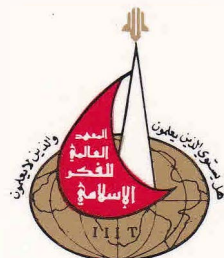


THE INTERNATIONAL INSTITUTE  
OF ISLAMIC THOUGHT



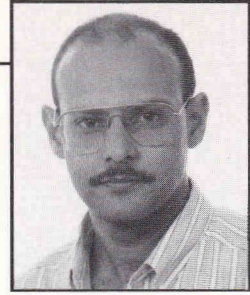
١٤٠١ هـ — ١٩٨١ م  
1401 AH—1981 AC

Academic Dissertations (2)

# Qur'anic Text: Toward a Retrieval System

Hānī M. 'Aṭīyah

## Hānī Muḥyi al Dīn 'Aṭīyah



- ◆ Born in Egypt on 10 Jumādā al Ulā 1381 AH/20 October 1961 AC.
- ◆ Received his B.Sc. in Physics from Cairo University in Ramaḍān 1404 AH/June 1984 AC.
- ◆ Received M.Sc. in Physics from Dundee University in Scotland, UK, in Ṣafar 1408 AH/June 1989 AC.
- ◆ Received Ph.D. in Islamic Informatics from University of Wales, UK, in Dhū al Qa'dah 1411 AH/June 1991 AC.
- ◆ This book is based on the dissertation he submitted for his Ph.D. degree.
- ◆ He has worked as an assistant professor at the International Islamic University in Malaysia, Department of Library and Information Science, during 1412 AH/1991 AC—1413 AH/1993 AC, where he helped to establish the Master Degree program and taught several courses in that department.
- ◆ He is presently an assistant professor at Imām Muḥammad ibn Sa'ūd Islamic University, Department of Library and Information Science.
- ◆ He has written a number of papers, has reviewed a selection of books in the field of Islamic informatics, and has had articles published in various academic journals. He has also participated in a number of conferences and seminars on various aspects of the Islamic paradigm.
- ◆ His forthcoming publications will include, in shā'a Allāh: *Elements of the Theory of Knowledge in the Qur'an* and *Classification Schemes of Early Muslim Scholars*.

*Academic Dissertations (2)*

# **Qur'anic Text: Toward a Retrieval System**

**Hānī M. 'Aṭīyah**

**International Institute of Islamic Thought  
Herndon, Virginia USA  
1417 AH/1996 AC**

Academic Dissertations (2)

© Copyrights 1416 AH/1996 AC by  
The International Institute of Islamic Thought  
555 Grove Street (P.O. Box 669)  
Herndon, Virginia 22070-4705 USA

Library of Congress Cataloging-in-Publication Data —

‘Aṭīyah, Hānī M., 1961 (1381)-

Qur’anic text : toward a retrieval system / HaHānī M. ‘Aṭīyah,  
p. x, 277 cm. 23 x 15

Includes bibliographical references (p. 25) and index.

ISBN 1-56564-212-0. -- ISBN 1-56564-213-9 (pbk.)

1. Information storage and retrieval systems--Koran. 2. Koran--  
Abstracting and indexing. 3. Islam--Sacred books--Abstracting and  
indexing. I. Title.

BP130.38.A85 1994

025.06'2971224--dc20

94-24729  
CIP

The views and opinions expressed by  
the author are not necessarily  
those of the Institute.

First Edition  
(1417 AH/1996 AC)

## *Dedication*

*To my father whose day starts with the Qur'an and ends with the Qur'an. To him I owe my academic guidance and spiritual support, which to me were the backbone throughout my life and during the preparation of this book.*

*To my mother whose blessings and prayers for me were the source that helped me to overcome many difficulties and strengthened me with more Imān.*

*To all Muslim Scholars and Information Scientists who are devoting their lives to serve the cause of Islam and who believe that the Islamization of the westernized Muslim mind is the only way ahead.*

# Contents

**Preface**

ix

**Chapter One: Introduction ..... 1**

|     |   |    |
|-----|---|----|
| 1.1 | Information Science: An Overview .....        | 1  |
| 1.2 | Information Science: An Islamic Context ..... | 2  |
| 1.3 | Early and Modern Qur'anic Indexes.....        | 13 |
| 1.4 | A Qur'anic Retrieval System .....             | 24 |
| 1.5 | Methodological Approach .....                 | 28 |

**Chapter Two: The Natural Language Problems**

**in Qur'anic Statements ..... 31**

|     |   |    |
|-----|---|----|
| 2.1 | Overview.....                           | 31 |
| 2.2 | Philological Aspects .....              | 32 |
|     | A. Ambiguous Statements.....            | 32 |
|     | B. Dis/continuous Statements .....      | 37 |
|     | C. Variant Readings.....                | 39 |
| 2.3 | Historical Aspects.....                 | 41 |
|     | A. Occasions of Revelation .....        | 41 |
|     | B. General and Specific Statements..... | 44 |
|     | C. Judaica .....                        | 47 |
| 2.4 | Theological Aspects .....               | 49 |
|     | A. The Sects.....                       | 49 |
|     | B. The Modernists .....                 | 58 |
| 2.5 | Conclusion .....                        | 61 |

**Chapter Three: Information Retrieval Problems**

**in Qur'anic Vocabularies ..... 63**

|     |   |    |
|-----|---|----|
| 3.1 | Overview .....                          | 63 |
| 3.2 | Homographs .....                        | 64 |
| 3.3 | Forms of Terms .....                    | 65 |
| 3.4 | Synonyms.....                           | 72 |
| 3.5 | Organization of Terms.....              | 76 |
|     | A. Single-word Headings .....           | 76 |
|     | B. Compound Headings .....              | 82 |
| 3.6 | The Structure and Display of Terms..... | 85 |
| 3.7 | Conclusion .....                        | 98 |

|   |            |
|---|------------|
| <b>Chapter Four: Information Technology Problems<br/>in Qur'anic Style.....</b>                   | <b>99</b>  |
| 4.1 Overview.....   | 99         |
| 4.2 Automatic Search Analysis.....  | 100        |
| A. Morphological Analysis.....  | 100        |
| B. Syntactical Analysis.....  | 102        |
| C. Semantic Analysis.....   | 108        |
| D. Knowledge-based Analysis.....  | 113        |
| 4.3 Automatic Indexing.....   | 118        |
| 4.4 Conventional Searching Devices.....   | 123        |
| 4.5 Conclusion.....   | 137        |
| <br>  |            |
| <b>Chapter Five: The Plan, Construction, and Design<br/>of the Qur'anic Retrieval System.....</b> | <b>139</b> |
| 5.1 Overview.....   | 139        |
| 5.2 Abstract Forms and Index Cards.....   | 139        |
| A. Abstracting Forms.....   | 140        |
| B. Indexing Cards.....  | 145        |
| 5.3 Construction Techniques.....  | 148        |
| A. Selection of Terms.....  | 148        |
| B. Compilers' Guidelines.....   | 156        |
| 5.4 Scheme Policy.....  | 166        |
| A. Exegetical Instructions.....   | 166        |
| B. Indexing Rules.....  | 169        |
| 5.5 Displays.....   | 175        |
| A. Printed Display.....   | 175        |
| B. Screen Display.....  | 179        |
| 5.6 Conclusion.....   | 181        |
| <br>  |            |
| <b>Chapter Six: Two Examples of Qur'anic Retrieval<br/>System: Display and Organization.....</b>  | <b>183</b> |
| 6.1 Overview.....   | 183        |
| 6.2 Example No. 1: Physics.....   | 184        |
| 6.3 Example No. 2: Family.....  | 195        |
| <br>  |            |
| <b>Bibliography.....</b>  | <b>253</b> |

| Appearance                   | Pronunciation |   |
|------------------------------|---------------|---|
| <b>Consonants</b>            |               |   |
| ا                            | '             | Pronounced as the English letter "a" in alight  |
| ب                            | b             | Pronounced as the English letter "b" in bite    |
| ت                            | t             | Pronounced as the English letter "t" in tight   |
| ث                            | th            | Pronounced as the English letter "th" in theft  |
| ج                            | j             | Pronounced as the English letter "j" in just    |
| ح                            | h             | No English equivalent                           |
| خ                            | kh            | No English equivalent                           |
| د                            | d             | Pronounced as the English letter "d" in doubt   |
| ذ                            | dh            | Pronounced as the English letters "dh" in that  |
| ر                            | r             | Pronounced as the English letter "r" in right   |
| ز                            | z             | Pronounced as the English letter "z" in zest    |
| س                            | s             | Pronounced as the English letter "s" in sight   |
| ش                            | sh            | Pronounced as the English letters "sh" in shout |
| ص                            | ʃ             | No English equivalent                           |
| ض                            | ɟ             | No English equivalent                           |
| ط                            | t             | No English equivalent                           |
| ظ                            | ʒ             | No English equivalent                           |
| ع                            | '             | No English equivalent                           |
| غ                            | gh            | No English equivalent                           |
| ف                            | f             | Pronounced as the English letter "f" in fight   |
| ق                            | q             | No English equivalent                           |
| ك                            | k             | Pronounced as the English letter "k" in kick    |
| ل                            | l             | Pronounced as the English letter "l" in light   |
| م                            | m             | Pronounced as the English letter "m" in might   |
| ن                            | n             | Pronounced as the English letter "n" in night   |
| هـ                           | h             | Pronounced as the English letter "h" in height  |
| و                            | w             | Pronounced as the English letter "w" in white   |
| ي                            | y             | Pronounced as the English letter "y" in yet     |
| <b>Short and Long Vowels</b> |               |   |
| ـ                            | a             | Pronounced as the English letter "u" in but     |
| ـ                            | i             | Pronounced as the English letter "i" in bit     |
| ـ                            | u             | Pronounced as the English letters "u" in put    |
| ا                            | ā             | Pronounced as the English letter "a" in bat     |
| ي                            | ī             | Pronounced as the English letters "ea" in beat  |
| و                            | ū             | Pronounced as the English letters "oo" in boot  |

Chart 1: The English transliteration of the Arabic writing system.



| No. | Chapters (Sūras)         |                       | No. | Chapters (Sūras)            |                       |
|-----|--------------------------|-----------------------|-----|-----------------------------|-----------------------|
| 1   | <i>Sūrat al Fāṭīḥah</i>  | The Opening           | 58  | <i>Sūrat al Mujādalah</i>   | The Disputation       |
| 2   | <i>Sūrat al Baqarah</i>  | The Cow               | 59  | <i>Sūrat al Ḥashr</i>       | The Mustering         |
| 3   | <i>Sūrat Āl 'Imrān</i>   | The Family of 'Imrān  | 60  | <i>Sūrat al Mumtahanah</i>  | The Tested One        |
| 4   | <i>Sūrat al Nisā'</i>    | The Women             | 61  | <i>Sūrat al Ṣaff</i>        | The Ranks             |
| 5   | <i>Sūrat al Mā'idah</i>  | The Table             | 62  | <i>Sūrat al Jum'u'ah</i>    | Friday                |
| 6   | <i>Sūrat al An'ām</i>    | The Cattle            | 63  | <i>Sūrat al Mundāfiqūn</i>  | The Hypocrites        |
| 7   | <i>Sūrat al A'rāf</i>    | The Ramparts          | 64  | <i>Sūrat al Taghābun</i>    | Mutual Disillusion    |
| 8   | <i>Sūrat al Anfāl</i>    | The Spoils            | 65  | <i>Sūrat al Ṭalāq</i>       | The Divorce           |
| 9   | <i>Sūrat al Tawbah</i>   | Repentance            | 66  | <i>Sūrat al Tahrim</i>      | The Prohibition       |
| 10  | <i>Sūrat Yūnus</i>       | Jonah                 | 67  | <i>Sūrat al Mulq</i>        | The Sovereignty       |
| 11  | <i>Sūrat Ḥud</i>         | Hud                   | 68  | <i>Sūrat al Qalam</i>       | The Pen               |
| 12  | <i>Sūrat Yūsuf</i>       | Joseph                | 69  | <i>Sūrat al Ḥāqqah</i>      | The Unquestionable    |
| 13  | <i>Sūrat al Ra'd</i>     | The Thunder           | 70  | <i>Sūrat al Ma'ārij</i>     | The Ascent            |
| 14  | <i>Sūrat Ibrāhīm</i>     | Abraham               | 71  | <i>Sūrat Nūh</i>            | Noah                  |
| 15  | <i>Sūrat al Ḥijr</i>     | Ḥijr                  | 72  | <i>Sūrat al Jinn</i>        | The Jinn              |
| 16  | <i>Sūrat al Naḥl</i>     | The Bee               | 73  | <i>Sūrat al Muzzammil</i>   | The Mantled One       |
| 17  | <i>Sūrat al Isrā'</i>    | The Night Journey     | 74  | <i>Sūrat al Muddaththir</i> | The Enshrouded One    |
| 18  | <i>Sūrat al Kaḥf</i>     | The Cave              | 75  | <i>Sūrat al Qiyyāmah</i>    | The Resurrection      |
| 19  | <i>Sūrat Maryam</i>      | Mary                  | 76  | <i>Sūrat al Insān</i>       | Mankind               |
| 20  | <i>Sūrat Ṭā Hā</i>       | Ṭā Hā                 | 77  | <i>Sūrat al Mursalāt</i>    | The Sent Angels       |
| 21  | <i>Sūrat al Anbiyā'</i>  | The Prophets          | 78  | <i>Sūrat al Naba'</i>       | The Tidings           |
| 22  | <i>Sūrat al Ḥajj</i>     | The Pilgrimage        | 79  | <i>Sūrat al Nāzi'āt</i>     | The Pluckers          |
| 23  | <i>Sūrat al Mu'minūn</i> | The Believers         | 80  | <i>Sūrat 'Abasa</i>         | He Scowled            |
| 24  | <i>Sūrat al Nūr</i>      | The Light             | 81  | <i>Sūrat al Takwīr</i>      | The Coiling Up        |
| 25  | <i>Sūrat al Furqān</i>   | The Distinguisher     | 82  | <i>Sūrat al Inṣṭār</i>      | The Cleavage          |
| 26  | <i>Sūrat al Shu'arā'</i> | The Poets             | 83  | <i>Sūrat al Mujaḥḥifīn</i>  | The Skimpers          |
| 27  | <i>Sūrat al Naml</i>     | The Ant               | 84  | <i>Sūrat al Inshiqāq</i>    | The Splitting Asunder |
| 28  | <i>Sūrat al Qaṣaṣ</i>    | The Story             | 85  | <i>Sūrat al Burūj</i>       | The Star Clusters     |
| 29  | <i>Sūrat al 'Ankabūt</i> | The Spider            | 86  | <i>Sūrat al Tārīq</i>       | The Night Star        |
| 30  | <i>Sūrat al Rūm</i>      | The Romans            | 87  | <i>Sūrat al A'lā</i>        | The Sublime           |
| 31  | <i>Sūrat Luqmān</i>      | Luqmān                | 88  | <i>Sūrat al Ghāshiyah</i>   | The Enveloping        |
| 32  | <i>Sūrat al Sajdah</i>   | The Prostration       | 89  | <i>Sūrat al Fajr</i>        | The Daybreak          |
| 33  | <i>Sūrat al Aḥzāb</i>    | The Confederate Clans | 90  | <i>Sūrat al Balad</i>       | The City              |
| 34  | <i>Sūrat Sabā'</i>       | Sheba                 | 91  | <i>Sūrat al Shams</i>       | The Sun               |
| 35  | <i>Sūrat Fāṭir</i>       | The Originator        | 92  | <i>Sūrat al Layl</i>        | The Night             |
| 36  | <i>Sūrat Yāsin</i>       | Ya Sīn                | 93  | <i>Sūrat al Duḥā</i>        | The Forenoon          |
| 37  | <i>Sūrat al Ṣāfāt</i>    | The Rangers           | 94  | <i>Sūrat al Sharḥ</i>       | The Dilation          |
| 38  | <i>Sūrat al Ṣād</i>      | Ṣād                   | 95  | <i>Sūrat al Tin</i>         | The Fig               |
| 39  | <i>Sūrat al Zumar</i>    | The Hordes            | 96  | <i>Sūrat al 'Alaq</i>       | The Blood Clot        |
| 40  | <i>Sūrat Ghāfir</i>      | The Forgiver          | 97  | <i>Sūrat al Qadr</i>        | The Decree            |
| 41  | <i>Sūrat Fuṣṣilat</i>    | Expounded             | 98  | <i>Sūrat al Bayyinah</i>    | The Elucidation       |
| 42  | <i>Sūrat al Shūrā</i>    | The Counsel           | 99  | <i>Sūrat al Zalzalah</i>    | The Earthquake        |
| 43  | <i>Sūrat al Zukhruf</i>  | The Ornamentation     | 100 | <i>Sūrat al 'Adiyāt</i>     | The War Steeds        |
| 44  | <i>Sūrat al Dukhān</i>   | The Smoke             | 101 | <i>Sūrat al Qārī'ah</i>     | The Rapping           |
| 45  | <i>Sūrat al Jāthiyah</i> | The Crouching Down    | 102 | <i>Sūrat al Takāthur</i>    | The Multiplication    |
| 46  | <i>Sūrat al Aḥqāf</i>    | The Sand Dunes        | 103 | <i>Sūrat al 'Aṣr</i>        | The Time              |
| 47  | <i>Sūrat Muḥammad</i>    | Muhammad              | 104 | <i>Sūrat al Humazah</i>     | The Traducer          |
| 48  | <i>Sūrat al Fath</i>     | The Victory           | 105 | <i>Sūrat al Fil</i>         | The Elephant          |
| 49  | <i>Sūrat al Ḥujurāt</i>  | The Chambers          | 106 | <i>Sūrat Quraysh</i>        | Quraysh               |
| 50  | <i>Sūrat Qāf</i>         | Qāf                   | 107 | <i>Sūrat al Mā'ūn</i>       | Benevolence           |
| 51  | <i>Sūrat al Dhāriyāt</i> | The Scattering        | 108 | <i>Sūrat al Kawthar</i>     | The Copious Provision |
| 52  | <i>Sūrat al Tūr</i>      | Mount Tūr             | 109 | <i>Sūrat al Kaḥfirān</i>    | The Unbelievers       |
| 53  | <i>Sūrat al Najm</i>     | The Star              | 110 | <i>Sūrat al Naṣr</i>        | The Succor            |
| 54  | <i>Sūrat al Qamar</i>    | The Moon              | 111 | <i>Sūrat al Masad</i>       | The Plaited Rope      |
| 55  | <i>Sūrat al Raḥmān</i>   | The Most Benignant    | 112 | <i>Sūrat al Iklās</i>       | The Pure Religion     |
| 56  | <i>Sūrat al Wāqī'ah</i>  | The Occurrence        | 113 | <i>Sūrat al Falaq</i>       | The Creation          |
| 57  | <i>Sūrat al Ḥadīd</i>    | Iron                  | 114 | <i>Sūrat al Nās</i>         | Mankind               |

Chart 2: The names and numbers of the chapters of the Qur'an

## Preface

The International Institute of Islamic Thought is pleased to present this volume on the subject of dealing at an intellectual level with the text of the Qur'an. This effort takes place within the framework of the Institute's efforts to develop new methodologies for dealing with and understanding the Qur'an. Indeed, the Institute has been quite active in encouraging the efforts of scholars in this important field and considers the subject a priority in the overall process of reforming Muslim thought.

The Institute's efforts in the field began with the commissioning and publication of Shaykh Muḥammad al Ghazālī's *Kayfa Nata'āmal ma'a al Qur'ān* (How We Deal with the Qur'an), which was followed by the Institute's publication of Fāṭimah Ismā'īl's doctoral theses *al Qur'ān wa al Naẓar al 'Aqlī* (The Qur'an and Rational Inquiry). These were followed by the publication of Dr. Muḥammad Jābir Fayyād's *al Amthāl fi al Qur'ān al Karīm* (Metaphor in the Holy Quran), and Muḥyi al Dīn 'Aḥiyah's *al Kashshāf al Iqtisādī li Āyāt al Qur'ān al Karīm* (An Index to Verses Concerning Economics in the Qur'an). The most recent publication in the series was Dr. Tijānī 'Abd al Qādir's *Uṣūl al Fikr al Siyāsī fi al Qur'ān al Makkī* (The Fundamentals of Political Thought in the Makkan Qur'an). All of these works have appeared in the Arabic language. The Institute has also published, in English and French, a work by Dr. Ṭāhā al 'Alwānī and Dr. 'Imād Khālid entitled *The Qur'an and the Sunnah*. In English, too, the Institute has published the proceedings of a conference entitled *The Qur'anic Concept of the Human Psyche*, edited by Ṣafar Ishāq Anṣārī.

Nor have the efforts of the Institute stopped short at publications. In fact, the Institute has supported the efforts of several research teams and seminars on the subject. Most recently, the Institute co-sponsored a seminar on the subject with the Sudanese University of the Holy Qur'an, which was held in Sudan in January 1996. Despite the increasing attention given worldwide to Islamic studies in general, and Qur'anic studies in particular, technological advances in the organization of information about the Qur'an brought on by the widespread use of computers have highlighted the need for the development of systems for the retrieval of Qur'anic texts,

systems which go beyond the syntactical and lexical to the contextual and conceptual, thereby enabling researchers to deal with problems of terminology, non-standard recitation, and interpretation. Indeed, were it not for advances in technology and the present intellectual revival that has awakened Muslims to the possibilities of the Qur'an as a source of knowledge, methodology, thought, and culture, many of these questions would never have arisen in the first place.

The need for computer-based information systems about the Qur'an was born of the methodological requirements of contemporary Islamic thought in dealing with the texts of the Qur'an itself. The methodology required must be capable of integrating subjective interpretations of the Qur'an and its verses and chapters with critical and analytical interpretations of its themes, issues, concepts, and categorizations. Such a tool will enable Muslim scholars to move beyond the stage of partial acquisition from contexts by dealing exclusively with lexical meanings and ignoring the essential relationships that exist between different elements of the same subject.

Dr. Hānī M. 'Aṭīyah's choice of this subject for his doctoral dissertation came from his conviction that the Muslim mind desperately needs to involve the informational sciences in its studies of the classical Islamic disciplines and, at the head of these, the science of the Qur'an itself.

We ask Allah Almighty to make this effort a beneficial one and a foundation upon which a truly Islamic information science may be erected.

The International Institute of Islamic Thought

## **CHAPTER ONE**

# **Introduction**

### **1.1. Information Science: An Overview**

All societies are information societies to the extent that one cannot imagine a community that has no pattern of information flow. Much of the information flow is personal communication and is therefore not confined intrinsically to speech or writing. For example, it could consist of music, pictures, sketches, colors, traffic signs and signals, highway and smoke signals, Morse code, sign language, club badges, uniforms, bodily posture, nods of the head, facial expressions, eye movements, and others. Moreover, in relation to cultural communication, the information flow deals with institutions and governments.

For any system or organization to survive, an effective flow of information is essential. This has become even more important in contemporary societies. The governments or quasi-official institutions of some societies have sought to regulate and control the information flow in pursuit of their own objectives and aims. Some provide information to specific groups, while others seek to provide it to a wider spectrum of society.

In the handling of common daily services and operations, organizations and governments use information for different purposes, such as planning, management, and decision making. The advent of computer-based data processing has made an immense flow of information available to institutions, the utilization of which can help them realize their goals. Information, therefore, is an indispensable resource for all societies and organizations. The field that has emerged to study this flow of information and its related services is known as “information science.” However, there are conflicting views about this term’s exact definition and who is entitled to be called an “information scientist.” The discus-

sion of such views is beyond the scope of this study.<sup>1</sup> In the context of this study, the term "information science" is defined in accordance with the definition that has been adopted by the American Society for Information Science (n.d.):

Information Science as a discipline seeks to create and structure a body of scientific, technological, and systems knowledge related to the information transfer chain ... As a discipline, Information Science investigates the properties and behavior of information, the focus that governs the transfer process, and the technology required to process information for optimum accessibility and use. Its interests include information representations in both natural and artificial systems: the use of codes for efficient message transmission, storage and recall: and the study of information processing devices and techniques such as computers and their programming systems.

It is an interdisciplinary field derived from and related to mathematics, logic, linguistics, psychology, computer technology, operations research, librarianship, the graphic arts, communications, management, and similar fields.

Information Science has both a pure science component, which inquires into the subject without regard to application, and an applied science, which develops services and products.

## 1.2. Information Science: An Islamic Context

Dealing with the challenges of the information age requires an in-depth understanding of and an ability to work in integrated and cooperative modes. New fields have emerged, and still others are emerg-

---

<sup>1</sup> For detailed reports and analyses, see H. Wellisch, "From Information Science to Informatics: A Terminological Investigation," *Journal of Librarianship* 4, no. 3 (1972): 157-87.

ing. While the language barrier has resulted in a delay factor, this has not stopped the production of publications from continuing to increase at an exponential rate. The end result of this process has been the realization that no single body or organization can cater to all the various information needs of a nation: what is needed is a system of co-operation between organized national centers and international information services so that each element may complement the other. Such a system is an essential part of any information policy.

This all seems to be rather elementary and self-evident. However, unfortunately, it must be said that such a system does not yet exist, for many services dealing with Islamic materials have developed in isolation from each other. In addition, the research output is often repetitive, of poor quality, and lacking in imagination. While Muslim librarians have incorporated such new disciplines as Islamic economics and Islamic anthropology, most of the Islamic legacy, which has been accumulating for fourteen centuries, has yet to be published. While there has been a recent exponential increase in the number of publications involving writings on Islam, books and articles on Islam and the Muslim world continue to be written in various languages and range, in format, from manuscripts to microfiche. Unfortunately, the fact remains that Muslim libraries and information services suffer from a lack of sufficient material resources that would allow researchers to take stock of this information explosion and of an intellectual strategy to overcome the related bibliographic deficiency. In addition, the resources of the Islamic legacy and literature that originated in Muslim lands are hopelessly dispersed, and Muslim librarians continue to waste their time and energy trying to adapt to the situation and modifying the information flow in such a manner that it will fit into the western system.

It is worth noting that library science, as developed in the West, is bound to reflect the image of western civilization and its ethos. The classification schemes, rules of cataloguing, lists of subject headings, and other library science techniques used to exploit the available material, all reflect the western way of life. A primary example of this is the restriction of "Islam" to the heading "Religion," as is the case with the

Dewey Decimal Classification scheme. Such a listing, based on a western conceptualization, violates the basic concept of Islamic principles. By listing such disciplines as "Islamic Economics" and "Islamic Anthropology" under the main headings of "Economics" and "Anthropology" in a Library of Congress Subject Headings List, a whole array of new disciplines has emerged that is not looked at in these lists. Furthermore, if the names of Muslim women are entered by their surnames, following Anglo-American cataloguing rules, Muslim researchers will be unable to find them.

Certain Muslim information scientists, as well as many Muslim thinkers, believe that as Islam has its own worldview, Islamic literature naturally has its own unique characteristics. However when innovative classification systems, bibliographical indexes, lists of subject headings, and other adaptations of these tools are adopted in ways that do not reflect properly the universality of Islam, any analysis and synthesis of Islamic material becomes impracticable. In the case of Muslim information scientists who are seeking to establish their own system, the paramount intellectual problem and the most devastating argument against the indiscriminate amassing of information stems from the attendant theory of knowledge. Knowledge, with a capital K, covers far more than a set of statements, for it is, in fact, a theory leading to an authentic classification of Knowledge. In other words, the classification of Knowledge is so fundamental to Islamic epistemology that it constitutes the first point in the dialectical relationship between facts and value.

Muslim scholars of the classical period were very well aware of the fact that information divorced from its value context is meaningless. This was so obvious to early Muslim thinkers that they propounded their own classification schemes before setting up their projects. The schemes of al Kindī (d. 252 AH / 866 CE), al Fārābī (d. 339 AH / 851 CE), al Khawārizmī (d. 380 AH / 990 CE), Ibn Sīnā (d. 428 AH / 1037 CE), al Ghazālī (d. 505 AH / 1097 CE), and Ṭāsh Kubrī Zādah (d. 968 AH / 1561 CE) must be regarded as the most notable. For example, al Fārābī and Ibn Sīnā divided the sciences according to whether they were theoretical (*nazarīyah* نظرية) or practical (*'amalīyah* عملية). Others, such as

al Khawārizmī, divided them into Arabic sciences (*'ulūm 'Arabīyah* علوم عربية) and foreign sciences (*'ulūm a'jamīyah* علوم أعجمية), while al Ghazālī adopted the bifurcation of all knowledge into revelational sciences (*'ulūm naqlīyah* علوم نقلية) and rational sciences (*'ulūm 'aqlīyah* علوم عقلية).

However, Muslims had undertaken specific subject-based classification much earlier than this general classification of Knowledge. For instance, the classification of Hadith in the first, second, and third *hijrī* centuries<sup>2</sup> by Mālik (d. 179 AH / 796 CE), al Ṭayālīsī (d. 204 AH / 820 CE), Ibn Ḥanbal (d. 241 AH / 855 CE), al Bukhārī (d. 256 AH / 870 CE), Muslim (d. 261 AH / 875 CE), Abū Dāwūd (d. 275 AH / 889 CE), al Tirmidhī (d. 297 AH / 910 CE), Ibn Mājah (d. 275 AH / 889 CE), and al Nasā'ī (d. 303 AH / 916 CE). Specifically, Mālik, al Bukhārī, Muslim, Abū Dāwūd, Ibn Mājah, al Tirmidhī, and al Nasā'ī arranged their collections according to juridical subject classification, while al Ṭayālīsī and Ibn Ḥanbal classified their collections according to narrator.

With the invention of printing, the philosophic classification of Knowledge became confused with the classification of books. In the words of Hulme (1911,446-47), this is because:

Books ... are concrete aggregates of facts selected from the common stock of knowledge, and are produced under the laws of supply and demand to meet the wants of the various bodies of community. The result is a welter of cross classifications and of overlapping areas of definition for the reception of which the framework of philosophic classification is quite insufficient.<sup>3</sup>

Hulme's alternative idea of classification according to a scheme of classes and subdivisions is justified and involves the establishment of

---

<sup>2</sup> The term *hijrah* (*hijrī* when used as an adjective) refers to the migration of the Prophet Muḥammad from Makkah to Madinah in 622 CE. This incident marks the beginning of the Muslim era.

<sup>3</sup> E. W. Hulme, "Principles of Book Classification," *Library Association Record*, no. 13 (1911): 446-47.



real classes of literature on particular subjects. In other words, the classification scheme is neither philosophically nor theoretically based, but empirically derived for the subjects of the publications. By any standards, bibliographical guides ought to belong to the least ideology-infected genre of academic writing. This, however, is not the case, for bibliographies present serious problems of structural, as opposed to personal, bias. For example, the first important Islamic bibliography, *al Fihrist*, was divided into ten areas of bibliographical Islamic literature: the Qur'an, grammar, history and belles-lettres, poetry, scholastic philosophy, law, philosophy and science, legends and fables, sects and creeds, and alchemy. The author, a bookseller known as Ibn al Nadīm (d. 385 AH / 995 CE), arranged his work in chapters divided on the basis of a particular subject, its development, and its authors (whom he had either met or about whom he had heard).

The next important work after Ibn Nadīm's *al Fihrist* belongs to al Ṭūsī (d. 460 AH / 1067 CE). This work, entitled *Fihrist Kutub al Shī'ah*, and known popularly as *Fihrist al Ṭūsī*, is based, for the most part, on the material provided by Ibn al Nadīm, but it completes it and adds more precise details on Shī'ah works and writers. The author, born five years after the death of Ibn al Nadīm, was a noted Muslim jurist and theologian of his period. This was one reason why he classified his *Fihrist* by the authors' names. The works of nine hundred authors, both contemporaries and predecessors of al Ṭūsī, were arranged by their authors' forename and the subjects covered included, among others, the fields of tradition, *rijāl* literature, interpretation of the Qur'an, scholastic theology, and jurisprudence.

Several centuries later, a monumental work of Islamic bibliography was compiled by Ḥājji Khalīfah (d. 1067 AH / 1657 CE), who was also popularly known as Kateb Çelebi, the Ottoman polilister. Entitled *Kashf al Zūnūn 'an Asāmī al Kutub wa al Funūn*, the author spent about twenty years collecting the material and then compiling the resulting information according to an alphabetical listing. This work included twenty-five thousand Arabic, Persian, and Turkish works belonging to nine thousand five hundred authors which he had either seen in the pub-

lic libraries of Istanbul or in the bookshops of Ḥalab as well as those he found mentioned in other books of *ṭabaqāt*. In its arrangement, he used the title of the work as the main entry, rather than the author's forename, which was a new approach. The concept of the title as the main entry later became a well-established practice in compilation and the preparation of library catalogues and bibliographies in the Muslim world.

Ibn Kahyr al Ishbīlī (d. 1179 AH / 1765 CE), a philologist and traditionalist of Seville, compiled the celebrated *Fihrist al Kutub wa al Ta'rif*. This bibliography listed one thousand four hundred Moorish and oriental works in Arabic on Qur'anic sciences, grammar, lexicography, literature, poetry, and a list of catalogues that had preceded his own. The books listed had either been read by the author or reported to him by trusted persons.

Ḥājji Khalīfah's attempt was continued in the early twentieth century in such monumental works as Brockelmann's *Geschichte der Arabischen Literatur*, Sezgin's *Geschichte des Arabischen Schrifttums*, and Storey's *Persian Literature*. Brockelmann's work was published originally during 1898-1902 and then expanded by three supplementary volumes during 1937-1942. The first and second volumes consist of a list of surviving Arabic manuscripts from printed catalogues of collections, while the third treats modern Arabic literature since Napoleon's invasion of Egypt. Sezgin's work, which appeared in 1967, deals with Arabic manuscripts and consists of several volumes covering Qur'anic sciences, theology and history, some of the natural sciences, and poetry of the period before 430 AH. Unlike Brockelmann, who organized his material on a chronological basis, Storey arranged his according to subject. In addition, the latter work covers material in Persian and contains notes on the principal editions, translations, and criticisms of the individual writers.

Listings of Islamic literature published after the invention of the printing press appeared first in Bianchi's (1821) *Catalogue des Livres Turcs, Arabes et Persans, Impriés à Constantinople, Deputis l'Introduction de l'Imprimère, en 1726-27, Jusqu'en 1820*, which included a chronological listing of sixty-eight Arabic, Persian, and Turkish books

printed in Istanbul during 1726-1820. This bibliography was completed by Hammer-Purgstall (1827-35) in his *Liste der Osmanischen Literature des 19ten Jahrhunderts*, which gave a chronological listing of ninety-eight works related to Ottoman literature printed in Istanbul during the period 1728-1830. Zenker's *Bibliotheca Orientalis* (1840) was a similar work that purported to provide the titles of all Arabic, Persian, and Turkish works that appeared after the invention of the printing press. Its second edition, which was published in two volumes, appeared in 1846-61 and contained classified listings with occasional annotations of 8,831 oriental language books and of translations from or into western European languages from the beginning of printing until 1957. Another book, which bore the same title as Zenker's bibliography and consisted of eight parts, was published by Friederici (1877-84), contains a complete list of books, papers, serials, and essays published during 1875-82 in England and its colonies, Germany, and France on the history, languages, religions, antiquities, literature, and geography of the East.

The first coverage of Islamic studies published in the West was presented by Schnurrer (1799) in *Bibliotheca Arabica: Actum Nunc Atque Integram*. Its second edition (1811) provided a subject classification along with detailed annotations and a chronological index of works published in Europe during 1505-1810. In fulfilling his ambition to continue Schnurrer's work, Chauvin (1892) compiled his twelve-volume *Bibliographie des Ouvrages Arabes au Relatifs aux Arabes Publiés dans l'Europe Chrétienne de 1810 à 1885*, which was published during 1892-1922. Klatt's four-volume *Literatur-Blatt für Orientalische Philologie* (1883-88) was another attempt to cover the literature on linguistics related to the Muslim world and was published in the West during 1882-86. This bibliography was continued in Müller's (1887-1911) *Orientalische Bibliographie*, which provided an annual classified listing of literature about Africa and the East between 1887-1911. Upon its termination, an attempt to fill the resulting gap of cover literature published between 1910-12 was made by Becker and Graefe. This triannual author listing of books and articles on Islam and the Muslim world was published under the section *Bibliographie* in the first three volumes of *Der Islam*.

Other distinguished works were Gebrael's *Manuale di Bibliografia Musulmana* (1916) and Pfanmüller's *Handbuch der Islam-Literatur* (1923). The latter, which provided a very detailed bibliographical essay of books and articles in oriental and western languages published up until 1923, had a great influence on the appearance of *Abstract Islamica*, which was established in 1927. The latter has appeared as a supplement to the *Revue des Études Islamiques* and provides abstracted information for European-language publications on Islam.

In 1954, the *International Bibliography for the History of Religions* was inaugurated with the assistance of UNESCO. Its mission was to provide an annual listing of books and articles on Islamic studies that had been published in western languages. This effort, however, was discontinued in 1973. Since that time, the *British Society for Middle Eastern Studies Bulletin* has published bibliographical reports and essays on a regular basis.

Pearson's *Index Islamicus*, a cumulative index first published in 1958, represents an attempt to list the periodical articles published during 1906-1955. In 1977, the *Quarterly Index Islamicus* was introduced as a recurrent series covering the growing volume of Islamic literature. Initially published by Mansell, an English publishing firm, it is now published in Amman, Jordan, by the Āl al Bayt Foundation, a subdivision of the Royal Academy for Islamic Civilization and Research. A distinguished volume by Behn (1989) sought to cover literature published during 1665-1957. An earlier work of his, the *Islamic Book Review* (1980), sought to cover all Islamic literature reviewed in the West.

Since 1970, Muslims have taken the lead in this effort to control their bibliographical destiny. In 1977, the Islamic Foundation (Leicester, UK) started publishing *The Index of Islamic Literature* as a supplement to its *Muslim World Book Review*. The *Index* provides material in English concerning works on Islam and the Muslim world that have been published in the West. In 1990, the *Index* became a joint publication of the Islamic Foundation and the International Institute of Islamic Thought (Virginia, USA). The great index *al Fihrist*, published in Beirut, started

issuing its bibliography in 1981 with the intent of covering Arabic periodicals published in the Arab world. Another work, *al Bibliografiyah al Mawdū'iyah al 'Arabīyah fī 'Ulūm al Dīn al Islāmī*, edited by 'Abd al Wahāb Abū al Nūr and published by the Arabic Organization of Education, Culture, and Science, attempts to cover as much literature on Islam as possible. Originally, the organization envisioned the publication of six volumes that would cover the fundamental aspects of Islamic sciences. However, as of this writing, only four volumes have appeared, namely, Qur'anic Sciences, Hadith Sciences, Fiqh Sciences, and Uṣūl al Fiqh Sciences. These volumes cover articles published in Arabic periodicals since the invention of the printing press until 1976. A more complete version is to be found in 'Abd al Raḥmān's (1984) four-volume *Index Arabicus*, which is published in Baghdad and covers literature published in Arabic periodicals and related to Islamic studies during 1876-1984. In 1989, a new quarterly bibliographical journal published in Nicosia, Cyprus, *al Kashāf al Islāmī*, began to cover publications in Arabic periodicals on Islamic subjects.

The production of similar, though somewhat specialized, bibliographies has been undertaken by Ali (1961), Usmani (1984), and Anees (1985) on the Qur'an; Hammadah (1965) on Sirah; Denffer on Hadith; Anees and Athar (1986) on Sirah and Hadith; Sardar (1977) on Hajj; Ahsan (1982) and Asaf (1983) on Da'wah and Islamic movements; Khan (1973) and Siddiqi (1981) on Economics; 'Aṭīyah (1985) on Education; Nasr (1975-78) on Islamic Sciences, Hamraneh (1964) on Medicine and Pharmacy; Anees (1988) on Muslim Woman and Family; Manzoor (1989) and Shahjahan (1988, 1989) on Philosophy; and Sardar (1987) on Librarianship.

From the foregoing state-of-the-art review of information management for Islamic literature, it is heartening to note that, in recent years, there has emerged an increasing awareness among Muslim individuals and organizations of the need to initiate some sort of bibliographical control. However the scope and method of the resulting efforts have been very limited. The material is varied and vast, new disciplines and languages have emerged, and the publication of books and articles on

Islam has become so extensive that it is now beyond the control of individual scholars in a given discipline. Global control of Islamic material, therefore, cannot be conceived of without the help of an integrated circuit of services or multinational organizations. Two decades ago, Pearson (1975) called for a "total bibliographic control of Islamic studies." However, his argument rests on western assumptions in dealing with Islamic bibliographies. A few years later, a similar concern was expressed by Anwar (1983), who suggested the establishment of a "World Islamic Bibliography Centre." Unfortunately, he did not recommend a conceptual design for the realization of such a universal bibliographical control of Islamic literature.

In the past few years, Muslim scholars have paid some attention to evolving new classification schemes. Earlier, Abū al Nūr (1973) developed an original scheme based on a literary warrant of publications in Arabic. This was followed by Sardar (1979), who worked out a scheme based on a general idea taken from Ranganathan's Colon Classification. But a review of these two schemes shows that neither can be used by all Islamic countries: Abū al Nūr initiated his scheme on the basis of the Arabic alphabet and Indian numbers, while Sardar based his on the English alphabet. Thus neither is suitable for use in the majority of Muslim countries. Sabzwari (1982, 16), who listed at least nine other proposed or currently used classification schemes in different Muslim countries and universities, concluded that "none of these could be adopted in toto by all Islamic countries." Instead, he suggested a Universal Islamic Classification, which has not yet been subjected to a full discussion or analysis.

Therefore it is the belief of the author of this book that the basic issue in initiating a method of global control for Islamic materials, both conceptually and physically, is the need to establish a full subject index of the Hadith collections and the Qur'anic text. Unless such indexes are provided, the subject-content of the rich yet diffuse literature of the Islamic legacy will remain confused in terms of classification criteria. This argument is based on two points. First, those writing in the field of Islamic studies commonly refer to Qur'anic verses and Hadith traditions in order to justify their claims or to prove their points. This requires two

things: an author must either refer to the Qur'anic and Hadith exegeses or depend on his own understanding and interpretation. If he refers to the Qur'anic and Hadith exegeses, it is probable that he would find it difficult to locate the desired information. As traditional exegeses follow the same arrangement of both the Qur'anic text and the Hadith collections, a researcher who did not have a proper subject index to guide him through the scattered related material would have no chance of finding all of the needed material. However, if an author depends on his own understanding and interpretation, he may need to review all of the material dealing with or relating to his subject. This requires instantaneous access to secondary sources, viz, books, articles, dissertations, and reports, which the user can find in bibliographies, periodicals, dissertation indexes, and conference proceedings. Without these materials being classified in a systematic way, which would allow the author to find what he needs without inconvenience, the user will again face difficulties. In contrast, seeking access to the primary sources (i.e., the Qur'anic texts and the Hadith collections themselves), still requires an enhanced subject index in order to gather the scattered related subjects. The second argument is that the actual works of these authors, from the point of view of librarians, refer to Qur'anic and Hadith subjects and, in order to be placed on the shelves in the proper and most suitable location, they require an analytical subject classification scheme. Again without a proper classification scheme, researchers will not be able to locate the desired information easily.

Therefore, as a first step, the proposed system requires the establishment of Qur'an and Hadith retrieval systems: this means analyzing contemporary Arabic-Islamic writings together with the Islamic legacy through the combined use of two retrieval systems. Through this method, the generation of a multiuse classification scheme of literature dealing with the Qur'an and Hadith topics will become a preliminary step in the task of establishing a bibliographical framework for the global control of Arabic-Islamic literature dealing with the Qur'an and Hadith literature. Both the classification scheme and the bibliographical base are classified according to microclassification classes of resources

of the Islamic legacy and civilization in non-Arabic languages, as well as those that originated in Muslim lands, can be combined in an integration system that will form the base of a universal Islamic classification scheme and a universal Islamic bibliography. The conception of a microclassification of classical and contemporary Arabic writings will result in a framework for an integration system that, in turn, will form the base of a national Arabic classification scheme and a national Arabic bibliography. Figure 1-1 outlines the steps in this integration system and the relation between its parts.

But we must still start with the Hadith collections and the Qur'anic text. The Hadith collections (i.e., those of Malik, al Bukhari, Muslim, Abu Dawud, Ibn Majah, al Timidhi, al Nasa'i, al Tayalisi, and Ibn Hanbal) have certain limitations: in many cases, one who is looking for a specific Hadith must have prior knowledge of the narrator or else he must guess which chapter contains the hadith. This has been appreciated by many scholars, notably Wensinck (1983, 69) and 'Abd al Baqi (1978), both of whom initiated an index for the above-mentioned Hadith collections. Wensinck compiled a comprehensive word index, while 'Abd al Baqi compiled a detailed subject index. An analysis of these works is beyond the scope of this study and must remain an area of future investigation.

As regards the Qur'anic verses, their arrangement does not follow any chronological order and the subjects are scattered throughout the texts. Several notable works have been compiled, both manually and automatically, in order to establish both a word and a subject index. However, these works were not undertaken for the purpose of an analytical subject classification, but for their own purposes. An analysis of these works will be made in the following section.

### **1.3. Early and Modern Qur'anic Indexes**

For Muslims, the Qur'an is the very word of God revealed to the Prophet Muhammad through the archangel Gabriel, as were the earlier revelations sent to the prophets David, Moses, and Jesus. The Qur'an plays an important role in the lives of all Muslims. Only the Arabic text



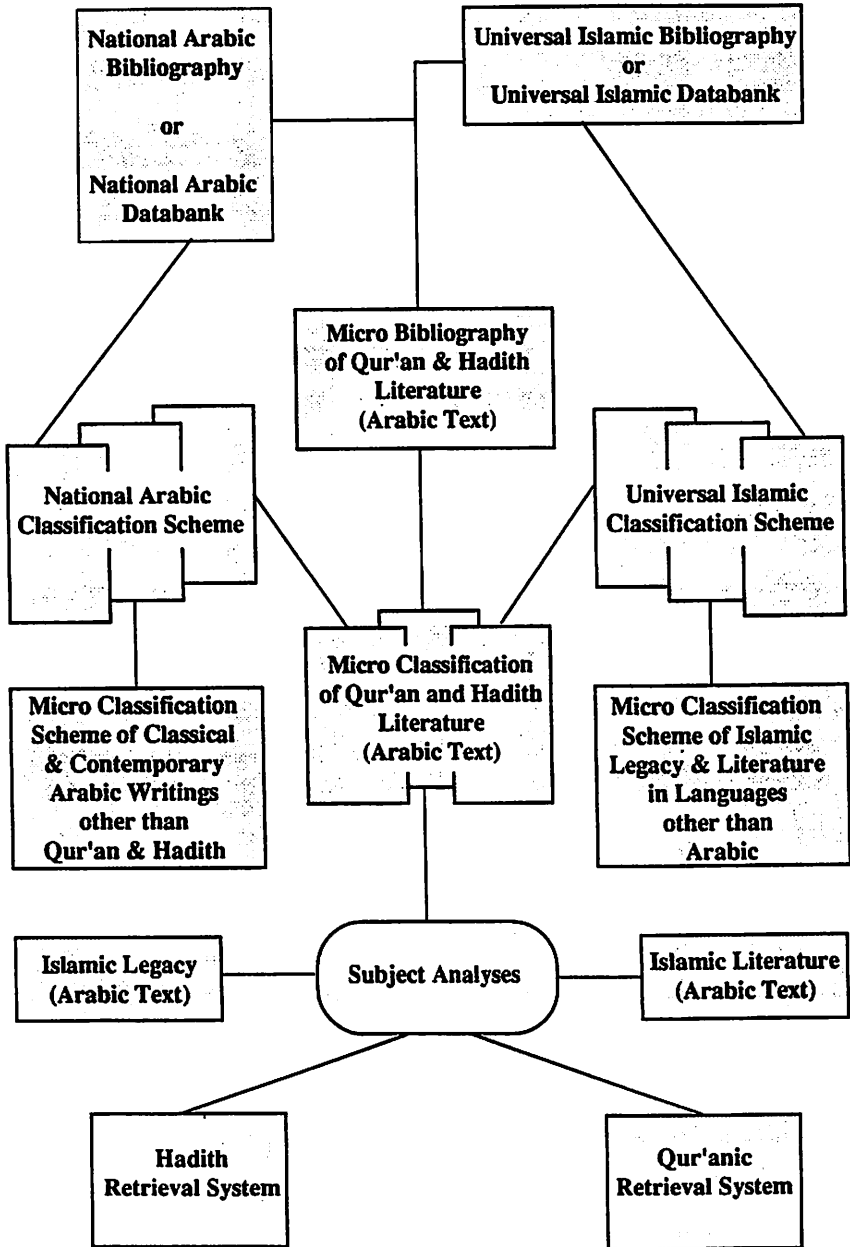


Figure 1.1: A chart representing the proposed integrated system for both conceptual and physical control of Islamic literature.

is considered authoritative and appropriate for recitation and revealing the word of God. Moreover the Qur'an itself, while reaffirming the truth of all previous revelations, comprises all truth for mankind. In furthering this aim, it seeks to create faith in God through the revelation of His signs and provides for the material, moral, and spiritual welfare of mankind. Through such a promotion of human welfare in every sphere, it expounds and explains all that is (or may be) needed by mankind for the complete fulfilment of human life. In the case of devotion, it explains the five pillars of Islam (submission to God, prayers, fasting, alms-giving, and pilgrimage). On sociological matters, it deals with such institutions as marriage, divorce, inheritance, and the care of orphans and widows. On the individual level, the Qur'an teaches good conduct, discusses Islamic ethics, and lays down rules for the conduct of interrelationships among the members of society. On the educational level, it informs people of what happened to their predecessors, some of whom were evil and some of whom were good, and provides examples of how to live properly by relating the lives of the prophets and the messengers. Finally on the spiritual level, the Qur'an is a constant reminder that life in this world is short and that each individual will have to account for his deeds in the hereafter.

In comparison to other religious scriptures, the Qur'an is a book of medium size. Divided into 114 chapters, arranged roughly according to length, the Qur'an is a mosaic of revealed verses uttered by the Prophet Muhammad at different places over a period of twenty-three years. The arrangement of the verses does not follow any chronological order, and the subjects, though scattered throughout the text, still show a unique kind of unity and homogeneity. In view of the absence of mass printing facilities at that time, the customary method for safeguarding the revelation and the tradition was for many people to commit it to memory. In that sense, the scattering of subjects over the Qur'anic text was not an obstacle to the verses' retrieval. The amount of information that can be passed on in this way, however, is limited. When it was discovered how to record information of various kinds in a relatively permanent form, a development that could be regarded as finding a substitute for human

memory, society began to advance. As a result, instead of having to rely on an individual's memory for the desired information, we now have the printed copy of the Qur'an. Instead of a natural memory, we now have an artificial memory: the index.

As a matter of fact, the term "index" originally comes from a Latin word that means "he who, or that which, points the way." In the sixteenth century, it became fully Anglicized and acquired its current literary sense (Knight 1980, 17). Literally, the term "index" means "that which shows, indicates, manifests, or discloses a token or indication" (Webster's 2d ed.). In more technical terms, an index is

a systematic guide to items contained in or concepts derived from a collection. These items or derived concepts are represented by entries arranged in a known or stated searchable order, such as alphabetical, chronological or numerical. (Rothman 1974, 286)

In the broader sense of the word, four eras have to be considered when dealing with Qur'anic indexes. The first is that of the polymath,<sup>4</sup> which represents the time when the sum total of human knowledge was sufficiently small enough for one person to comprehend it all. Throughout history, generations of Muslim scholars have devoted themselves to interpreting the Qur'anic text. However, for interpretation depends upon the meanings behind the actual words, this led to the use of linguistic and related disciplines. Undoubtedly, this was due to the tremendous need felt by Muslims to know Qur'anic Arabic, which became the official language of established Islamic states. Consequently, the study of Arabic grammar was intimately linked with Qur'anic analysis and especially with the science of *al qirā'āt* (القرآيات readings). By the end of the first century AH / eighth century CE, grammarians organized large-scale linguistic surveys to collect as much linguistic data as possible. They

---

<sup>4</sup> This term has been borrowed from Foskett *The Subject Approach to Information*, 4th ed. (London: Bingley, 1982), 3.

then refined and systematized the earlier generation's methods of analysis. From these surveys, they were able to delineate the exact linguistic structure and its effect on the communicators of the language in their efforts to capture its various representations, all of which had to be taken into account while formalizing the structure of the Arabic language.<sup>5</sup>

As a consequence of these formalizations, attempts to compile the contents of the Qur'an were subjected to further analysis. This, in turn, gave rise to the study of a variety of issues, among them readings and modes of recitations, taken up by Ibn al Jazrī (d. 833 AH / 1430 CE) and al Dānī (d. 444 AH / 1052 CE); synonyms and homographs by al Damaghānī (d. 478 AH / 1085 CE) and al Tha'ālibī (d. 429 AH / 1038 CE); abrogation and abrogated verses by Ibn al Bārīzī (d. 738 AH / 1337 CE) and Ibn Du'āmāh (d. 117 AH / 735 CE); ambiguities by al Asfahānī (d. 502 AH / 1109 CE) and Ibn 'Abd al Salām (d. 660 AH / 1262 CE); the circumstances surrounding the revelation of an individual verses(s) by al Suyūṭī (d. 911 AH / 1505 CE) and al Wāḥidī (d. 468 AH / 1076 CE); metaphors by Abū 'Ubaydah (d. 209 AH / 824 CE); similes by Ibn Nāqiyā (d. 485 AH / 1092 CE); repetitions by al Asādabādī (d. 415 AH / 1024 CE) and al Karamānī (d. c. 500 AH / 1107 CE); and norms of reading styles by Ibn al Anbārī (d. 328 AH / 940 CE).

This vast literature of specialized exegesis has served as a crude form of thematic accession to the Qur'anic sciences (*'ulūm al Qur'ān* علوم القرآن) as well as a systematic expression of the Qur'an's relevance to issues of everyday life. This could, of course, be of purely historical significance, but from the technical point of view, it may also be regarded as a first attempt at indexing.

The second era, according to Foskett (1982), was the problem-oriented era, which, in effect, started with the development of printing. The new technology associated with writing and illustrating instruments advanced considerably during this period, which lasted until well into this century. This era was characterized by the need to solve particular

---

<sup>5</sup> See, for example, 'Abd al Raḥmān Muḥammad Ibn Khaldūn, *Muqadimat Ibn Khaldūn*, 4th ed. (Beirut: Dār al Qalam, 1981), 437-40.

problems, using whatever disciplines might be necessary, regardless of whether they belonged together or not. This could probably be referred to as the beginning of word indexing.

The first attempt by a Muslim to compile a Qur'anic index was the work of al Wirdārī (d. 1054 AH / 1637 CE).<sup>6</sup> His work, known as *Tarṭīb Zibā*, indexed all of the key words in the Qur'an with reference to their verse and surah numbers. The significant words, as they appear in the Qur'an, were arranged alphabetically by their first letter.

As a distinguished work of its time, such scholars as al Nabulṣī (d. 1062 AH / 1645 CE),<sup>7</sup> al Naqshabandī (d. 1227 AH / 1810 CE),<sup>8</sup> and al Walī (d. c. 13th century AH / 19th century CE)<sup>9</sup> have attempted to recompile it in different styles. Works similar to al Wirdārī's index have been compiled by Kāzīm (1859) and Shāh (1906).

The first European attempt to index the Qur'an was undertaken by the German orientalist Gustav Flügel who, in 1812, published his *Concordantia Corani Arabicus*. This is one of the most concise and comprehensive Qur'anic concordances ever compiled by an European. Flügel arranged the words alphabetically according to their roots, a decision that resulted in one noticeable disadvantage: as he followed no systematic criterion and did not adopt any well-known arrangement, many words are rendered incorrectly in terms of their original roots. A further problem is that he used a self-designed verse numbering system, which he derived from a special copy of the Qur'an prepared by him specifically for his concordance. The end result was confusion, for a researcher who used the accepted version of the Qur'an was confronted with unfamiliar verse numbers. Works that are similar to Flügel's concordance are those of Qaṣṭamūnī (1905), al Maqdisī (1906), and Yakin

---

<sup>6</sup> al Ḥafīẓ Maḥmūd ibn al Mulā Darwīsh al Widārī, *Tarṭīb Zibā* (Saudi Arabia: Imām Muḥammad ibn Sa'ūd Islamic University, n.d.), ms. no. 3609.

<sup>7</sup> Ismā'īl ibn 'Abd al Ghānī ibn Ismā'īl al Nabulṣī, *Unwān al Āyāt* (Saudi Arabia: Imām Muḥammad ibn Sa'ūd Islamic University, n.d.), ms. no. 2989.

<sup>8</sup> Muḥammad Najīb ibn al Ḥājj 'Umar al Naqshabandī, *Tashīl al Tarṭīb* (Saudi Arabia: Imām Muḥammad ibn Sa'ūd Islamic University, n.d.), ms. no. 6017.

<sup>9</sup> al Ḥafīẓ Muṣṭafā ibn Sulaymān, *Tahzīb al Tarṭīb fī Fahrasāt Āyāt al Qur'ān* (Saudi Arabia: King Saud University, n.d.), ms. no. 2903.

(1984). The latter has tended to arrange his work by abjad, as opposed to alphabetic, order.<sup>10</sup>

'Abd al Bāqī (1945), who translated Flügel's concordance into Arabic, included notations concerning the errors found in the original work. He also revised the verse numbers and adopted the system used in the structured copy of the Qur'an printed by the King Fu'ād press. 'Abd al Bāqī's concordance became, and remains until our own day, one of the most common, widespread, and admired works in the Muslim world. However, it also contains some notable disadvantages related to the use of the root. First, an intellectual effort has to be made to reduce the words to their roots. Second, words may be reduced to an incorrect original root. Third, since the concordance is constructed on the basis of three radical roots, confusion may occur when dealing with words having two, four, and five radical roots. Fourth, the reduction of words to their root forms is extended to names, which are abstract and thus can be misplaced.

Other works that are similar to 'Abd al Bāqī's concordance have been produced by al Ibiyārī and Marzūq (1953-69) and Barakat (1957), both of which follow 'Abd al Bāqī's arrangement. Al Ibiyārī and Marzūq's work, however, is part of a five-volume encyclopedia of the Qur'an. In addition, al Miṣrī (1941), Majma' al Lughah al 'Arabīyah (1953-69), and Ibrāhīm (1961) have produced Qur'anic concordances in which the lexical meaning accompanies the words. The latter, having recognized 'Abd al Bāqī's weakness of reducing names to roots, provides them in alphabetical order and without derivation(s). One English-language work that is similar to 'Abd al Bāqī's is that of Kassis (1983), which has the distinctive characteristic of being divided into two parts. Part one consists of a list of brief Qur'anic citations in which the name of God occurs. It is arranged in transliterated Arabic alphabetic order according to the key Arabic term with which the holy name is associated. A similar list of citations covers the occurrences in the Qur'an of all other words

---

<sup>10</sup> The difference between abjad and alphabetic order is discussed in chapter 3, section 3.5A.

except common words (i.e., prepositions and articles). Part two consists of an index of all English words, except common words that occur in Arberry's translation of the Qur'an, arranged as in the first part.

Other attempts, which feature a back index of the copy of the Qur'an, are provided by al Ḥimṣī (1984) and Ramyār (1964). Less similar, yet following the same approach are the works by Nāzīm (1867), al Nikhābī (1920), al Dimashqī (1927). Naṣṣār (1965), al Shāfi'ī (1972), Hāshim (1979), al Bindāq (1981), and 'Aḫīyah (1984). These works attempted to list the Qur'anic verses alphabetically rather than simply compiling the words. However, the approaches are not as far removed from a word methodology as to deter one from calling them Qur'anic word indexes.

Turning to the multilingual Qur'anic dictionaries, both Penrice (1985) and Nadawi (1983) have supplied Arabic-English Qur'anic dictionaries. Penrice, basing his work on Flügel's, locates each word under its verbal root. Where none exists, the word is placed alphabetically. The vowel of the aorist is given where it is known. One major drawback, which Penrice points out in his preface, is his failure to notice the variant readings of disputed passages or the numerous interpretations of the same passages, which abound in the commentaries. The dictionary of al Nadawi, is no different, although it does cover this gap. In the preface, the author lists the titles of the commentaries that he used and also provides supplemented indexes for the roots of the words used. Similar works have been undertaken in Urdu (Maudūdī 1979), Indonesian (Syamsu 1977), and in a Turkish-English-Arabic-Persian glossary (Eckmann 1976).

The third era, referred to by Foskett (1982) as the discipline-oriented era, is characterized by the division of knowledge into more or less watertight compartments or disciplines, a process that has resulted in an increasing specialization and a development of disciplines in their own right. Based on this development, works on selected Qur'anic topics have been published in the English language by Kherie (1979), Afzalur Rahman (1981), and Sherif (1985). Attempts to list relevant verses under specific headings have been undertaken by Afzalur Rahman (1983) and

Khan (1987). However, the first comprehensive subject headings of the Qur'an was introduced by the French orientalist La Beaume, in his work *La Koran Analyse*, who divided his work into eighteen broad major subject headings and then subdivided them further into three hundred fifty items. The work was translated into Arabic by 'Abd al Bāqī under the name *Tafṣīl Āyāt al Qur'ān al Ḥakīm*. During the translation process, 'Abd al Bāqī (1969) added a further one hundred items from Montet's French translation of the Qur'an. Works similar to 'Abd al Bāqī's were produced by Muhannā (1983), Barakāt (1985), and Muḥammad (1984). Muhannā divided his work into eighty-seven subjects, Barakāt divided his into twenty-four broad subject headings, and Muḥammad into twenty-nine broad headings.

Other attempts to produce a back index of the copy of the Qur'an have been undertaken by al Ḥimṣī (1984), who divided his work into fifteen broad headings, and Ramyār (1964), who divided his into thirteen. In his index, Ramyār also listed all of the unknown persons mentioned in the Qur'an separately at the end.

Finally, Fānī and Khorramshāhī (1986) compiled an exhaustive and comprehensive Persian-Arabic subject index. This work, which is based on an alphabetic arrangement of subjects, is not divided into broad subjects but rather indicates the entries directly in alphabetical order. The index includes seven thousand entries, together with cross-references. Also, variant readings, the circumstances surrounding the individual revelations, and different interpretations are taken into consideration in the number indicated by the authors in the preface. Although this effort is without a doubt the most successful and advanced Qur'anic subject index, it has three main disadvantages. First, the entries are arranged alphabetically and not under broad headings. Thus the user has to guess the word before searching for it, a process that reduces his recall of subjects to his ability to recall particular words. Second, the reference to the verse and chapter numbers alone without the Qur'anic text requires the user to refer to a copy of the Qur'an before finding his material. Though this ultimately reduces the size of the index, it creates the irritation of doubling the effort. Third, the fact that the two authors are Shī'ites



means that their Shī'ite thought is reflected on several occasions in the index as, for example, referring to the claim that some verses were revealed in the cause of Fāṭimah and 'Alī (see the index under Fāṭimah and 'Alī).

Despite the fact that they could be regarded as unique attempts in the area of Qur'anic word and subject retrieval, all of them, with the exception of the work of Fānī and Khorramshāhī (1986), have some important deficiencies. First, the division of the subjects does not reflect the actual range of the Qur'an's contents, but is rather arbitrary in nature. This is reflected in the variety of divisions found in these works. For instance, eighteen broad headings are used in La Beaume, fifteen in Rāmyār (n.d.), and twenty-nine in Muḥammad (1984). Second, the selection of broad headings does not follow any consistent criterion, as can be seen by the appearance of some specific topics alongside their general headings. For instance, *al akhlāq al dhamīmah* الأخلاق الذميمة (immorality) and *al nifāq* النفاق (hypocrisy) by Barakāt. Third, these indexes depend on only one popular reading in the Islamic world: "the reading of Ḥafṣ," as is shown clearly by al Ḥimṣī. This eliminates other possible interpretations of the verses that could have resulted from different readings. Fourth, the surrounding circumstances of the individual revelations are not taken into consideration, even though they could highlight some historical interpretations of the verses. Fifth, the works do not say whether the placement of verses was decided upon by the compiler or whether it was the result of consulting specific commentaries. Sixth, cross-referencing is poor and the lack of control vocabulary (i.e. plural-singular, derived nouns, synonyms) is apparent in the selection of the key words. It is evident, however, that the authors of the above-mentioned works did not commit themselves to establishing an analytical subject classification. Rather, they were doing it for their own satisfaction alone.

Considerable effort has been made to accommodate contemporary technology to the Qur'an in order to produce a screen presentation of the Qur'anic text and to facilitate easy search and retrieval. Among such

undertakings are *Salsabeel*,<sup>11</sup> *al Qur'ān al Karīm* (Zerosoft),<sup>12</sup> *the Alim*,<sup>13</sup> *al Qur'ān al Karīm* (al 'Alamīyah),<sup>14</sup> and *al Qur'ān al Karīm* (Apple Mac)<sup>15</sup>. *Salsabeel* is essentially a software word index that provides whole word matching. However, it does not retain the Qur'anic calligraphy or display the essential diacritical marks. Similarly, *al Qur'ān al Karīm* (Zerosoft) provides the Qur'anic text in Arabic along with its English translation. Yet the search strategy is applied only to Arabic text and is either by whole word matching or through a menu selection. Opposite to this, *the Alim* software presents the Qur'anic text in Arabic, but the searching is applied only to its English database. This software provides whole word matching, string matching, and synonym searching via Boolean operators. The more advanced *al Qur'ān al Karīm* (al 'Alamīyah) maintains the diacritical marks and original calligraphy of the Qur'an. The search strategy is by whole words and string matching and roots, vowelled and unvowelled. It also provides for searching through subjects along with some other features that help the user in the field of Qur'anic studies. Finally, *al Qur'ān al Karīm* (Apple Mac) is quite similar to that of al 'Alamīyah, though it is directed to serve Mac users.

However, now that we are in the fourth era, the mission-oriented era (Foskett 1982) in which demands for information can span a whole range of disciplines, the old barriers between disciplines have been removed, even if shadow boundaries among specialized fields have been retained. This process (i.e. analyzing the Qur'an with the intent of establishing an analytical subject classification system), therefore, emphasizes that the distinction between word and subject indexing is worth defining. Specifically, in a word index, the indexer is tied inextricably to the text's

---

<sup>11</sup> Applied Microsystem Technology, *Salsabeel: An Electronic Index of Qur'an*, (London: 1987).

<sup>12</sup> Zerosoft, *al Qur'ān al Karīm: A Qur'anic Concordance* (1990).

<sup>13</sup> Advent Technologies, *The Alim: An Islamic Literature Database* (Texas: 1992).

<sup>14</sup> al 'Alamīyah, *al Qur'ān al Karīm: An Electronic Library of Qur'an* (Kuwait: 1988).

<sup>15</sup> Apple Mac, *al Qur'ān al Karīm with Tafīr al Jalāllin: An Electronic Library of Qur'an* (Cairo: 1993).

words, regardless of their meanings: if the word appears in the text, it is selected; if it does not appear, it is not selected. In that sense, word indexing requires little skill in the field of indexing and can be done by an indexer who has had no subject-matter training. Although word indexing leads to the dispersal of related subjects, since the words' meanings are not taken into account, it is this type of indexing that is performed effectively by the computer.

In contrast, an indexer who is engaged in a process of subject indexing appeals frequently to the words' meanings and then makes a careful selection of the terminology used in the subject enquiry in order to avoid a dispersion of related subjects under different headings. This type of indexing, therefore, requires a great deal of subject-matter training and necessitates a thorough competence in indexing as well. In terms of machine searching, it has had little success and still requires a great deal of work.

In the present study, the process of indexing concentrates not simply on selecting words or keywords in the text, but on the study, emphasis, organization, and expansion of the text's contents. It requires the indexer to abstract those subjects that are apparent in the text, as well as those that are hidden. He must also convert the words into single-word form or into combinations of words that best express the ideas. As such an index has not lent itself to machine searching in the past, a number of researchers are now actively examining its potential. In the context of this study, this is the form of index that will be referred to as a retrieval system.

#### **1.4. A Qur'anic Retrieval System**

A Qur'anic retrieval system, as explored in this study, is a system designed to augment human knowledge and to aid human activity by storing items from the Qur'anic text that need to be processed, searched, retrieved, or disseminated to various users. To compile a Qur'anic retrieval system, two main factors should be considered: the abstraction of hidden allusions to the subject from the Qur'anic text, and the ability to identify them by words even if the text mentions them in a different

phraseology. In the construction process, two complementary features are involved: various interpretations applied to a verse(s) by different commentators, and the form and organization of the headings that are likely to describe the verses in the retrieval system. From the indexing point of view, the difference in interpretations could lead to placing such verses under different headings. On the other hand, the retrieval system's efficiency would be affected seriously if the compiler did not use the headings in a strictly controlled way in order to provide the user with easy access to the system. In many cases where the criteria for verse placement and heading selection are not specified in the construction of the retrieval system, the user may feel that the information sought is not provided when it is, in reality, located somewhere else. This is probably the most important aspect of subject indication.

In terms of its general form (i.e., for the purpose of a library catalogue) Cutter (1949, 66-75), who followed an empirical approach to the basic problems of subject indication, provided four rules to be used for governing the formation of classes by literary warrant: polytopical books must all be entered under a distinct subject heading, each work must be entered under its subject-heading and not under the class heading that includes that subject, each work must be entered under the word that best defines its subject, whether it occurs in the title or not, and the preferred heading in the case of synonyms is the one that is most familiar to the class of people who consult the library. According to Lancaster (1972, 15) Cutter's work represents an attempt to devise a generalized set of codes for the construction of alphabetical subject headings that would serve as a starting point for the pattern of vocabulary control, namely, a controlled list of index terms, which was dominant in subject catalogues in libraries for many years.

As the subject matter of publications tends to be complex, Lancaster (ibid., 23) claims that the vocabulary of the compound subject headings suggested for library cataloguing is not adequate for the more exhaustive technique of subject indication, that is for a "thesaurus." Originally, the term "thesaurus" came from the Greek concept of storehouse or treasury (Vickery 1960, 181). Subsequently, it came to be used in connection

with the construction of an English dictionary, such as *Roget's Thesaurus of English Words and Phrases*. Roget's *Thesaurus*, first published in 1852, arranged the words according to the ideas expressed in order to provide the user with a number of alternative words for a similar concept that might provide an equally appropriate word for that concept. Arabic works that are of a similar nature to Roget's began to appear in the third century AH, viz., al Sākit's (d. 244 AH / 858 CE; 1877) *Tahdhīb al Alfāz*; Ibn Qutaybah's (d. 276 AH / 889 CE; n.d.) *Adab al Kātib*; al Hamdānī's (d. 320 AH / 932 CE; 1885) *al Alfāz al Kitābīyah*; Ibn Qudāmah's (d. 337 AH / 948 CE; 1932) *Jawāhir al Alfāz*; al Rumānī's (d. 384 AH / 994 CE; 1988) *al Alfāz al Mutarādifah al Mutaqāribat al Ma'nā*; and Aḥmad Ibn Fāris' (d. 395 AH / 1005 CE; 1970) *Mutakhayyir al Alfāz* were all compiled in an attempt to provide the writer with a list of synonyms from which to select.

Given this function, these types of works cannot be regarded as a thesaurus for information retrieval from the information scientist's point of view (Lancaster 1972, 25). The most distinctive function of the information retrieval thesaurus is to control synonyms, homographs, generic levels of meanings, and spurious relations between terms. According to Joyce and Needham (1958, 194), investigators at the Cambridge Language Research Unit in England began to discuss the applicability of the thesaurus concept to information retrieval in 1956. According to Vickery (1960, 181), the first time the word "thesaurus" was used in connection with information retrieval was in 1957. Nowadays, the thesaurus is used more widely than the list of subject headings. It has become so popular that Schreider (1965, 222) defines information as "the degree of change of the thesaurus."

In principle, subject heading lists and thesauri are similar in that both consist of alphabetically arranged terms with accompanying cross-references. Also, some specialists use the terms "subject headings list" and "thesaurus" interchangeably because of their evident similarity of design. Foskett (1982, 440-41) expresses this similarity thus:

The only difference between a thesaurus and a list of subject headings is that the former normally excludes headings for composite subjects; with some examples, even this is not true, and one is forced to the conclusion that the name is intended to signify a distinction which does not exist. Many lists of subject headings have left much to be desired, and thesaurus constructors have perhaps endeavored to persuade us (and themselves) that their lists were free from fault.

Yet the fact remains that traditional subject heading lists, such as that used by the Library of Congress, are meant to deal with human knowledge in its entirety, while most, if not all, modern thesauri treat only a subset of this knowledge, namely, a particular discipline or field of study. In the context of this study, subject heading lists and thesauri will be considered equal in terms of their similarity in design for the purpose of analyzing their weaknesses and strengths in relation to Qur'anic vocabularies.

The foregoing considerations make it plain that, in principle, a great deal can be gained from using vocabulary control devices. In practice, it may emerge that accuracy and consistency are difficult to maintain. Not only must indexers be intimately aware of the available indexing vocabularies and practices, but they should also be familiar with collection characteristics and possess a high degree of sophisticated training and experience. More often than not, however, resulting index entries are incomplete or lack specificity. They are also time-consuming and expensive. By far the most common alternative technique is to use little or no vocabulary control, namely, to use an automatic natural language database.

In an automatic natural language database, the entire text of the document collection has to be fed into a machine-readable form. Normally, this would be kept in a special storage memory under code numbers. A special program is then prepared to locate the items of information on request by means of a program-searching strategy. An automatic natural language database can serve as a successful replacement of a manual system due to the potential advantages of using the language of discourse. In addition, it pro-

vides just the right kind of expression to denote each particular concept and may be carried out more rapidly, and more cheaply, than indexing based on a controlled vocabulary. However, a critical problem in constructing an automatic natural language database is textual accuracy. This has led investigators into the field of computational linguistics in an attempt to impose various rigorous linguistic controls on natural language systems. In dealing with various linguistic levels, one must consider the morphological level at which individual words can be recognized and formed as well as the syntactic and semantic levels in which the sentence's grammatical structure can be identified and interpreted. It is important, therefore, to be aware of the automatic methods currently used to process natural language texts, although the full scope of language understanding may not be needed in information retrieval. In the context of this study, the various levels of linguistic methods as well as the commonly used grammatical theories will be tested in relation to the Qur'anic text. Also the main features of the Arabic language will be analyzed in relation to current software programs.

### **1.5. Methodological Approach**

The present study attempts to relate the ideological value content presented in the Qur'anic text to the empirical facts established in the mission-oriented era. The study is neither philosophical nor empirical, but is rather based fundamentally on an analysis of, and suggested solutions to, those problems that are likely to confront the compiler in his analysis of the issues faced by scholars seeking access to the scriptures as well as those problems that are likely to face compilers seeking to construct a retrieval system in relation to the Arabic language, both in terms of a manual and an automatic system (Figure 1-2). Also, there is an attempt to determine guidelines for compilers on the basis of recommendations and principles based on a discussion and analysis of the problems presented. However, as the fruits of this work are intended to be taken as the initiation of an analytical subject indication of the Qur'anic text, examples are provided on the basis of natural science and social science disciplines.

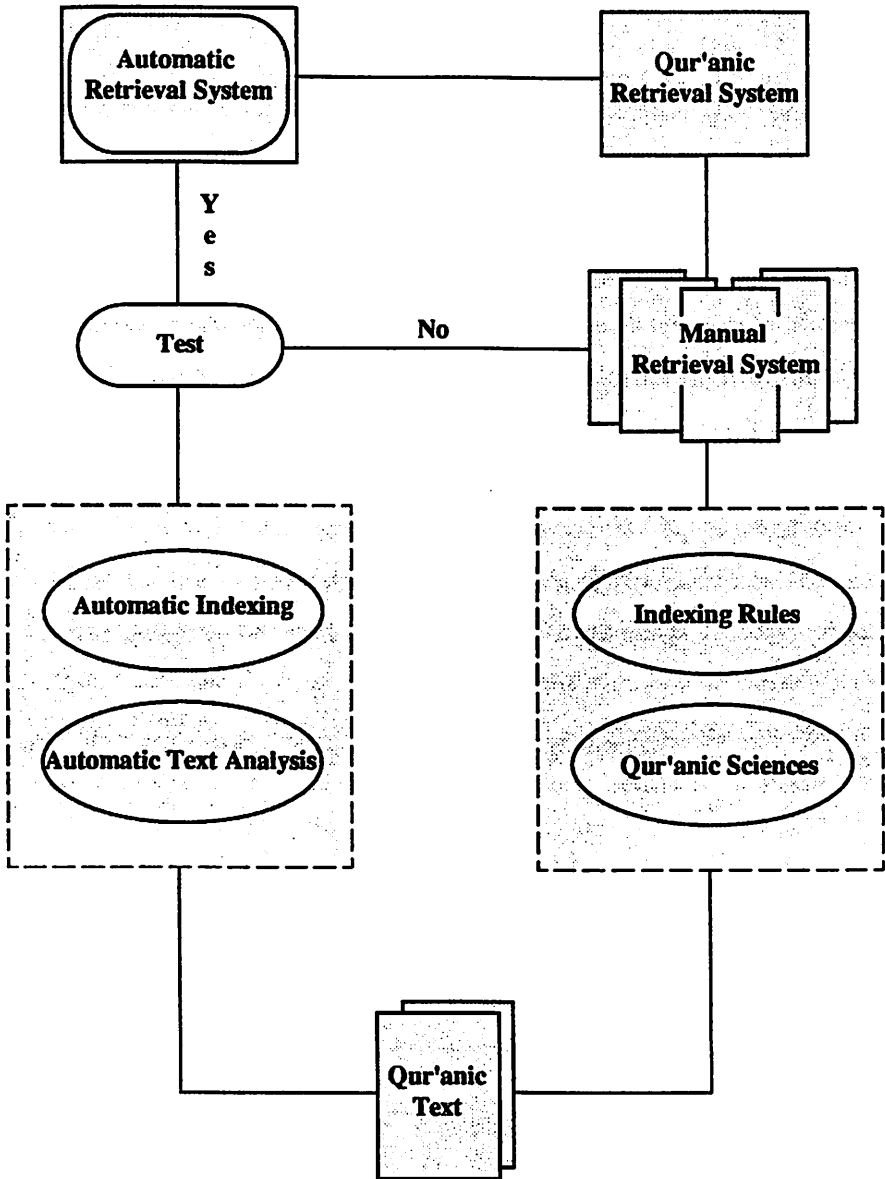


Figure 1-2: The chart represents the steps that have been taken through this study to initiate a Qur'anic retrieval system.



In this study, it is not practical to conduct experimental work on all aspects. Thus the choice has been narrowed to accepting such appropriate prerecommendations as principles of interpretation, norms of the Arabic language, suggestions of information retrieval specialists, and preestablished design rules of computer specialists. In practice, of course, many design decisions will be taken in line with these recommendations. Preliminary decisions about system requirements and constraints will color the development process and, therefore, may influence strongly the design of the retrieval system under construction. Simply knowing, in formal terms, what the system will look like is not enough, for the design guidelines will also determine the type of people needed for the work and how they will perform it. Finally, examples are proposed as a useful intellectual exercise to encourage compilers to check whether all significant points have been taken into consideration during the analysis.

The following chapters are organized in such a way as to confront the problems and the solutions encountered in attempts to design a Qur'anic retrieval system.

The second chapter highlights the major philological, historical, and theological considerations, as indicated by various interpretations, and tests the impact of these opinions on the retrieval system's performance. The third chapter analyzes the function of vocabulary control as applied to Qur'anic terminology and measures, through the use of known experiments, the variables that control the compiler's judgement in selecting the terms. The fourth chapter examines treatments in the computational analysis area that have a bearing on the Qur'anic style of calligraphy and structure. It also focuses on the function of the man-machine interface in the interests of user satisfaction and the content representation of Qur'anic software. The fifth chapter presents the guidelines and recommendations of the Qur'anic retrieval system. Finally, the sixth chapter applies examples for the Qur'anic retrieval system as suggested in the natural and social sciences.



## About This Book

The feeling that Islamic literature deserves and requires an Islamic classification scheme has been met with a widespread dissatisfaction among information scientists in the Muslim world. However, indexing the classical Qur'anic exegeses and Ḥadīth collections still requires an initial operation list of subject headings of both the Qur'anic and Ḥadīth texts.

This book is based on an investigation of the terminology in the Qur'anic text for the purpose of designing a retrieval system. It makes use of conceptual verses and words as partial examples for the required task. These examples are used to test the factors affecting the design at both the documentary and the computation levels. At the documentary level, the examples are used to examine the effects of Qur'anic terminology on the commentators and to see how it affects the performance of the retrieval system. Also it examines the characteristics of the Qur'anic vocabulary, against the problems known to be encountered in constructing an efficient information retrieval system. On the computation level, the examples are used to examine the possibility of the Qur'an, in its stylistic form, being processed by the computer. As a result, the study offers guidelines and recommendations with two examples for the natural and social sciences, as a model for constructing a retrieval system for Qur'anic text.

