

A Criteria to Evaluate the User Friendliness of Search Interfaces of Text Retrieval Packages based on End User Behaviour

Wathmanel Seneviratne
Librarian
Central Bank of Sri Lanka

In its modern perspective Information retrieval is a process, which is becoming increasingly dependent upon physical mechanisms. When the computer retrieval is involved Information Retrieval System (IRS) consists of both hardware and software. The predominant feature of such software packages is the Search Interface, which is also called 'User Interface' that permits non-programmer to expert user to comprehend and interrogate the data stored. In other words the interface mentioned so should be able to accommodate any kind of user comfortably.

The Search Interface is the place where the Man-Machine Dialogue has taken place physically, conceptually, perceptually. Search Interface should provide the user with an essential overall feature called 'User Friendliness' for keeping this dialogue effective.

As the popularity gained over the years by the Information Retrieval Packages, the term User Friendliness became the most over used phrase in the software industry, since the concept is treated as the main feature of selection criteria for software packages. the criteria of user friendliness focused around the user's mental model which is important not only in determining the success of a search interface but also in determining the design of a search software. The friendliness is achieved when inexperienced users find the system to which they attach is easy to learn and easy to use while experienced users are able to use it quickly and efficiently. Galitz defines the qualities of easy to use software as adaptive; transparent; comprehensible; natural; predictable; responsive; self explanatory; forgiving; efficient; flexible and available.

This criteria to evaluate user friendliness of the search interfaces was developed using an experimental study conducted on End User Behaviour in an actual search environment. A prototype database was designed for the purpose. Two search interfaces of two text re-

trieval packages were used for the study (CDS/ISIS and INMAGIC).

The evaluation criteria was designed as mentioned below.

1. Users studied were divided into categories as mentioned below.

Expert Users - the category consists of users who has experience in computerized systems and are experts in the computer field.

Intermediate Users - a group consists of users who has moderate knowledge in using computers.

Novice Users - This group consists of ordinary users who has no prior experience or knowledge is using computers.

2. The search interface is divided into broad aspects and under each broad aspect specific attributes were separated for easy evaluation as follows. These broad aspects were decided observing the behaviour of the users when proceeding towards searching.

1. Logging on to the system
2. Searching
3. Help and error messages provided
4. Representation of the out put

Sub attributes studied under each broad aspects above are indicated in a chart at Fig.I.

3. Behaviour of the users of three categories were observed and evaluated against the features or attributes that are considered essential to be available in a search interface.

4. The users selected for the experiment were given pre tested and equal number of searches. In this experimental study the users were given three pre-tested searches which were to be performed within nine search attempts. The number of search attempts shall be reduced or increased depending on the extent of evaluation.

MAIN ASPECTS STUDIED FOR A SEARCH INTERFACE OF TEXT RETRIEVAL SOFTWARE

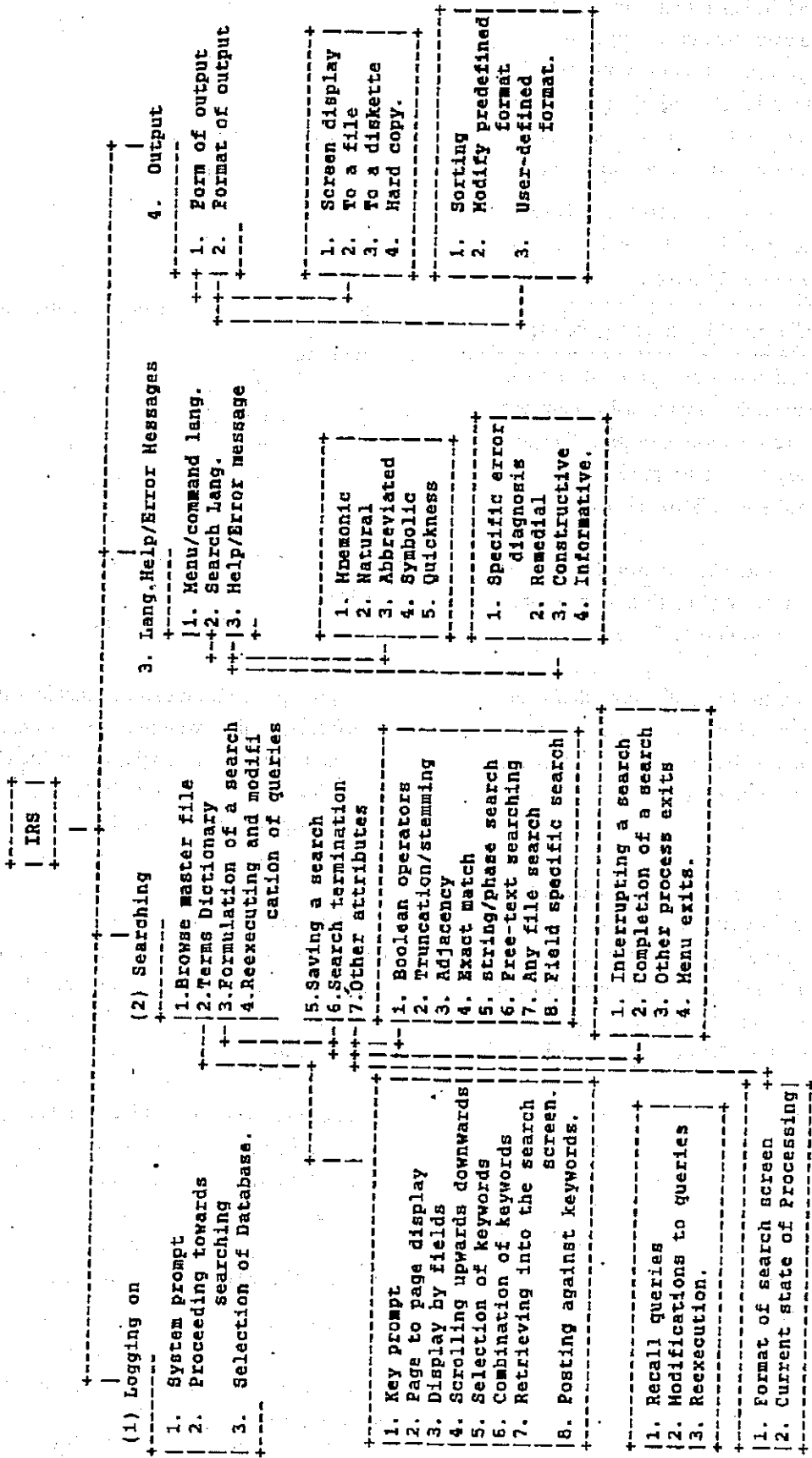


Fig. 1

Method of analysis

Searching pattern and the behaviour of end-user were observed in obstrusive way. Search attempts made by the users, their moves, gestures and comments were also observed and studied while responding against each function available in the search interface. This analysis was named as “ **Measures of Search in Progress** ”.

The hit rate (number of search results obtained in each search session) and time spent for each search were also analysed as “ **Measures of Outcome** ”.

Measures of search in progress:

The user was observed from logging on to the system through performing expected searches, using help and error messages and representation of output. The performance of the searches were measured using two basic parameters, **Certainty & Uncertainty**.

The clearness and confidence the user show when passing or attempting a step or function in the search interface was named as ‘Certainty’.

The other parameter is the level of intermediary help sought while searching, which means the user is not certain or do not feel free to use the options available in the user interface. Hence the parameter is named as ‘Uncertainty’.

The Certainty and Uncertainty were measured as signing weights to each move made by the user against each step or function he passed in the searching process. Clearness level and the Weights assigned for each move are given below. These weights are assigned to each attribute as the user travels through the interface.

NH	-	No help given	LH - Little help given
MH	-	Maximum help	SP - Stopped progressing (as no help available)
NA	-	Function is not available	

CERTAINTY WEIGHTS

where,	C	=	3
	NVC	=	2
	NC	=	1
	NA	=	0

UNCERTAINTY WEIGHTS

where,	NH	=	0
	LH	=	1
	MH	=	2
	SP	=	3

The weights were used in the following way:

When the,

P = C,	H = NH
P = NVC,	H = LH
P = NC,	H = MH
P = NA,	H = SP

where,

P = Performance of Certainty
H = Intermediary help sought.

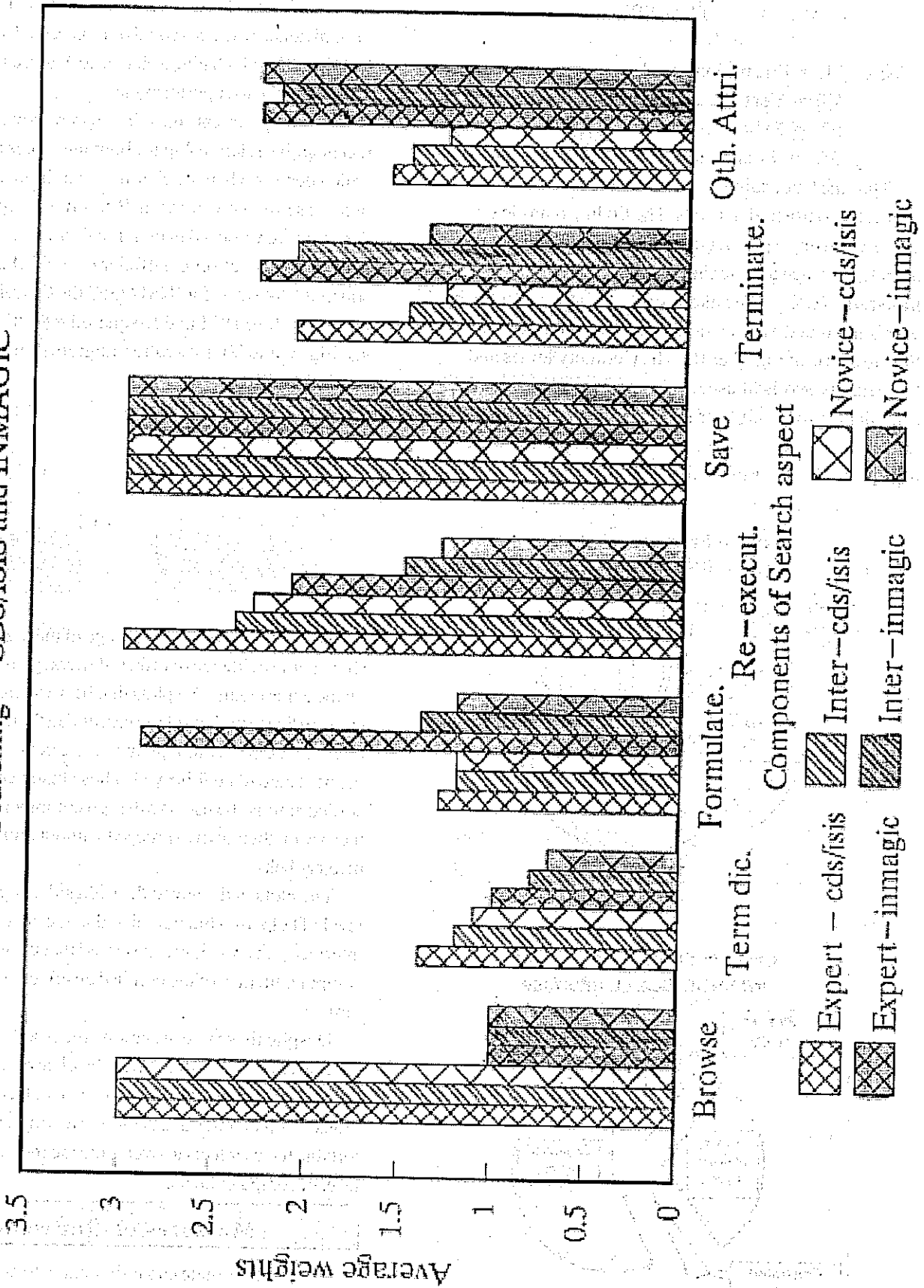
When expected features are available within the search interface, Certainty weights, C, NVC, NC and NA are assigned to the each step passed by the user. Accordingly Uncertainty weights NH, LH, MH were assigned when a help is available and SP was assigned when help is not available.

All the features and attributes of the search interface (please see Fig.1) were observed and measured for clearness level. In this experiment, selected two TXT packages, CDS/ISIS and INMAGIC were made available respectively to each user selected from the sample. The level of clearness or the Certainty level highly represented the level of friendliness of the search interface. It was observed that, the inexperienced user was not able to get through some functions when it is not very clear as expected from a friendly interface. This measurement is also complemented by the Uncertainty level measured based on the level of intermediary help used.

Uncertainty level measurement leads to calculate the “Frustration Level” (FL) faced by the user when using a search interface. The values could be obtained in the following way.

When H = NH,	FL = VP
H = LH,	FI = NVP

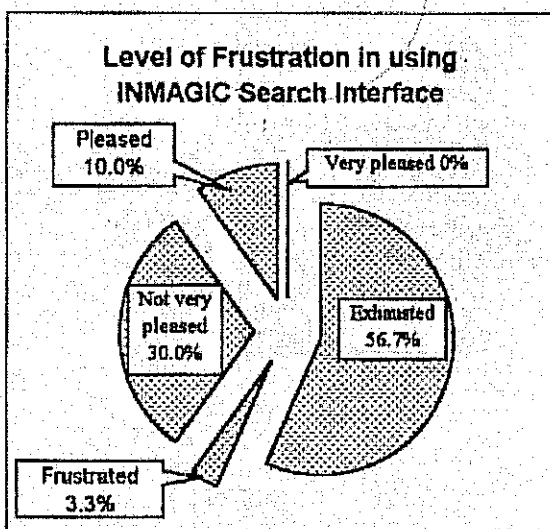
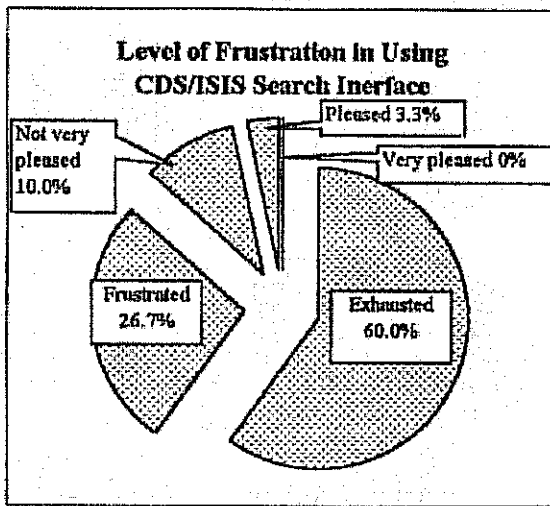
Certainty weights obtained by three user groups
Searching - CDS/ISIS and INMAGIC



H = MH, FI = EX
 H = S, FL = FR

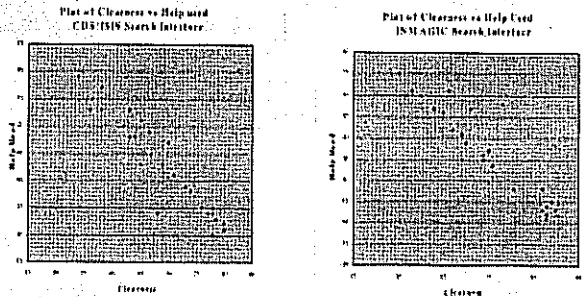
Where FL = Frustration Level
 VP = Very Pleased
 EX = Exhausted
 FR = Frustrated

Thus the Uncertainty weights could be assigned and connected to frustration levels. The findings can clearly be viewed using a pie diagram. If "Very Pleased" (VP) mark level is high the interface is user friendly and if "Exhausted" (EX) / "Frustrated" (FX) level is high the user is frustrated with an unfriendly user interface. See the pie charts at Fig. II & III. The Certainty levels and Uncertainty levels of usage of users were represented graphically for each user category in Fig. IV. for "Searching Aspect." Other aspects are also calculated and can be shown graphically



The relationship between Certainty and Uncertainty levels could be calculated using Correlation Coefficient and the individual evidences could be represented using Scatter Diagrams which give a chance to analyse the factors affected to the user performance.

In this experiment, the relationship shown was a direct negative relationship for both interfaces observed, which depicts when the clearness was high, help level used was low and vice versa. But some exceptions and deviation were noticed in individual cases. Eg. for CDS/ISIS the Correlation Coefficient = 0.972827 and R-squared = 94.62%, for INMAGIC the Correlation Coefficient = 0.931821 and R-squared = 86.83%. Please see Fig. V and VI for scatter diagrams for CDS/ISIS and INMAGIC.



The plots lied outside the range of the regression line show some evidence of effect of literacy level of users on the experiment. The plots that lie away from the Regression line to the left side represent that the users utilised low level help eventhough they were not certain about the next step to be followed. These instances were recorded among the users with high computer literacy as they try on their own to proceed without online or intermediary help.

The plots lied away to the right side, depicts higher level of help used but not related to the clearness of the interface. Anyway it was observed that the low level of computer literacy of the user had an effect over this performance.

Despite these two situations recorded overall performance indicates that when the clearness was high, help level used was low and vice versa. Hence it is obvious that if an interface is more friendlier, more comfortable the user feels to overcome the incompetence in level of computer literacy.

Measures of Out come

The ultimate capability expected of a retrieval system should be its ability to produce the desired output. Measuring this capacity of a system should be de-

signed very carefully as the evaluation deals with the satisfaction derived by the user against the search results obtained by using that particular system.

Measuring the outcome of the search should evaluate not only the retrieval capacity of that particular software package but also the searching behaviour of the user including probable mental model of the searcher relating to each search.

Measures of Outcome designed in this experiment were only based on the searching behaviour and the users' mental model. Two methods were used to obtain desired evaluations.

(1) Designed three searches confining to the most searched aspects by users.

Author search (eg: "Wickrama singhe, L.A.")

Title search (eg: " ")

Keyword search (eg: "AS
BESTOS")

(2) Scheduled the exact hits (number of records available) available in the database for each search.

Process used to obtain data for analysis:

1. Users were allowed nine search attempts (three times for each search) for three pre-tested searches.
2. The hits obtained for each search attempt were recorded.
3. The hits obtained by users were compared and evaluated with the actual hits available in the database.
4. Time utilised were calculated for the first search and for the subsequent searches.

Analysis of hits:

Number of records recovered (Hits) in trial runs were analysed in three ways.

- a) Number of records retrieved were compared with the exact number of records available in the database for each search.
- b) Calculation of Irrelevant hit level.
- c) Calculation of Zero hit level.

Using b) & c) above, False Drop level was calculated for each user group which represent how the level of friendliness of a search interface affects the productivity or the relevancy of the hits recovered for a particular search conducted. This measurement indicates the level of incompleteness of the search help level provided as well as the desperation or satisfaction of the user in

working with the interface. In this particular study it was observed that when the interface is friendly all the users obtain low False Drop Level and vice versa.

The formula in Bradford's Law also can be used with this data to calculate the relationship of the hit level produced to measure the relevancy if friendliness of the interface to be analysed against the relevant hit level.

Analysis of the time utilised:

A formula was also worked out to calculate the time utilised for each search. The time fraction used by any user could be formulated as follows. The formula is based on the three search attempts provided for the sample users in the experiment, but may be adjusted according to the number of trials expected to be given to the subjects of any experiment. The criteria is formulated in the following way.

$$T = \frac{TN}{T1 + T2 + T3}$$

where, TN = Any consecutive search

T1 = First attempt of Title/ Author/
Keyword search

T2 = Second attempt of Title/ Author/
Keyword search

T3 = Third attempt of Title/ Author/
Keyword search

and;

$$T1 = ft + t + m + o$$

$$T2 \text{ \& } T3 = fs + t + m + o$$

where; ft = time taken the search for first trial

f = time taken to formulate the
search in progress

t = system time elapsed while
the search in progress

m = time taken to use
intermediary help

o = time taken to read online help
messages

User of this criteria could accommodate different types of users for any search interface using a fair time fraction and with pre-tested queries. It was also observed that the time utilised within permitted searches had shown remarkable decrease exponentially even with the average figures. However the familiarisation curve of the

user gives rising tendency with recurring search attempts provided. Hence this model analysis how fast a user gets familiar with the search interface after first search depending on the literacy level of the user.

In summing up the two measures used to study the friendliness of search interfaces, not only provide an evaluation criteria for user friendliness but also guide a researcher or a system analyst towards systematic and efficient arrangement of user friendly features within an user interface which is always essential to build a productive man-machine dialogue.

References:

Brian C. Vickery and Alina Vickery, Information science in theory and practice, London: Butterworths, 1987.

C.J. Van Rijsbergen, Retrieval effectiveness in information retrieval experiment, London: Butterworths, 1991.

Charles T. Meadow, The making of an Information retrieval Interface, International Forum on Information and Documentation. Vol. 16, No.3 (July 1991) 22.

D.Alasdiar Kemp, Computer based knowledge retrieval, London: ASLIB, 19

L. Trenner and A.B. Buxton, Criteria for User friendliness. In Proceedings of hte 9th International Online Meeting, Oxford: Learned Information, 1985.

Loran B. Doyle, Information retrieval and processing. Los Angles: Melville Publishing Co., 1975.

Peter J. Vigil, Online retrieval: analysis and strategy, New York: John Wiley & Sons, 1988.

Raya Fidel. Database design for information retrieval: a conceptual approach, New York: John Wiley & Sons, 1987.