

# Khojee

*A Search Engine*

By

Shahnaz Dogar

A thesis submitted in partial  
fulfillment of the requirement  
for the degree of

M.Sc. in Computer Sciences

Bahria Institute of  
Management and Computer Sciences  
Islamabad, Pakistan



Bahria Institute of  
Management and Computer Sciences  
Islamabad, Pakistan

**Khojee**  
*A Search Engine*

By

**Shahnaz Dogar**

## Acknowledgments

Dedicated to my mother...the most precious gift from the creator to the man. I would not have been possible. I remain indebted to them for all their support. I would also like to thank my teachers who helped me and encouraged me through out my tenure here in Bahra

## Acknowledgments

Without the kind help of my friends Saeed, Zafar and Sagar ...this all would not have been possible. I remain indebted to them for all their support. I would also like to thank my teachers who helped me and encouraged me through out my tenure here in Bahria.

# Abstract

Khojee is a very smart and efficient Meta-Search Engine. It calls the "infind" search engine that further calls ten of the leading Internet search engines. Each of these search engines is automatically called in parallel, and retrieves the maximum number of results each engine will allow. But the main feature of Khojee that distinguishes it from currently available search engines is that it maintains a user profile that enables its registered user to store data resulting from the user search criteria for future use. Then user can login to their account any time and can get hold of previously stored data, user also has an option that he can view the result page online or he can download a single zip file and view it offline. Main idea behind this search engine was to save user's precious time.

## Chapter 2

2. Research	3
2.1. History of Search Engines	3
2.1.1. Archie and Veronica	3
2.1.2. Robots	4
2.1.3. The first Web Directory	4
2.1.4. Spider	6
2.1.5. Search Directories	6
2.1.6. The Big-Guns	6
2.1.7. Meta-Engines	7
2.1.8. Slowing Relevancy	7
2.2. Next Generation Web Search	8
2.2.1. Introduction	8
2.2.2. Site Search and Tasks	11
2.2.2.1. The Importance of the Task	12
2.2.2.2. Metadata	13
2.2.3. An Example: Epluribus	14
2.2.4. Example: Yahoo	18
2.2.5. Integrating Search	20

# Table Of Contents

---

2.2. Other Approaches to Site Search 22

Acknowledgments ii

Abstract iii

2.4. Integration with General Web Search 26

## Chapter 1

---

1. Introduction 1

1.1. What a Search Engine is? 1

1.2. Different Search Approaches 1

## Chapter 2

---

2. Research 3

2.1. History of Search Engines 3

2.1.1. Archie and Veronica 3

2.1.2. Robots 4

2.1.3. The first Web Directory 4

2.1.4. Spider 5

2.1.5. Search Directories 5

2.1.6. The Big-Guns 6

2.1.7. Meta-Engines 7

2.1.8. Skewing Relevancy 7

2.2. Next Generation Web Search 8

2.2.1. Introduction 8

2.2.2. Site Search and Tasks 11

2.2.2.1. The Importance of the Task 12

2.2.2.2. Metadata 13

2.2.3. An Example: Epicurious 14

2.2.4. Example: Yahoo 18

2.2.5. Integrating Search 20

2.2.6. Example: Biomedical Text	20
2.3. Other Approaches to Site Search	23
2.3.1. Specialized Interfaces	24
2.3.2. Question Answering	25
2.4. Integration with General Web Search	26
4.3.4. Image and Text Quality	48
<b>Chapter 3</b>	48
<hr/>	
<b>3. Requirement Analysis</b>	<b>28</b>
3.1. Comparison with currently available solutions	29
3.2. Main Features	30
3.3. Tools & Technologies	31
3.4. Proposed Deliverables	31
5.1. Application Architecture	62
<b>Chapter 4</b>	62
<hr/>	
<b>4. Technology Overview</b>	<b>32</b>
4.1. Microsoft Active Server Pages	32
4.1.1. Introduction	32
4.1.2. A Little History	32
4.1.3. What are ASP Components?	33
4.1.4. Choosing the Language : Java, C++, or Visual Basic	36
4.1.5. How it works?	37
4.1.6. When to Use ASP?	38
4.1.7. The Environment	38
4.1.8. Choosing ASP	39
4.2. Java Servlets	39
4.2.1. What are Servlets?	39
4.2.2. A Little History	40
4.2.3. Why Use Servlets?	41
4.2.4. Enter Java Servlets to save the Day	43

4.2.5. What About Perl or Microsoft ASP?	44
4.3. Macromedia Flash	45
4.3.1. What is Flash? Why developers use flash?	45
4.3.2. Flash Animations are Faster	46
4.3.3. Size	46
4.3.4. Image and Text Quality	46
4.3.5. Flash Presentations	46
4.3.6. Vector Graphics	47
4.3.7. Raster Graphics	47

## Chapter 5

---

<b>5. Development Overview</b>	<b>49</b>
5.1. Application Architecture	49
5.2. Application Design	50
<b>6. References &amp; Literature</b>	<b>58</b>