# Basic Model For Information Retrieval Based On Free-Text In Database Application

## Mohd Adan Omar[1], Md Zahir Mat Cha[2], Mohd Rushdi Idrus[3]

*[1,2,3] Graduate Department of Information Technology*
*College Arts and Sciences, Universiti Utara Malaysia, 06010, UUM Sintok, Kedah*
*e-mail: adan@uum.edu.my, zahir@uum.edu.my, rushdi@uum.edu.my*

### ABSTRACT

*The purpose of this article is to design the basic model for information retrieval based on free-text domain in a particular database environment. It is an alternative method to the usage of key that is widely used in today database applications. At the end of this paper, we show the prototype that has been developed to support the free-text model verification. The result from the testing process is also given.*

### Keywords
*Information retrieval, Free-text, Basic Model.*

### 1.0 INTRODUCTION

Researchers in the field of information retrieval spend a lot of time in finding the technique to optimize the abilities of index and tool in retrieving information needed in the shortest period. Several techniques used such as using algorithm, key word and structured query language. Research toward information retrieval in database system focuses to the usage of the key as retrieval sources. As alternative to the key use, information retrieval using free text domain in database environment is proposed. The literature review shows, several theories relate to the usage of free text as information retrieval domain. However, most of the views discussed the retrieval in Internet environment.

Until the research conducted, there is no model for information retrieval based on free text in a database application except Internet based information retrieval. So, the research focuses to the technique of data retrieval using free text without link to any key attribute as primary key in database application environment. Most of the data search in database application individually or multi user environment will link to the key as main field. This approach, sometime gives the implication that brings to system halt, slow and fail to search information needed in certain period.

So the objective of this research is to develop and propose the model of information retrieval based on free text domain in database application. The model proposed will help in providing an alternative to data retrieval process using free text without tied to certain field. Used to free text as

alternative key, the process of information retrieval directly browses the data table without rely on unique prime key.

### 2.0 INFORMATION RETRIEVAL

Information retrieval is an approach to allow user makes item searching that knows already exist in a specific subject. The ability to search information toward specific topic is important part of information retrieval system (Brinkley and Burke, 199+,). There are two early models of information retrieval which are query based information retrieval model and browsing information retrieval model (Fox, 199.,. Information retrieval model using query is based on classical model as shows in Figure ) (Belkin and Croft, 1012 Fox, 199.,.
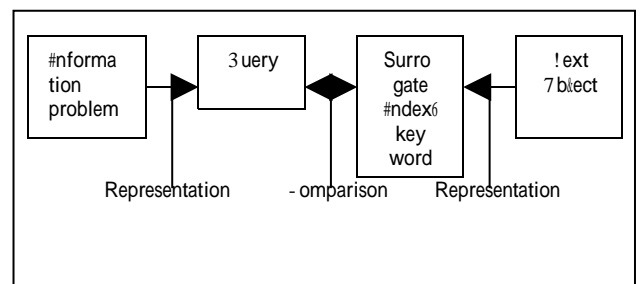


*Figure 1:* Query Model for Information Retrieval Sources4 (Belkin and Croft, 1012 Fox, 199.,

In this model, index and key word was used. Researches after this time use this model as reference model as Bate ()*0*, and Cooper ()*00, concluded that, although this model work well and can be suit with current technology, but lot of problem existed in it application such as, difficult to formulate query and a variety of important query components and also term dependency. So that Borayska and Fox ()**., added a system based on query formula contains limitation which leads to mismatch between human and information processing.

Fox ()**., and Al Hawamdeh et al. ()**), in their study proposed browsing paradigm for various task in information retrieval because of its effectiveness and

efficiency. #n browsing model, the usage of database is independent where user does not need to know the query formula and will move to whole of database. !he model proposed by -ox ()**.,, provided the interactive searching activities and the system user determined the search direction based on system respond. "owever, both models show negative usage of free text in their information retrieval system but the original concept of free text information retrieval is suit with browsing base information retrieval.

## 3.0 FREE-TEXT INFORMATION RETRIEVAL

!he method of information retrieval using full text is normal several decades ago. At the beginning, the method was using key word which refers to the information wants to retrieve. !he latest approach was using word indexes which record the information position in database. !he list of indexes allowing retrieval process toward all data documents rely on matched query and index ( 8 ikipedia, 9::+,. #t is clear that, key words, indexes and queries was becoming the key component in this method. !he next version of method was trying to trim the retrieval component usage such as key word, index and query. -hisholm and ; oida ()***, for example proposed <ector Space %ethod to look the frequency of word in a query. -h=vez Arag>n et. al. (9::?, developed @ovel "ybrid %ethod to strengthen the index ability until object image.

8 en (9::?, said, information retrieval based on text in web environment using search engine still used key word and index as reference resource. %ost of search engines in web environment are based on 'oolean Retrieval %odel where the user needed to determine a query as 'oolean combination for search word (Aones and 8 illet, )**1,. !he major problem in this method was the difficulty in formulating the query. So researcher try to replace to new method called information retrieval based on concept (Alsaffar et al., )***,. !he study about the system of information retrieval based on concept was introduced by %c-une et al. ()*0+, through the system called #nformation Retrieval 'ased on Rule. /anguay and Raghavan (9::?, introduced information retrieval based on concept but the definition of concept was generated by decision tree and allowed user to choose the definition fixed by expert.

Searching system using free text allows user to find and retrieve the information (document, by using word combination while the system uses key word need index to scan all document in database. #n the latest system, user will retrieve the information using controlled word combination as predicted by 'lair and %arron ()*0+,. Retrieving the information using free text means in the database, user will key in the relevant word without specifying the location of word fields (etde.org, )***,. Bimmerman ()**:, concluded that, the main key of free text in database is it ability to browse fast through many items. !he information retrieval based on free text proposed in this study is specific to database application

with the use of idea in document environment and also web search engine.

## 4.0 THE MODEL

!he literature review shows that free text mainly used in information retrieval to get the document such as conference proceeding, journal and article in library environment. !here is no (during research period, free text in database environment developed by tables. "owever, an S3C still used in this model, but combine together with algorithm to allow free text used as searching subject.

!he approach for interface is similar (in concept, with practiced in search engine where words can be typed in searching space and result will be displayed orderly. 8 e adopt this approach into the process of information retrieving using free text. A space to fill word on a form is provided and search manager will do the browsing process into database tables. #f in search engine, user can choose the information needed by clicking the link, free text environment also allows the same process.
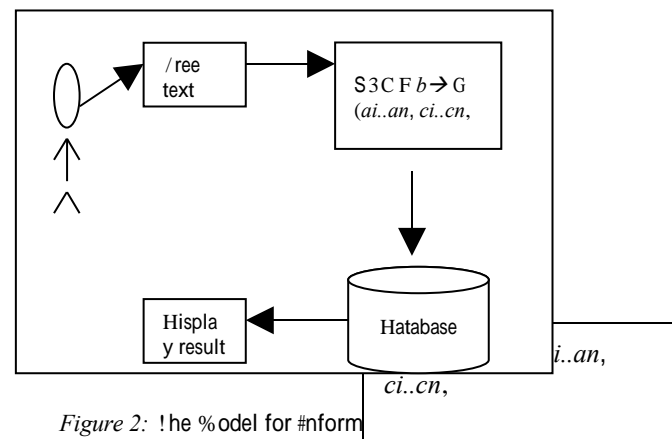


*Figure 2:* !he %odel for #nform 'ased on /ree !ext

/ig 9 shows the model for information retrieval based on free text in database environment. /ree text box refer to the process text entry for searching purposes. S3C still be used together with retrieve algorithm where $b$ is referred as text entered bye the user. Searching system will browse through tables and cells in database is showed by $ai..an, ci..cn$. Retrieval will be made if matched text occurs. !he result will be displayed in order form and user can click the order for specific information.

## 5.0 THE PROTOTYPE

!o prove the model produced in this research, we develop a prototype which focuses on free text search module. !he prototype developed together with database for Salary System. 8 e used %icrosoft <isual 'asic D.: and %icrosoft Access as development tool. /ig ? shows the prototypeЕs search menu module interface.
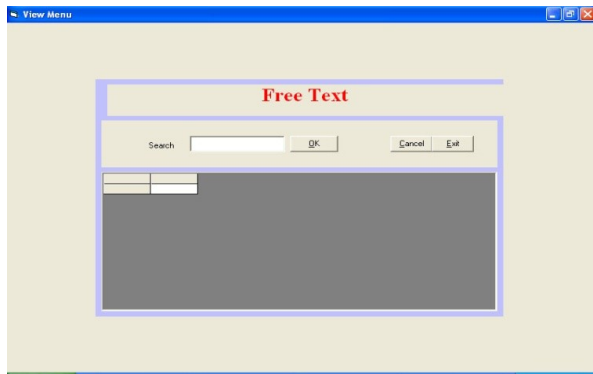
*Figure 3:* Search %enu %odule #nterface

## 6.0 THE TEST RESULT

!he purpose of testing process is to make sure the prototype work properly and support the model of information retrieval based on free text as proposed. !he !esting is also to see whether the developed prototype is able to generate search results as describe in the model. 8e do experiment by entering text in free text space on the form and record the result. !he example of test experiment shows in /ig ... !he result of selected experiments is show in !able ).
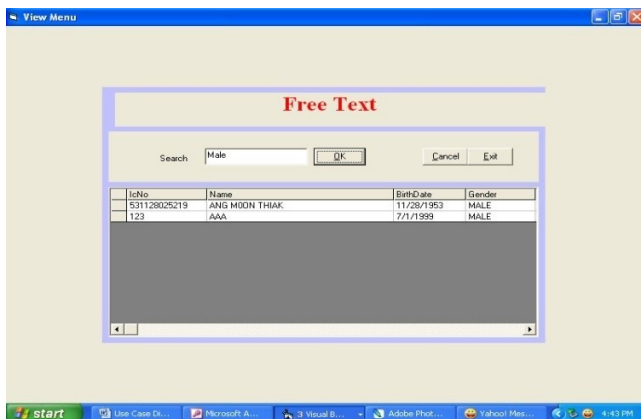


K%aleL text has been entered into search space. Search result shows two rows which contain word K%aleL in

*Figure 4:* Ixample of test experiment

!able )4 *Prototype Testing Experiment Result*

| !esting | Ixpected Result | Actual Result | -omment |
|---|---|---|---|
| $ser click search menu | /orm will be displayed | Jes | |
| | | | |
| $ser enters the word | Ixpected result will be displayed | Actual result displayed | |

| | | | |
|---|---|---|---|
| K%aleL | Jes | Jes | |
| K):::L | Jes | Jes | |
| K)::L | @o | @o | |
| KAngL | Jes | Jes | |
| KSintokL | Jes | @o | @o data |
| KAohorL | Jes | @o | @o data |
| K):L | @o | @o | |

## 7.0 CONCLUSION

#n general, the ob&ective of the research is to develop the basic model for information retrieval based on free text in database environment. After we develop the prototype to prove the model and, we do the testing experiment to validate the prototype and the model. "owever several aspects do not touched in this research such as the ob&ect of image instead of text. !he developed model is able to improve the aspect of information retrieval and help in database system development.

## REFERENCES

Al "awamdeh, S., de <er, R., Smith, 5. M 8illet, N. ()**),. $sing @earest @eighbor Searching !echnique to Access /ull !ext Hocument. *Online Review,* <ol. )+(?6.,, pg. )1? )*).

Alsaffar, A., Heogun, A., Raghavan, <. M !raylor, H. ()***,. Inhancing #nternet Search Ingines to Achieve -oncept 'ased Retrieval. *Forum 1999-Improving the visibility of R & D Information.* 7ak Ridge.

'ates, %.A. ()*0*,. !he design of 'rowsing and 'errypicking !echnique for the 7nline search #nterface. *Online Review*, <ol. )?(+,, pg .:1 .9..

'elkin, @.A. M -roft, 8.'. ()*01,. Retrieval technique. *Annual Review of Information Science and Technology* (AR#S!,, <ol. 99, pg ):* .0*.

'lair, H. -. M %arron, %. I. ()*0+,. An Ivaluation of Retrieval effectiveness for a /ull !ext Hocument Retrieval System. *Communication ACM.* Available4 http466portal .acm.org.

'rinkley, %. M 'urke, %. ()**+,. #nformation Retrieval from #nternet4 An Ivaluation of the tools. *Internet research: Electronic Networking Applications and Policy*, <ol +, @o ?, pg ? ):. Available4 http466www.emarald.com

-h=vez Arag>n, A., %a, A. A., Starostenko, 7. M Behe, A. (9::?,. @ovel "ybrid %ethod for #mage Retrieval by 7ntological Hescription of Sub regions.

*Proceeding of Artificial Intelligence and Doft Computing 2003*. Acta Nress.

Cooper, W.S. ()**0,. Setting Beyond Boole. *Information Processing and Management*, <ol. 9.(?,, pg. 9.? 9.0.

Cox, ;. ()**.,. *Information Retrieval by Browsing*. City Nolytechnic of Hong ;ong.

Itde.org. ()***,. *Free-Text of Fields Searching*. www.etde.org.

Languay, R. M Raghavan, <. (9::?,. Senerating Rule Based Trees From Hecision Trees for Concept Based #nformation Retrieval. *Applications, Products and services of Web.*

Sorayska, W. M Cox, ;.R. ()**.,. *Expert System as Extensions of the Human Mind in Artificial Intelligence and Society.*

Aones, ;. M Willet, N. )**1. *Introduction to Chapter Five. Reading in Information Retrieval*. %organ ;aufmann., San Fransisco.

%c Cune, W., Tong, R., Hean, A. M Shapiro, H. ()*0+,. RUBRIC: A System for Rule Based #nformation Retrieval. *IEEE Transactions on Software Engineering*, <ol. ))49, pg *?* *...

Wen, J J. (9::?,. *A Search Engine for Image. Project Report*. Faculty of Computer Science, University of %aine.

Wikipedia. (9::+,. *Keyword Searching*. www.wikipedia.org

Bimmermann, %. ()**:,. *Notes of Free-Text Information Retrieval*. www.his.com