational Linguistics*, 13-19.
- Saneifar, H., Bonniol, S., Poncelet, P. & Roche, M. 2014. Enhancing Passage Retrieval In Log Files By Query Expansion Based On Explicit And Pseudo Relevance Feedback. *Computers In Industry*, 65, 937-951.
- Sholikah, R. W., Arifin, A. Z., Purwitasari, D. & Fatichah, C. Co-Occurrence Technique And Dictionary Based Method For Indonesian Thesaurus Construction. 2017 2017. IEEE, 1-6.
- Wibowo, A. 2014. Peningkatan Performansi Sistem Temu Balik Informasi Dengan Metode Phrasal Translation Dan Query Expansion. Batam: Teknik Multimedia Dan Jaringan Polteknik Negeri Batam. [Http://P2m.Polibatam.Ac.Id/Wp-Content/Uploads/2012/05/Ari-Wibowo-Peningkatan-Performansi-Sistem-Temu-Balik-Informasi.Pdf](http://P2m.Polibatam.Ac.Id/Wp-Content/Uploads/2012/05/Ari-Wibowo-Peningkatan-Performansi-Sistem-Temu-Balik-Informasi.Pdf). Diakses Pada 15 Januari 2018
- Zulen, A. A. & Purwarianti, A. Study And Implementation Of Monolingual Approach On Indonesian Question Answering For Factoid And Non-Factoid Question. *Paclac*, 2011. 622-631.