WEB-BASED ACCESS TO ONLINE DATABASE VENDORS

by

LESIBA STEPHEN LEDWABA

Submitted in partial fulfilment of the requirements for the degree of

Magister Informationis

in the

Department of Information Studies

at the

RAND AFRIKAANS UNIVERSITY

SUPERVISOR: PROFESSOR P.A. VAN BRAKEL

June 2000

SUMMARY

This research investigated the role played by Web-based interfacing in improving online searching. A comparative analysis was undertaken to investigate end-user searching in both conventional online systems and Web-based services. The results of the analysis necessitated further improvements in Web interfacing. In fact, this study identified areas in which online searching poses problems and finally suggested features which need to be incorporated into further developments of Web interfaces to online systems.

DEDICATION

This work is dedicated to my parents Sepheka Mogale Johannes and Raesetja Maria. They were always the pillars for my education.

ACKNOWLEDGEMENTS

I would like to express my indebtedness and sincere appreciation to the following persons for their contribution to the successful completion of my research:

- Professor P.A. van Brakel, Chairman of the Department of Information Studies at RAU, for his constructive scholarly advice and unwaning patience throughout my study.
- Mr. J.K. Tsebe, University Librarian and Mrs M.M. Chuene, Head of Acquisitions at UNIN who, through their encouraging words served as inspirational forces in pursuing this project.
- My friend, Skepe Kgope for his encouragement throughout this study.
- Colleagues in the Acquisitions Section of the University of the North Library for their support and understanding during my absence.
- The University of the North for the financial assistance provided by the Research Administration Center.
- My wife, Mpheye and my son, Ronny for their unqualified moral support.
- Finally, to the Almighty God for always being there for me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND TO THE RESEARCH PROBLEM

The information industry has been growing at an alarming rate. This is partly due to the developments in information technology. Nowadays, it is possible to gain access to the references of millions of documents world-wide by means of a single network line.

The information industry is made up of players such as database producers, database vendors, information professionals, and end-users. These players function as a system and interact with other players to form the information industry. However, Estabrook (1981:1377) maintains as early as 1981 that in the past more emphasis in the information industry was placed on the supply of information products, and on the technology required to deliver information, than on the needs of end-users.

Producers compile and publish databases in specific fields. Most of these collections were originally intended for the production of printed indexes. The introduction of machine-readable formats had set a trend in that collections of bibliographic material could be generated electronically.

Among the large producers are the National Library of Medicine (NLM), and Institute for Scientific Information (ISI). The producer negotiates with the vendor to make its product accessible and available to the potential international market. It should be noted that database producers could at times also act as database vendors.

Database vendors form a link between database producers and information professionals, including end-users. Database vendors seem to be uncomfortable with the notion of being in the information industry for profit, as Dole (1987:129) maintains, they regard themselves rather as providers of information than sellers of their services or products.

These database vendors tailor their products to the end-users' needs through designing search software which end-users would find easy to use in manipulating the products. Through search software, database vendors co-ordinate and make accessible a myriad of databases through a common node and thus offering information from a variety of fields and different producers. Products and services supplied by database vendors include among others, CD-ROM and variable disc products as well as bibliographic, full text and real-time online information systems. Database vendors include inter alia, The Dialog Corporation, and LEXIS/NEXIS.

As indicated previously, database vendors form a link between database producers and information professionals. Information professionals, through their companies, subscribe to specific products from the database vendors. Information brokers on the other hand, subscribe to specific products from these vendors as well. The products subscribed to by these companies serve to meet particular information needs of end-users. On the other hand, end-users are consumers of information, that is, people who need information for decision-making, some of whom rely on intermediaries for information searching.

With current developments in the information industry, more and more information systems and services are being designed with the end-user in mind. This is because end-users do constitute a potentially huge and financially attractive market.

Nicholas et al (1987:109) indicated more than a decade ago that *"there is a danger that the very real advances in information technology might be lost on a public whose information-handling skills and perceptions are clearly neither known nor fully understood"*. On the other hand, Ross (1984:3) noted that the importance of end-user requirements has not been adequately appreciated other than the technical factors, which dominated the information industry until recently.

Ross (1984:3) further indicates that "... *the implications and ramifications of putting the user in direct touch with the data flow are being overlooked in the general rush to sell"*.

Through the advancement in information technology, more database products became available online, packaged in the form of CD-ROMs, or available via the Internet. This advancement in information technology set a new trend in the information industry in that a shift is now towards improved end-user access. This aspect was the driving force for the design of improved interfaces for end-user searching.

Though database vendors were engaged in designing "user-friendly" interfaces in the past, they seemed not to be sure of the friendliness of their interfaces, since more training manuals were produced to supplement many of these interfaces (Steele and Tseng 1992:56). As indicated earlier, today's developments in information technology has made it possible for end-users to interact directly with the database vendor's product.

1.2 STATEMENT OF THE PROBLEM

The introduction of the World-Wide Web (WWW) has prompted database producers as well as vendors to make their information products available via the Internet. In order to access these information products via the Internet, database vendors need to develop Web-based interfaces.

On the other hand, the WWW brings authors, publishers and end-users closely together without the intermediary role played by vendors (Tenopir 1998:35). The same author argues that if the user can go directly to a particular Web site, some people wonder why they should go through database vendors such as The Dialog Corporation or LEXIS/NEXIS and pay their high charges.

On the contrary, other people recognise the value of database vendors.

Therefore the problem identified in this paper can be formulated as follows:

"How is Web interfacing improving online searching?"

The following sub-problems can be derived from the research problem:

- The first sub-problem focuses on end-user's needs for online information retrieval and how these needs were addressed in the past. That is, to what extent were the traditional features end-user directed?

- The second sub-problem is concerned with initial improvements on typical interface features. It is an overall view of the basic features inherent in front-end systems.

- The third sub-problem focuses on Web-based interface features. Web interfacing is so far a widely used medium in online searching.

- Since the Web has been introduced, how are the traditional features of online vendors being accommodated via a Web interface? These issues are the focus of the fourth sub-problem.

1.3 RESEARCH METHODOLOGY

Literature surveys were conducted to investigate the role of database vendors in the online industry with regard to end-user searching. The nature of Web-based interfacing was studied as well as how this interfacing differs from traditional interfacing. Interface features of both conventional online services and the Web were compared to determine how end-user searching was done in the past, and at the present, as a guide for future end-user searching.

1.4 SUMMARY OF CHAPTERS

The research project is organised as follows:

Chapter 1 is an introduction giving a background to the study and a statement of the problem.

Chapter 2 discusses how end-users' online information retrieval needs were addressed by database vendors in the past. Included in this chapter are the features and requirements of a typical interface to a computer system or rather software in general.

Chapter 3 discusses the development of front-end software. Technological developments leading to the introduction of front-ends were also investigated.

Chapter 4 focuses on Web-based interfaces. This chapter also concentrated on how end-users' information needs are accommodated via Web-based interfaces.

Chapter 5 describes the methodology used to compare conventional online services and Web-based interface features. The results of such a comparison were presented and recommendations geared at future user interface development are provided.

Chapter 6 includes a summary of the report. Recommendations for further research areas are provided.

CHAPTER 2

BASIC FEATURES OF COMPUTER INTERFACES

2.1 INTRODUCTION

This chapter investigates the features of a typical interface to computer systems. These features are essential to this study because they explain how end-users interact with computers. The focus of this chapter is also on the traditional approach to human-computer interaction. The interface concept appears in different forms. For example, when two people are talking face to face with each other, they are in a way interacting, and therefore there is an interface between them. The focus of this study is basically on human-computer interaction. The purpose of this chapter is to investigate how end-users' online information requirements are being addressed through the typical features of computer systems.

A human-computer interface is software that enables the end-user to use what is available in the system, formulate specific requests for information and receive and display the results of a search. The introduction of graphical user interfaces (GUIs) have brought improvements in interface software in the sense that it is much easier to understand and use systems, and enable many more tasks to be performed.

Interfaces can be customised to meet the needs of different types of end-users, that is, from expert to novice users. Since the focus of this chapter is on the features of a typical interface to computer systems, it is import to first understand what interfaces are.

On the basis of the developments in human-computer interfacing, Ratnakumar and Haravu (1994:15) defined a user interface as an interactive system through which a novice user can communicate with the software irrespective of its complexity, with little or no knowledge of it, and exploit its capabilities. In simple terms, a user interface refers to the methods and devices that are used to accommodate interaction between machines and the human beings who use them. Despite taking many forms, user interfaces therefore accomplish two fundamental tasks, namely:

- Communicating information from the machine to the user, and
- Communicating information from the user to the machine.

Furner (1997:9) remarked that the goal of a user interface is to *"...provide effective support for the interaction, or communication of instructions, requirements and responses, that necessarily takes place between human user and the retrieval mechanism"*.

The balance of this chapter consists of an overall view of the basic features and requirements of a typical interface to software applications.

2.2 EVOLUTION OF USER INTERFACES

The first computers had user interfaces that were as rudimentary as the computers themselves. The implication was that only trained computer specialists could actually communicate directly with computers. These interfaces were not meant for end-users.

In the next stage of this evolution, computers were communicating to users through printing devices, and the users communicating to the computer through punch cards. This improvement was rather cumbersome and inefficient because it was still necessary that computer specialists communicate directly with computers.

There was a major breakthrough when video screens or monitors were used to communicate information between the computer and the user and typewriter-style keyboards were used to communicate information to the computer. This innovation was a major step forward for ordinary users to communicate directly with computers. Video screens were limited to displaying only the characters that were found on keyboards, and the usefulness of user interface was thus constrained by the same limitation. Despite the limitations, a great deal of training was still required in the use of the computer. User interface was improved by the introduction of graphic user interfaces (GUIs). Therefore limitations which were inherent in video screens were significantly reduced.

2.3 GRAPHICAL USER INTERFACES (GUIs)

A GUI is a platform in which graphics are used during the interaction process. The invention of GUI complemented text-only interfaces which were common in dumb terminals. Two major factors that separated the new trend from the old were:

- The use of graphics to communicate information to computer users.
- To present a number of options to the user rather than requiring to memorise and manually enter commands from a virtually unlimited set of options.

Unlike character-based interfaces, the GUI allowed an interface to be focused on end-users. Therefore, the learning curve was significantly reduced. GUIs are a natural choice of interfacing because they are user-friendly, colourful, window-based approaches which rely on icons and natural language processing (Head 1997:21). The introduction of GUI has set a trend for online services to redesign their old command-based software to encompass GUI elements. However, these GUI elements do not ensure that interfaces which incorporate them will be intuitive and easy to use.

For example, the *"Home"* icon can be confusing in the sense that a novice user might think of it as referring to the top of the page.

GUIs, whether they are software or hardware interfaces, are end-user directed. They focus the user interface on the needs of end-users rather than mandating that end-users conform to the needs of the computer. Head (1997:22) argues that a well-designed GUI reflects the following functional components:

- Multiple windows that can be opened simultaneously for multi-tasking.
- Icons that allow for direct manipulation.
- Navigation by pointing or clicking with a mouse.
- Consistency across applications that ensures reusability and fluency in future versions.

It has been indicated in the second paragraph (under 2.3) of this Chapter that GUI elements do not ensure that interfaces which incorporate them will be easy to use. This is because some of the icons might be confusing to novice users. Ease of use is important because it contributes to the user's acceptance of an interface, especially in the case of a novice user.

Ease of use also indicates limited power in searching. The easiest option in searching WilsonDisk, for example, is through subject headings. As noted by Tenopir (1997:35), "ease" may apply for the simple reason that search and print options are displayed on a toolbar or as an icon on every screen.

Users familiar with online searching should be able to click on a particular icon to start a new search, print, or go backwards, even if the icons and wording for each function differ. To other searchers, "ease" may lie in consistency. SilverPlatter, for example, recommends local loading so "users can simultaneously search multiple SilverPlatter databases while using the familiar interface, commands, and search methods of their local Z39.50 search systems" (Tenopir 1997:35).

The implication is that even in "easy-to-use" systems, it is not easy to know which function is needed next, to formulate a search strategy, to narrow or broaden a search, or to understand the content of a bibliographic record.

GUIs have design components that match the cognitive abilities, expectations, and limitations that end-users expect from a system. These cognitive design components include the following (Head 1997:22):

- Providing multiple methods for completing the same task. In other words, alternatives are provided and end-users are allowed to tailor their needs against their choices.
- Using icons that rely on end-users recognition of "real-world" objects. Recognition and association are easier cognitive tasks than recalling commands.
- Creating visual, auditory, or tactile feedback that quickly sends information back to the end-user about whether the system is processing, lagging, or not functioning at all.
- Enhancing screen visuals with colour, font, shape, arrangement and contrast. All these components help end-users focus attention on tasks at hand.

In view of the cognitive components, most visual elements of the GUI are thought of as idioms. A scroll bar, for example, is an entirely new construct and it performs an obvious function, that is, its operation is easily mastered and end-users easily remember how it works.

Davies and Hepworth (1993:151) maintain that with the developments in interface design, most of the online search services have enhanced their features, but noted that many were geared primarily to the needs of the information professional and thus remain intimidating to the potential user. Davies and Hepworth's statement contrasts with the current situation because nowadays the developments in online search services are more end-user directed rather than focusing on intermediaries, as will be indicated in Chapter 4.

The following categories of interfaces are examples of end-user approaches:

- Command-based systems.
- Menu-driven systems.
- Direct manipulation systems.
- Combination of commands and menu-driven systems.

2.3.1 Command-based interfaces

Command-based interfaces require the end-user to have a reasonable degree of system knowledge, necessitating the use of commands and symbols with very little contextual help (Harter 1986; Van Brakel 1988; Davies & Hepworth 1993). These interfaces provide the user with a prompt to which a response is made through a command line with necessary parameters, or a file name with user-defined commands.

Online systems such as The Dialog Corporation are traditional examples which allow the user to activate a command at the prompt or enter an option on one of the users shell programs (Ratnakumar & Haravu 1994:16). It should be borne in mind that shell programming techniques also allow the user to develop menu-driven systems. A point worth mentioning is that these information storage and retrieval systems, such as The Dialog Corporation, BRS and ORBIT all began as character-based systems.

The fact that a user needs to know the commands in searching an online system is a limitation on the part of command-based interfaces. Due to developments in interfaces, command-based interfaces were later simplified (see Figures 1.1 and 1.2). These are menus derived from ERUDITE and The Dialog Corporation systems respectively.

It is evident from Figure 1.1 that in order to search information from the menu, knowledge of the commands is necessary.

The *.T* command is executed when displaying and recording of temporary searches. Each command functions differently and as a result, end-users experience difficulties in comprehending all of these commands.

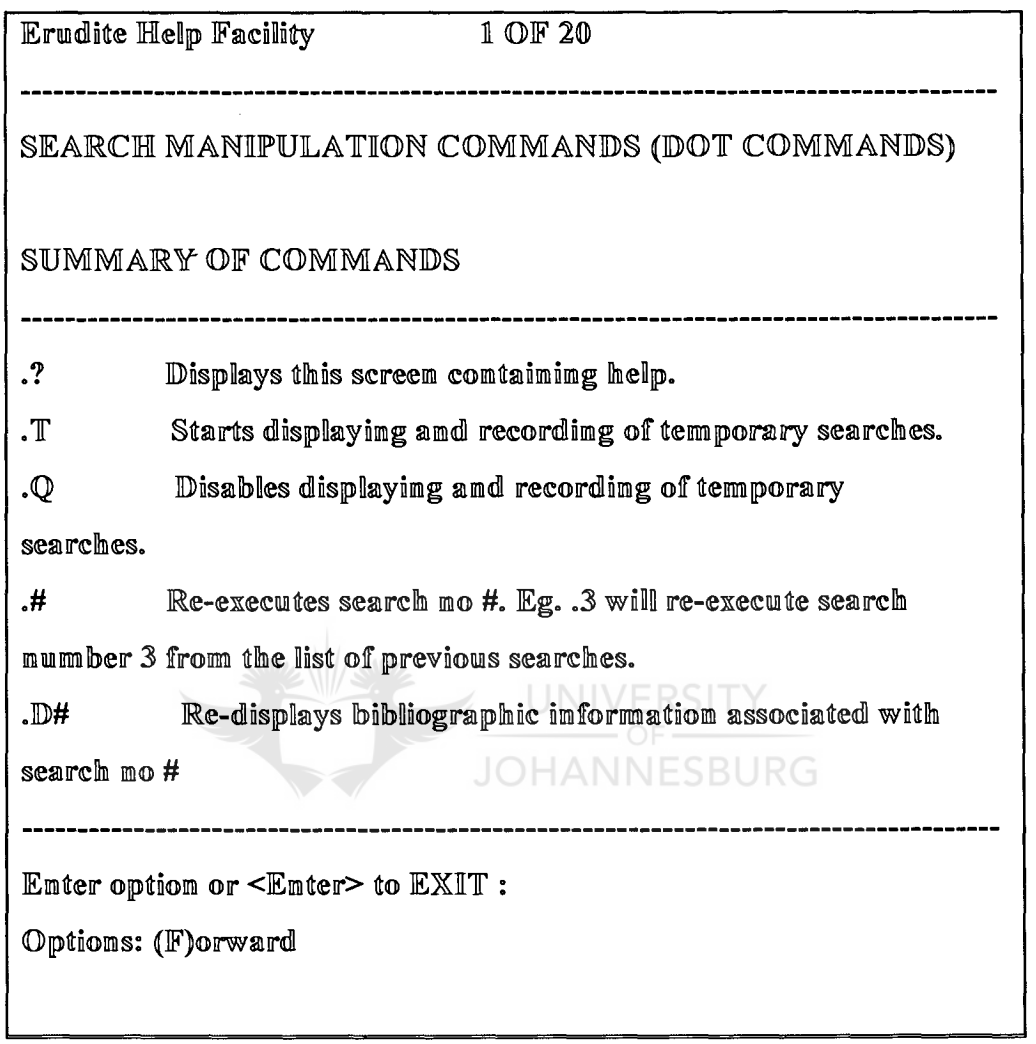

FIGURE 1.1: ERUDITE Command-based interface

Like Erudite's command-based interface (Figure 1.1), The Dialog Corporation has its own commands. An example depicted in Figure 1.2 is the *TYPE* command. Unlike in Erudite whereby the *.T* command was used to display search results, The Dialog Corporation requires a *TYPE* command to display results. It is therefore necessary for end-users to learn how to use these different systems.

TYPE

Example: View records 1-10 from set S3 in Format 5.

Command: t s3/5/1-10

The TYPE command displays your search results.

TYPE statements include the following elements:

Set Number

You can enter the set number of any set created since your last BEGIN command.

Format

The format determines which fields of each record will display. Available formats include predefined formats and user-defined formats.

Items

Item numbers specify the record(s) displayed by the TYPE command. You can specify a Single record, a range of records (1-10), non-sequential groups of records (1,4,7,12), or a Combination of a range of records and non-sequential groups of records (1-10,14,19).

Note: The number of items specified in your TYPE command cannot exceed 99 items. To view all the records in a set, if the set contains less than 99 records, enter the word ALL or a range that includes all the records in the set. If an item number is not included in your TYPE statement, only the first record in the set displays.

From

When searching multiple databases in OneSearch, you can specify file numbers in the TYPE statement by including the word FROM and the file number. For example, to TYPE Records 1-10 from set S2 in Format 6 from File 47, enter t s2/6/1-10 from 47. If you want records from each database in the session, include the words FROM EACH in the type statement. The FROM and FROM EACH options must be the last elements in the TYPE statement.

Accession Number

If you know the Dialog accession number of the record you want to display, you can retrieve it directly in most databases. For example, to display in Format 5 the record with accession number 1234567, enter t 1234567/5. To retrieve a OneSearch record using the accession number, specify the accession number, the format, and the term FROM followed by a database number. For example, t 1234567/5 from 470.

FIGURE 1.2: Dialog Corporation's Classic interface

2.3.2 Menu-driven interfaces

Menu-driven interfaces resolve the problem associated with command-based interfaces by providing the user with a range of choices, thereby obviating the need to remember individual commands and command syntaxes (Harter 1986; Davies & Hepworth 1993).

In these interfaces, menus provide the user with a list of options and selection may be made either by typing a number or character representing the option, or simply pressing the *Enter* key when a particular option is highlighted. In graphical interfaces, selection is made by a graphic device such as a mouse or graphic tablet.

On the contrary, scrolling through a list of menus can be tedious since every option selected may lead to other options before arriving at the required option. As a result, much time is consumed scrolling through menus. An example of a menu-driven interface is provided in Figure 2.1.

Options in executing queries are provided. That is, the end-user may prefer searching information through the title, series name and the ISBN or ISSN.

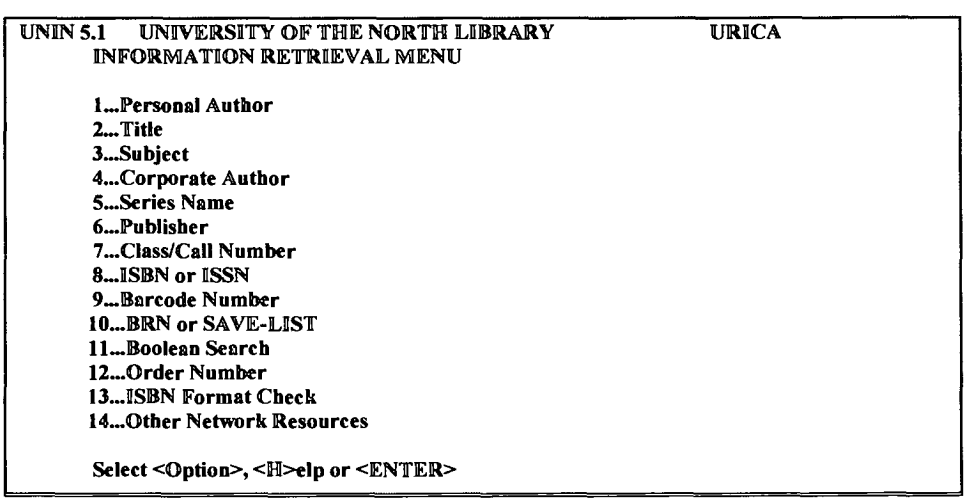

FIGURE 2.1: URICA's Acquisitions main menu

2.3.3 Direct manipulation interfaces

Direct manipulation interfaces are object-oriented interfaces in the sense that an object is chosen first, upon which an operation is performed (Ratnakumar & Haravu 1994:16). Direct manipulation interfaces are object-oriented in nature. Object orientation in the human-computer interface allows the use of real-world objects.

The underlying principle is that humans must fully comprehend the real-world object, or at least the mental model of it, before they can carry that knowledge across to the manipulation of the corresponding objects on the screen. Ratnakumar and Haravu (1994:16) base their argument on the assumption that *"...it is easier to see something and point to it than to remember and type it"*.

Direct manipulation interfaces minimise the user's effort to learn commands, to eliminate typographic errors and to keep the attention of the user on screen display. In general terms, icons allow humans to think in the familiar terms of the application domain rather than those of the medium of computation. As indicated in the preceding paragraph, it is quite comprehensible to refer to physical actions or icons rather than having to deal with complex syntax. The following advantages of direct manipulation interfaces, as seen by Shneiderman (1992:205), have been attributed to the fact that icons are preferred over complex syntax:

- Novices can learn basic functionality quickly through a demonstration by an experienced user, because the system lacks complex syntax. It should be noted that less syntax results in reduced error rates.
- Users can immediately see if their actions are furthering their goals, and, if the actions are counterproductive, they can simply change the direction of their activity.
- Users experience less anxiety because the system is comprehensible and because actions can be reversed so easily. Exploration is therefore encouraged.

- Users gain confidence and mastery because they are the initiators of action, they feel in control, and the system responses are predictable.

Direct manipulation interfaces have their own shortfalls. The disadvantages are as follows (Shneiderman 1992:204-205):

- Spatial or visual representations are not necessarily an improvement over text, because they may be too spread out, causing tedious scrolling on displays.
- Similarly, direct manipulation designs may consume valuable screen space and thus forcing valuable information offscreen, requiring scrolling or multiple actions (Shneiderman 1992:204).
- Direct manipulation interfaces have difficulty distinguishing the depiction of an individual element from a representation of a set or class of elements.
- Direct manipulation interfaces have problems with accuracy. The user has to control the action with precision.
- Another problem with direct manipulation interfaces is that they require substantial knowledge from the real world. In other words, in order for one to effectively derive meaning from direct manipulation interfaces, knowledge of the real world objects is required. For example, one must be able to attach a meaning on the *scissors* icon, which in object manipulation means "*cut*".

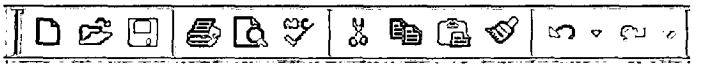

FIGURE 3.1: Object oriented interfaces (icons)

FIGURE 3.2: Object oriented interfaces (menu and icons)

Figures 3.1 and 3.2 are examples of direct manipulation systems. Direct manipulation interfaces require knowledge of the real world objects. For instance, the icons on Figure 3.1 cannot be comprehended unless one has knowledge on how these icons are applied to the world of interfaces. Figure 3.2 is a combination of both menu and graphics interfaces. This combination is user-friendlier in the sense that each icon has a description and the end-users need not struggle comprehending a particular icon.

Many companies have based their products on direct manipulation interfacing. Microsoft Office is one example of the products which uses icons.

2.4 SUMMARY

Human-computer interfacing is an area which keeps on developing. It is therefore necessary to continually evaluate products and compare their interfaces with the latest versions available. Research indicates that the manner in which an interface is designed has an impact on the particular producer's database. The implication is that an "unfriendly" interface will have a negative impact on the use of such database, and vice versa.

Chapter 2 was a general description of interface features. The focus was on how end-users information needs were addressed by typical features of computer interfaces. Chapter 3 will concentrate on front-ends to online systems.

CHAPTER 3

FRONT-ENDS TO ONLINE SYSTEMS

3.1 INTRODUCTION

Improvements in technology have led to software development and introduction of new software into online industry. As a result, front-ends were introduced as improvements on interface features. In other words, loopholes in previous interface features prompted the development of front-end software. This chapter focuses on the development of front-ends to online systems. Online services will be referred to in this study as conventional or traditional online services to differentiate them from other online services which have been developed more recently.

3.2 FRONT-END SOFTWARE DEVELOPMENT

Front-ends refer to software which mediate between the user and the online search service, in order to facilitate the process of information retrieval. Front-end software products have been developed as a result of attempts to simplify user interfaces. There has been terminological confusion between gateway software and front-end software. The following is a working distinction between front-end and gateway software as seen by various authors (Williams 1986; Van Brakel 1988; Efthimiadis 1990).

3.2.1 Front-ends

Front-ends mediate between the searcher and the databases provided through online services.

Efthimiadis (1990:221) maintains that front-ends have been developed with varying degree of sophistication reflecting the level of the complexity of the interaction implemented.

All interfaces have the same intention, namely to relieve the user of some of the difficulties and complexities in retrieving information from the system. In other words, front-ends enhance the features of a gateway by allowing steps such as pre-search editing, uploading of search formulation, and downloading to be simplified.

In some instances, the boundaries between front-ends and gateway software (which will be discussed in 3.2.2) categories can become blurred. It is because one search service can become a front-end while being a gateway at the same time.

There are additional advantages of front-ends over earlier interface development inherent in dumb terminals (Hawkins & Levy 1985; Efthimiadis 1990; Davies & Hepworth 1993):

- Front-ends can enable customised interfaces tailored to the needs of particular user groups.
- They allow maximum flexibility in the design of sophisticated interfaces and could include the incorporation of natural language, or artificial intelligence techniques.
- They can run under widely used software platforms such as the MacIntosh or Windows operating environments.
- They enable the integration of online searching with other software applications on the PC, such as downloading of retrieved information into word-processing, spreadsheet or database software for further processing.
- Communication software may be incorporated into the package thus making the initial communications set up relatively easy to perform.
- Front-ends may enable access to more than one host.
- Automatic dialling and automatic logon to a host computer were most felt benefits offered by front-ends.
- Help features.

These features can be in the form of separate screens available to the user on demand, instructions and prompts on the screen as the search progresses, or toll-free numbers to call for human assistance. The benefit of these features lies in the fact that many users have neither the time nor the inclination to read voluminous manuals. The following are the features applicable in front-end systems (Hawkins & Levy 1985; Efthimiadis 1990; Davies & Hepworth 1993):

- Pre-search editing and uploading.
  This feature allows the user to prepare the search strategy before connecting to the databank. Connect time and search costs are therefore saved through this feature.

- Database selection.
  The implication is that the system selects the database for the user by asking a series of questions, that is, selecting the database or providing a list of databases from which the user makes the final selection.

- Downloading.
  Many users wish to download search results and then reformat or manipulate them for further processing.

- Post-processing the search results.
  This feature goes further than downloading bibliographic citation in the sense that it involves performing statistical calculations, creating reports, or formatting bibliographies.

There are however, some disadvantages associated with front-ends. The obvious disadvantage of front-ends is in maintaining currency, both in terms of developments in the host service and in the data it presents.

For example, on mainframe-based systems it would be a relatively simple task and invisible to the user to correct minor software bugs, or implement system developments (Davies & Hepworth 1993:153). The start-up costs for front-ends are significantly greater than for traditional online services.

On the other hand, new versions of the software may be distributed to customers by diskettes, and this involves a continuing cost and the spread of computer virus.

3.2.2 Gateway software

Gateway software is an interface between the end-user and the host that provides auto-logon via a telecommunications network and possibly allowing uploading and downloading (Efthimiadis 1990:222).

A gateway performs the following functions:

- Dialling the telephone call.
- Selecting a communications network.
- Connecting to the host computer.
- Sending the user's login password (Efthimiadis 1990:223).

A gateway can be offered either at the searcher's end or as a service provided by an organisation which mediates between the searcher and the host computer, for example DialogLink.

According to Davies and Hepworth (1993:151), the earliest online interfaces evolved as communications software packages which were tailored to access particular online search service. Some of these packages facilitated searching, whereas others provided functions such as easy download of textual and numeric data.

In other online services, data was to be sent to the user's PC and all the processing was to be performed locally on the PC. A move towards easy accessibility prompted for the development of a user-friendly interface.

The following characteristics as described by Ratnakumar and Haravu (1994:20) are central to the development of Online Public Access Catalogs (OPACs), but most of them apply to online interfaces as well.

- The interface should be simple to use and understand.
- It should provide help messages, wherever appropriate.
- It should allow the use of Boolean operators (e.g. AND, OR, and NOT) during a search in a transparent manner as possible.
- The user should be allowed to select relevant records from a search set while browsing through the items in that set.
- The user should be able to perform the three traditional ways of searching a catalogue (by author, title and subject), and in addition users should have the option to do a simple two-parameter search.
- The user should be able to browse through indexes and selects relevant records based on terms in the indexes.
- The user should be able to save search sets to disc so that this information could be used at a later stage.

One important factor worth mentioning is that both the front-end software and the data it retrieves should be compatible with the user's technical environment, that is, whether the user is on UNIX, OS2, or Microsoft Windows platform.

3.3 SUMMARY

Through the developments in the online database industry, intelligent terminals, and later personal computers were introduced. Commands, menus, and graphical user interfaces (GUIs) were significant during this stage. These interfaces signalled the development of front-end software.

This chapter focused on the development of front-end software to online systems. Technological advances which led to the introduction of front-end interface features also formed part of this chapter. Chapter 4 will concentrate on Web-based interfaces.

CHAPTER 4

WEB-BASED INTERFACES AND THE IMPLICATIONS TO END-USERS

4.1 INTRODUCTION

Web-based interfaces refer to software that enables end-users to access information on the Web. Web interfacing was introduced as further improvement on end-user searching. Chapters 2 and 3 have been respectively focusing on the basic features of the typical interfaces to computer systems, and on front-end software development. Chapter 4 will be concentrating on Web based interfaces and how these interfaces differ from the traditional ones. It should be noted that the introduction of the Web has set a new trend for database producers and vendors to make their databases accessible through the Web. Web interfacing should not be seen as replacing traditional online systems, but rather complementing them.

One distinctive characteristic to be noted between the Web and conventional online systems is that in conventional online services, information is a product, whereas in the Web, information is sometimes a means to an end. O'Leary (1997) maintains that free online information on a Web site is a means to achieve other commercial ends, especially advertising or promoting a product or company.

Another difference is that the Web provides a common interface to all users, hypertext linking, and excellent image loading, none of which can be said of the conventional online platform. Web publishers and companies which develop sites on the Web can therefore perform tasks easily which conventional systems can do either with difficulty or not at all (O'Leary 1997).

The introduction of the Web has been received with mixed reactions. Some of the conventional online services were hesitant to adopt the Web medium because they saw the Web as a threat to their market due to free information and unlimited pricing available on the Web. Quint (1998) argues that instead of traditional online vendors taking advantage of the Web, they tend to fight it. With the Web becoming a major information tool, vendors started basing their products on the Web in order to remain in business.

Online services which could not make it to the Web were closing down. NewsNet is one example of such online services which were closed down. One reason given for the closing was that there is too much competition on the Web (Krumenaker 1997). In other words, most of the information provided at a fee by NewsNet was available for free on the Web. For instance, as Elston (quoted by Krumenaker 1997) puts it, out of one thousand (1000) titles on NewsNet, two hundred (200) of them available on the Web might be the most important titles. It is therefore meaningless for the company to continue charging for its services, whereas the very same services are free of charge on the Web. End-users will ultimately go for free of charge services and as a result the selling company will be out of business.

Web interfacing includes browsers (client programs) and search engines (search tools). What follows is the discussion of Web browsers, search engines and the current status in the movement towards Web interfacing by online services.

4.2 WEB BROWSERS

Web browsers are software programs that act as an interface between end-users and the Web. Browsers are regarded as "client" programs because they take commands from the end-user and then retrieve information and services by sending requests to "server" programs. Server programs include World-Wide Web (WWW), Gopher, and FTP.

Web browsers can be divided into two basic categories, namely text-mode and Graphical User Interfaces (GUIs). Text-mode browsers require less sophisticated computers or terminals and retrieve information faster than GUI browsers (Ellsworth & Ellsworth 1995:67). Lynx is one example of the text-based browsers.

The GUI browsers on the other hand, can perform many of the same tasks as text-based browsers, but they are accomplished largely through mouse point-and-click operations. Most GUI browsers have mouse scroll bars along the bottom and side of the window, allowing movement around the loaded page (Ellsworth & Ellsworth 1995:96). In addition to scroll bars, there are usually a variety of configuration options including colours and choice of which screen elements, such as toolbars and ribbons, are shown on the screen. Netscape Communicator is an example of GUI browsers.

The World-Wide Web is now the most popular server programme because it is capable of displaying text and graphics. As a result, Lemay (1995:5) defined the WWW as "a global, interactive, dynamic, cross-platform, distributed, graphical hypertext information system that runs over the Internet". This definition is the outcome of search results. The researcher included this definition under browsers to confirm that search tools (engines) search the Web through browsers. Lemay's definition is explained in the following paragraphs.

4.2.1 The Web is a hypertext information system

The World-Wide Web is a vast collection of text, graphics, sound, and video files (McGuire, Stilborne, McAdams and Hyatt 1997:8). The Web is a hypertext information system in the sense that the information selected links to another document which also contains links to other documents.

The implication is that the Web is neither rigid nor linear, instead documents can be selected at random. That selection may also have other links to it which also have links to other documents. It is possible to link back to the pages referred before.

4.2.2 The Web is graphical and easy to navigate

The fact that the Web combines text and graphics is an underlying principle which makes the Web very popular. As argued by Lemay (1995:5), prior to the Web, using the Internet involved simple text-only connections. Since the Web incorporates so much more than text, it is not enough to refer to the Web as "hypertext" system.

The Web can also be referred to as a hypermedia system since it includes other media, such as sound, graphics, and video.

4.2.3 The Web is cross-platformed

The Web can be accessed from whatever workstation the user is running on. In other words, as long as the user can access the Internet, access to the World-Wide Web is obvious irrespective of the type of the computer system. Since the Web started on UNIX systems, much of the use of the Web still takes place on UNIX systems (Lemay 1995:8). The implication is that since UNIX is still the dominant platform, most of the "newer advances in the Web technology are taking place on UNIX first".

4.2.4 The Web is distributed

The Web provides much information because that information is distributed globally across thousands of different sites, each of which contributes the space for the information it publishes (Lemay 1995:8). In other words, the consumers of information do not have to install or change disks in order to get whatever information they need.

4.2.5 The Web is dynamic

Since the Web is a conglomeration of information or full text pages, information can be updated at any time. New versions can be installed with little or no charges at all. New features are provided regularly which make the pages on the Web dynamic and up to date.

4.2.6 The Web can access many forms of Internet information

The World-Wide Web uses HyperText Transfer Protocol (HTTP), which is a protocol that allows for hypertext documents to be transferred quickly over the Net between the Web browsers and servers. The Web therefore provides a new system for publishing and distributing information.

In addition to the forms of Internet access, Powell (1994:59) indicates that the World-Wide Web supports more sources of networked and local electronic information than any other networked information retrieval tool.

4.2.7 The Web is interactive

The act of selecting a link and retrieving another screen of information defines the Web as an interactive system. The Web also enables people to design screens that look like forms and those who browse these screens can select from several choices filling in information in slots, or selecting a button to perform a particular operation.

From what has been discussed above, it can be deduced that the World-Wide Web is an amalgam of multimedia content, connected by hyperlinks and providing an easy, graphical interface for searching information. It should be noted that when the Web was introduced, navigation was basically through character-based browser (Lynx), until such time that GUI browsers were developed.

4.3 SEARCH ENGINES

Search engines are databases that search the Web to identify relevant files. The major area of database development on the Web has been in search engines. As a result, there are many search engines on the Web, most of them highly specialised and targeted at special subjects. Such search engines include Altavista, Excite, Infoseek, Hotbot, Lycos, and many more. Search engines are large automatically generated Web databases.

While the search engines represent the major mechanism to find Web-based information, Notess (1997) maintains that other strategies and directories provide quicker access to relevant sources than search engines.

Since search engines find and index as many pages as required, it is for this reason that they are most effective for searches with very unusual keywords, for combining keywords, for using advanced features such as field searching and limiting, and for finding pages within a Web site. The implication is that search engines are the most sophisticated tool to search the Web, and relatively easy to use.

Notess (1997) further indicates that search engines can be used for tracking down top-level pages for organisations when neither guessing nor the subject directories help, but as he noticed, they require a different approach.

All the search engines are aimed at information retrieval. There is no single search engine which can be regarded as self-sufficient. It is because what the searcher needs can be retrieved using different search engines depending on the type of information the researcher needs. For example, for mere browsing Yahoo can be preferred, whereas, for serious research Altavista or Infoseek can be recommended. It does not mean that Altavista or Infoseek are ends in themselves, as end-users could also use other search engines.

Search engines differ in terms of features and options they offer to end-users. Many of these search engines function reasonably well, although most have significant flaws (Notess 1997:58). It should be emphasised that a specific search engine produces better results if indexed manually as well. In this context, Notess (1997:58) maintains that a search engine that searches high quality sites in a specified subject area can be much more efficient and effective for a given discipline than a general Web-wide search engine. He argues that this approach can be combined with the best search engine features and thus produce search results of good quality.

It must be noted that search engines differ according to the:

- size of the index
- frequency of updating the index

- search option
- speed of returning a result set
- result set presentation
- relevancy of the items included in a result set, and
- overall ease of use.

4.4 CURRENT STATUS IN THE MOVEMENT TOWARDS WEB INTERFACING BY ONLINE SERVICES

The already existing need of friendly environments for effective information access has been further enforced by the increased growth of Web-based information. Since the introduction of the Web, there has been a drastic change in both the kind of people who access the information and the types of information itself. The quantity of information sources available on the Web is enormous, and it is required that the interaction mechanisms be friendly and easy to use. In relation to GUIs, there is a need for both effective visualisation of multimedia information and retrieval tools to be able to overcome the existing dichotomy between browsing and querying (Catarci 1996).

In most systems, the end-user has the duty of locating information by browsing multiple sources which could be complex and disorganised through too many links. Moreover, database information is traditionally structured and retrieved by constructing search strategies in a specific hosts' command language. In addition, information residing on the Web is retrieved by following hypertext links.

As indicated previously, the popularity of the Web prompted a number of prominent online vendors to incorporate this technology to enhance their online services. Eaton (1997:13) lists the following reasons as advantageous when online searching migrates towards Web technology:

- Despite its limitations, HyperText Markup Language (HTML) offers more rapid development opportunities than the complexities of creating large custom-built Windows programs.
- For end-users, HTML is a richer medium that plain American Standard Code for Information Interchange (ASCII) text and documents can be delivered directly to PCs in their original as-published format, namely as Acrobat Portable Document Format (PDF) files.
- For most information professionals, the Web browser has become the friendly environment for manipulating information, regardless of the platform being used.

Despite the benefits of the Web, there are also some disadvantages associated with the Web. Other disadvantages are as follows (Ochsner and Thomas 1996:478):

- It is difficult to find what is actually needed on the Web. The Web does not have any obvious boundaries to categorise information into relevant areas of interest. In other words, the user has to struggle locating a particular piece of information from the mass available. As indicated previously, the amount of information is so vast that scrolling through the data to measure its relevance is time consuming. On the other hand, overloaded telecommunication systems can further slow down the retrieval process.

- It is often difficult to measure authority and quality of the output.
  Information on the Web can be published by anyone and it can therefore seem authoritative. If sources are unknown, it is difficult to check the true value of the data provided.

- There are many search tools on the Web. There is a range of different Web crawlers and search engines available for use. All these search tools offer a variety of different search aids.

- Lack of standardisation of the output formats. It is usually difficult for the results to be presented in a manner that fits the users environment. This is mainly because information from the Web comes in a multitude of different formats.

On the question of searching in both the traditional online services and the Web, Van Brakel (1997:235) maintains that Web-based online visibility still needs extensive developing and testing when compared to the relevance and precision ratios of a commercial search. With the move towards push technology, it is hoped that Web-based interfacing will also be improved.

Despite the high recall and low precision ratios in the search process, the Web has solidified its position as the default platform for electronic information publishing and distribution. Search engines, subject guides, and other Net-based tools were thrust into the spotlight, though many were not ready to adopt the Web as a publishing medium.

On the other hand, traditional databases and some proprietary online services which did not take advantage of the Web were closed down and only a limited number made their way through the Web. The implication is that these traditional online services had to consolidate their databases to remain viable on the Web. Through consolidation within the database industry, some of the duplication among database offerings on different online hosts will be eliminated.

Database producers and vendors will be prompted to focus on the Web and on Web-oriented alliances and licensing agreements. University Microfilm International (UMI) ProQuest Direct and Dow Jones Interactive (now Factiva), for example, have shared content but for different markets. UMI is focusing on the academic, public, and school libraries, while Dow Jones gets the corporate and other special libraries market. This approach is directed to end-users.

UMI ProQuest Direct has added the full text of the Wall Street Journal all the way back to the 1860s, including Baron's and other Dow Jones (formerly known as Dow Jones NEWS/RETRIEVAL) content that had been exclusively available from Dow Jones Interactive (Tenopir 1997:37). Dow Jones is offering ABI/INFORM and the many full-text journals and magazines offered by UMI, including the image versions. Sharing of content is beneficial to both UMI and Dow Jones, and to the customers in particular. In other words, customers need not learn another system.

Since the purpose of this research was on Web-interfacing of commercial database vendors, examples of such vendors are provided below. The discussion also includes examples of major database producers.

4.4.1 Vendors on the Web interface platform

The introduction of the Web led most database vendors to opt for Web interfacing. One of the first vendors who opted for Web interfacing was The Dialog Corporation. Some vendors later followed and others are still experimenting with the Web. Most vendors seem to be committed to long-term development of their Web capabilities. According to Tenopir and Barry (1997:28), database vendors have already moved into testing their second-generation Web products, both to improve and update them.

The first generation of the database vendors' Web products was limited in terms of search capability and print options (Tenopir & Barry 1997:28). The new Web versions in the database market involve, for example, the adoption of sophisticated Java programming methods.

Since the database market has been intensified by the introduction of the Web, database vendors are now embarking on an alliance to collectively exploit the Web.

As maintained by Van Brakel (1997:235), it is not easy to distinguish if a specific product originated from a vendor or producer. It should be borne in mind that the Web is just another platform by which database companies can bring their products to end-users. It does not mean that traditional online vendors should get rid of their online and CD-ROM services, but to enhance these services through the Web platform.

Conventional online hosts should adopt the Web platform because it is a step in the progression of technology. Hosts which ignore the development lose out, for example, NewsNet. By adopting a Web platform, conventional online services will benefit as follows (Ochsner & Thomas 1996:481):

- The advanced search engines which conventional online services developed can be translated to the Web environment.
- Users will benefit from tools which perform complex searches automatically without the users having to develop their own search techniques.

The Web, for the most part, is an enhancement to the variety of platforms most companies already offer. The following are examples of vendors that opted for the Web-interfacing. Dialog is excluded as it is described in more details in the next chapter.

4.4.1.1 UMI's* ProQuest Direct

UMI's ProQuest Direct is an information system which combines the search facilities for its electronic sources and information delivery in a single, easy-to-use desktop package.

The competition within the database market has led to various database producers and vendors to form an alliance.

** UMI has been acquired by Bell & Howell Information and Learning*

On the other hand, some of the services have been bought by other bigger companies. An example of such services is The Dialog Corporation which is now trading under Thomson Corporation.

As indicated previously, UMI ProQuest Direct and Dow Jones Interactive have shared content but for different markets. On one hand, the interface in UMI provides access to databases such as those in academic, public, school libraries, and the full-text of Wall Street Journal. On the other hand, the interface in Dow Jones Interactive provides access to databases such as those which were not available from Dow Jones (that is, those formerly offered by UMI) including ABI/INFORM and other full-text journals and magazines.

4.4.1.2 Ovid Online

As indicated by Jacso (1996:25), Ovid Technologies, Inc. was the first online vendor that developed a fully functional Web interface that can manage search sets, display thesauri and search large databases such as MEDLINE and EMBASE with ease.

With the Ovid Web Gateway, an Ovid server's communications can be easily converted into the Web's page-based document language. The conversion results in any of the Ovid databases to be easily accessible through the Web browsers.

4.4.2 Database producers on the Web

Database producers took advantage of the Web by purchasing or developing a search engine and, thus making their databases directly accessible to information professionals and end-users. The direct accessibility of the databases from the producers to the end-users questions the potential role of database vendors.

It does not necessarily mean that database vendors will cease to function, but that their influential role in database market might be adversely affected. The following are database producers who took advantage of the Web:

4.4.2.1 Sociological Abstracts, Inc.

The Sociological Abstracts, Inc. has Web Search Service as its Web platform. Through the Web Search Service, access to various databases produced by Sociological Abstracts, Inc. is offered. Each database includes citations and abstracts of journal articles, books, book chapters, conference papers and citations to book reviews and dissertations (Kassel 1997).

According to Kassel (1997), the Web databases date back to 1986, whereas traditional online vendors provide backfiles from 1963.

4.4.2.2 Information Access Company (IAC)

Information Access company has produced InSite Pro as its Web platform. Kassel (1997) indicates that InSite Pro contains 2,500 journals from six of IAC's databases. Through InSite Pro, users are able to search the databases separately or conduct multi-file searches simultaneously.

The latest features of InSite Pro include searching by date range and two advanced search fields in Health and Wellness InSite. InSite Pro's PowerlinkSM technology provides fielded searching, refined search, and highlighted search terms with a user-friendly interface, which guide end-users in the search process.

In addition to the features mentioned above, InSite Pro also displays hyperlinked Universal Resource Locations (URLs) for the companies users are searching. In contrast to traditional online services, this Web product uses fixed-fee and subscription-based billing and no hourly charges and search fees are applicable (Kassel 1997).

Table 4.1 indicates some of the database producers who are already on the Web platform. Their Web platforms and URLs are also provided.

Database Producers	Producers Web Platform	Producers' URLs
CSA	CSA-IDS	www.csa.com
IAC	InSite	www.iac-insite.com
ISI	Web of Science	www.isinet.com
Sociological Abstracts, Inc.	Web Search Service	www.socabs.org
UMI	ProQuest Direct	www.umi.com

TABLE 4.1: Database producers and their databases on the Web platform

Regardless of the availability of the Web platform, online database producers continue to offer their databases through the major commercial vendors. This is in part, due to the fact that individual database publishers' Web services complement traditional search services. However, database vendors need to shape their products in line with users requirements in terms of speed, ease of use, and low charges.

With the development of technology, it is quite evident that the usage of the Web will continue to grow. Conventional online services are therefore aware that the Web is not simply a threat but a dissemination medium of enormous potential.

4.5 SUMMARY

Both database producers and vendors took advantage of the Web by producing their databases based on the Web standards. This is due to the fact that the Web seems to become the default platform for publishing information. Most companies developed Web-based interfaces to enhance their services. As highlighted earlier, the Web complements some of the platforms most companies were already offering.

The availability of free services on the Web led small companies to close down, whereas other companies decided to form an alliance. Some of the database producers started offering their databases directly to the information professionals and end-users. Though these database producers are in touch with the customer market, they still offer their databases through vendors.

Chapter 5 focuses on the comparison of interface features as outlined in chapters 2 to 4. Traditional and Web interface features will be closely looked at to determine disparities between them.

CHAPTER 5

CONVENTIONAL ONLINE AND WEB INTERFACING: COMPARATIVE ANALYSIS

5.1 INTRODUCTION

The study has thus far dealt with interface features in general. Chapter 3 focussed on front-ends to online systems, whereas Chapter 4 concentrated on Web interfacing. Various interface features were discussed in these chapters.

The purpose of this chapter is to investigate how the traditional features of online database vendors are being accommodated via a Web interface. In other words, the extent to which typical front-end features are being addressed through a specific system's Web interface. This chapter constitutes a comparative analysis of interface features discussed in Chapters 3 and 4.

The first part deals with the description of the methodology which was used to provide a comparison between front-ends and Web interface features. This discussion will conclude with recommendations regarding future interface development.

5.2 METHODOLOGY FOR COMPARING SEARCH RESULTS

In this research, a typical front-end and a Web-based online search service were compared to determine how a certain group of commands were executed.

DialogLink (DialogClassic) and DialogWeb were selected for the experimental work. DialogLink represented a typical front-end whereas DialogWeb was recently introduced as a Web-based service. The two approaches are both from The Dialog Corporation and have been selected due to the fact that this service has been the standard in online searching (Van Brakel 1997:238). A literature analysis has also been done in addition to actual searching.

The search area was limited to three searches so that different levels of the search process could be observed. The three search levels include:

- A basic single word search. For this purpose, a "basic single word search" is defined as a search which consists of a single word.
- An intermediate search using natural language terms, Boolean logic and truncation.
- A complex search comprising of natural language terms, Boolean logic, truncation, proximity and field limitation.

Searches were first conducted via DialogLink on DialogClassic to determine how interface features effectively interact with end-users. The same searches were also conducted on DialogWeb. DialogWeb required searchers to follow the "*Display Sets*" option for displaying the hit count for every term appearing in the search statement. This option proved to be a tedious procedure as compared to automatic hit count display in DialogClassic. The hit counts displayed on the search examples were, however, obtained from DialogClassic and DialogWeb. The searches were based on the ERIC database. The following is a discussion of the results:

a) A basic word search

A basic phrase search was based on the concept *"outcome-based education"*. The strategy was entered as follows:

```
Outcom? and base? and educ?
```

Processing of this query resulted in the following output:

DialogClassic

```
outcom      56
base        1009
educ        569
Total       294
```

DialogWeb

```
outcom      71
base        1239
educ        700
Total       1000
```

The total number of hits on DialogClassic and DialogWeb resulting from the search output is depicted in Table 5.1.

	Number of hits
DialogClassic	294
DialogWeb	1000

TABLE 5.1: Basic word search

The search output in Table 5.1 indicated that there were 294 records matching *outcome based education* on DialogClassic and 1000 records on DialogWeb. The terms in the search query were truncated to retrieve all the possible references. As a result, processing of the query resulted in fewer records in DialogClassic than in DialogWeb.

The researcher evaluated the first ten hits in both DialogClassic and DialogWeb to establish relevancy. It was found that most of the records within this set were relevant to the query, as depicted in Table 5.2.

Relevancy	
DialogClassic	90%
DialogWeb	90%

TABLE 5.2: Relevancy establishment in a basic search

Within the first ten records, seven records appeared to be duplicates. The finding is that in this basic search, DialogClassic and DialogWeb produce more or less the same results.

b) Intermediate search

"Distance education through the Web or Internet" represented an example of an intermediate search. The strategy was input as follows:


```
(distanc? or correspond?) and (educ? or teach?)and (web or
world(1w)wide(w)web or www or int?net)
```

The search output resulted in the following:

DialogClassic

```
distanc          39
correspond       1321
educ             569
teach            43
web              1541
www              1182
int?net          830
Total            213
```

DialogWeb

distanc	40
correspond	1431
educ	700
teach	43
web	1742
www	1392
int?net	880
Total	880

The total number of hits on both DialogClassic and DialogWeb resulting from the search output is provided in Table 5.3.

	Number of hits
DialogClassic	213
DialogWeb	880

TABLE 5.3: Intermediate search

DialogClassic resulted again in fewer hits than DialogWeb. The high quantity of hits produced does not necessarily imply that the relevancy of records produced will also be high. The implication is that even within the first ten records not all the references match the search query. The abstracts of the first ten records were evaluated and their relevancy to the query is depicted in Table 5.4.

Relevancy	
DialogClassic	70%
DialogWeb	80%

TABLE 5.4: Relevancy establishment in intermediate search

Only four records appeared to be duplicates in both DialogClassic and DialogWeb. The implication is that as the query becomes more narrow by means of the addition of more keywords, the number of duplicates is also reduced. For example, a search in "*distance education*" will retrieve more duplicates as compared to a search in "*distance education through the Web*". This is because the query "*distance education through the Web*" is more specific than just "*distance education*".

c) A complex search

At an advanced level, the following topic served as an example:
"*Outcome based education and teacher training or education"*

The search strategy was entered as follows:

```
outcome(w)based and (educ? or teach? or train?) and
(curricul? or learning(w)content)
```


Processing of the query resulted in the following hits:

DialogClassic

outcom	56
base	1009
educ	569
teach	43
train	770
curricul?	1821
Total	112

DialogWeb

outcom	71
base	1239
educ	700

```
teach       43
train       781
curricul?   1831
Total       621
```

Table 5.5 indicated the number of hits on both DialogClassic and DialogWeb.

	Number of hits
DialogClassic	112
DialogWeb	621

TABLE 5.5: Complex search

It can be seen from Table 5.5 that there were far fewer hits from DialogClassic than from DialogWeb. Thus, the differences in the number of hits could be attributed to the use of features or commands as discussed later in this chapter. As in the two search levels, the researcher also evaluated the first ten hits in both DialogClassic and DialogWeb and then established their relevancy. Few records were irrelevant to the search query as shown in Table 5.6. On the other hand, only three records appeared to be duplicates. This implies that in a complex search the query becomes more specific and thus, results in least duplication as compared to the other two searches.

Relevancy	
DialogClassic	70%
DialogWeb	80%

TABLE 5.6: Relevancy establishment in complex search

In general, all three searches from DialogClassic and DialogWeb indicated that when queries become more specific, they tend to have fewer duplicates.

Thus DialogWeb resulted in more relevant hits than DialogClassic. The search output indicated that both DialogClassic and DialogWeb accommodate the use of synonyms, for example "Web" and "Internet". When terms were truncated, more hits were produced. Apart from the empirical tests to indicate the differences in search results, the following features should also be discussed to establish similarities and differences in searching on DialogLink (DialogClassic) and DialogWeb.

5.3 GENERAL INTERFACE FEATURES

The following interface features are applicable during the above comparison:

5.3.1 Use of differemt techniques during an online search

Keywords were identified in the search queries used in the searches discussed in paragraph 5.2. In addition to keywords, truncation and Boolean operators such as AND, OR, NOT were also used during the searches. The implication is that the searcher was able to use a combination of terms and synonyms in every search query. For example, in searching for *"distance education through the Web"*, the user was able to use synonymous terms such as "*correspondence, teaching, WWW and Internet"*. All these options are supported in DialogWeb and DialogClassic. DialogWeb also provides a check box whereby the user marks the required searching parameter.

5.3.2 Database selection

It should be noted that DialogLink is software which gives access to DialogClassic. As a result, the actual searches were performed on DialogClassic. In DialogClassic, databases are selected manually, that is, by first filing, finding and consulting the correct database. These procedures require end-users to type in the correct commands.

On the other hand, DialogWeb offers the capability of linking directly to the required database. Therefore, it becomes easier for end-users to locate files from DialogWeb than in DialogClassic.

5.3.3 Automatic logon

An online system should provide users with automatic login facility which includes auto-dialing (Van Brakel 1987:118). The user should be able to store and send his/her identification codes and passwords to the vendor's database.

Van Brakel further maintains that *"...a program must be able to accommodate all the user ID's and passwords for the various systems the searcher is going to use"*. In other words, the system should also provide access to more than one host (that is, various databases from a particular vendor or publisher). Both DialogLink (DialogClassic) and DialogWeb give access to the user as soon as the user's details are authenticated by the system. Users' details are stored and identification codes are issued. In addition to the identification process, DialogLink (DialogClassic) and DialogWeb indicate to the searcher the status of the connection.

5.3.4 Uploading strategy

This feature allows for the offline development of a search strategy. The searcher can type and modify the search without actually going online. More offline functions such as offline storage and printing of results are available in DialogClassic than in DialogWeb. A type-ahead feature available in both systems enables users to type their search query ahead. This feature is depicted in Figure 5.1.

5.3.5 Searching and relevancy establishment

In general, the search output from DialogClassic resulted in fewer relevant hits than those from DialogWeb. This was proved on browsing through the first ten retrieved hits. The test was performed in all three search levels.

In addition to relevancy, DialogClassic and DialogWeb have a "*refine search*" feature which allowed the user to improve the search query. For example, the search on "*distance education*" retrieved many records, which were not easy to browse through. Using this "*refine search*" option in edit mode, it was possible to modify the query and thus retrieve results that closely match the query in the mediate and complex searches.

5.3.6 Storing

This feature allows the user to save the results for later use. Results can be more easily stored in DialogWeb than in DialogClassic. This is because in DialogClassic searchers have to pull down the file menu from the browser and thereafter be prompted for further instructions. In addition to commands, searchers need to remember the "Keep" command which allows for the creation of "Set O" which assists in the retrieval process. The session capture (a facility which enables retrieved results to be captured and consequently stored) available in DialogClassic enables the searcher to save the results directly on to the hard disk.

That is, the searcher has to click on the "*save*" icon or pull down the file menu from the browser, selects "*save*" and thereafter specifies the location of the file. Once stored, it is easier to retrieve results in DialogWeb than in DialogClassic.

5.3.7 Logoff

An online system should have a logoff feature which enables the searcher to quit from the system. It implies that the searcher has to end the session and thus logs off from the system. This feature is available via both DialogLink (DialogClassic) and DialogWeb.

5.3.8 Search hold

The search hold facility suspends the search and therefore reduces the costs involved in searching. DialogClassic suspends connect charges from the time the searcher clicks on a "*Help*" button to the time another command is executed. That is, DialogClassic technically issues a "*search hold*" command which turns off connect time charges. In DialogWeb, all the links to other functions can be read without costs. It is only when the system is processing queries that the searcher has to pay.

5.3.9 Session capture or log-to-disk

This feature allows retrieved results to be captured and stored on the workstation's disk. Session capture is applicable only in DialogClassic. In other words, DialogClassic requires that in capturing the results searchers need to remember the "*Keep*" command which allows for the creation of one "*Set O*" of all the required records. In DialogWeb, the searcher has to remember to pull down the file menu from the browser, select "*Save as*" then be sure to save the file with "*.txt*" and not with "*.html*" extension.

5.3.10 Postings features

DialogClassic has posting features which allow searchers to use the "*Set commands*". It is also possible to use several commands in one search. DialogWeb does not make use of posting features. That is, a command has to be fully executed before the next one can be entered.

5.3.11 Capturing of costs

This feature allows for costs to be accumulated for the same searcher during the same search session. As a result, the searcher is able to see how much he/she has spent within a given time. The capturing of costs feature is fully supported in DialogClassic and DialogWeb. That is, the next session by the same searcher will start with new costs.

5.3.12 Help features

These features can be in the form of separate screens available to the searcher, instructions and prompts on the screen as the search progresses. Both DialogClassic and DialogWeb have help features which guide the user throughout the search process. In DialogWeb, help features are referred to as guided search features.

5.3.13 Combination of commands in a search (stacking)

In DialogClassic, the user can input a combination of several commands within a search query. This is one of the features which helps users to modify their queries. Stacking of commands is not supported in DialogWeb. That is, commands have to be typed in one at a time.

5.3.14 Offline facilities

Offline facilities allow for the creation and storage of new search strategies, editing downloaded search results, offline printing and accounting functions. All these functions are fully supported in DialogClassic. However, DialogWeb possesses few of these facilities, such as the *Type-ahead buffer*. Figure 5.1 is a *Type-ahead buffer* which represents one of the offline facilities.

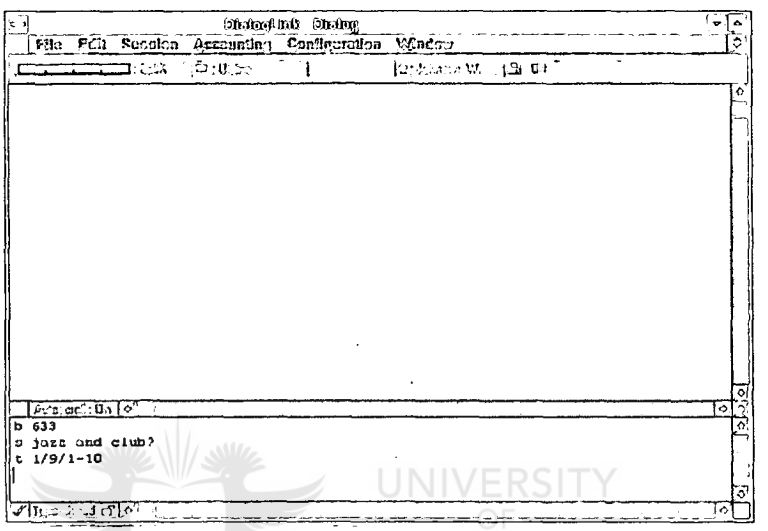

FIGURE 5.1: A Type-ahead buffer

5.3.15 Ranking

This feature allows for results to be ranked according to their relevancy to the query. The best matching references are usually ranked first. Both DialogClassic and DialogWeb support this feature.

5.3.16 OneSearch

The OneSearch feature allows for the multiple searching of databases in a single mode. This approach relieves the searcher of the difficulties associated with first closing the current database and then logging on to another. DialogClassic and DialogWeb have each one interface which gives access to more than 450 databases.

5.3.17 Check box

The check box allows the user to click next to the databases or search options he/she chooses to search. This feature is applicable in DialogWeb only. DialogClassic does not support this feature, but instead databases are selected by using the "*Set commands*". Figure 5.2 is an example of a check box.

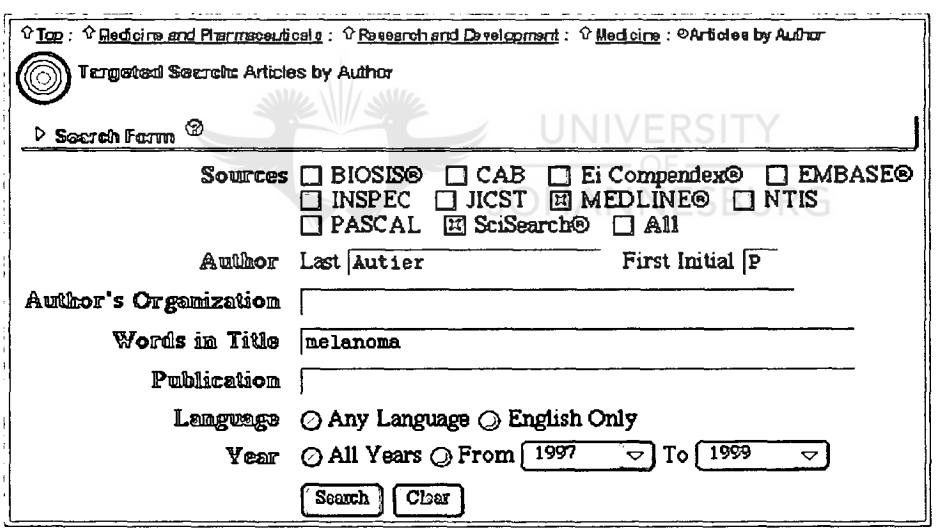

FIGURE 5.2: Check box

5.3.18 Direct link to *Blue Sheets*

With this feature, the user is able to click directly to the *Blue Sheets* of his/her choice.

DialogWeb offers the capability of linking directly to any *Blue Sheet* with its context-sensitive hyperlinks, as in Figure 5.3. DialogClassic also offers the capability of linking directly to *Blue Sheets*. The implication is that The Dialog Corporation has made it possible for users to link directly to *Blue Sheets*, under "*products*" from its website.

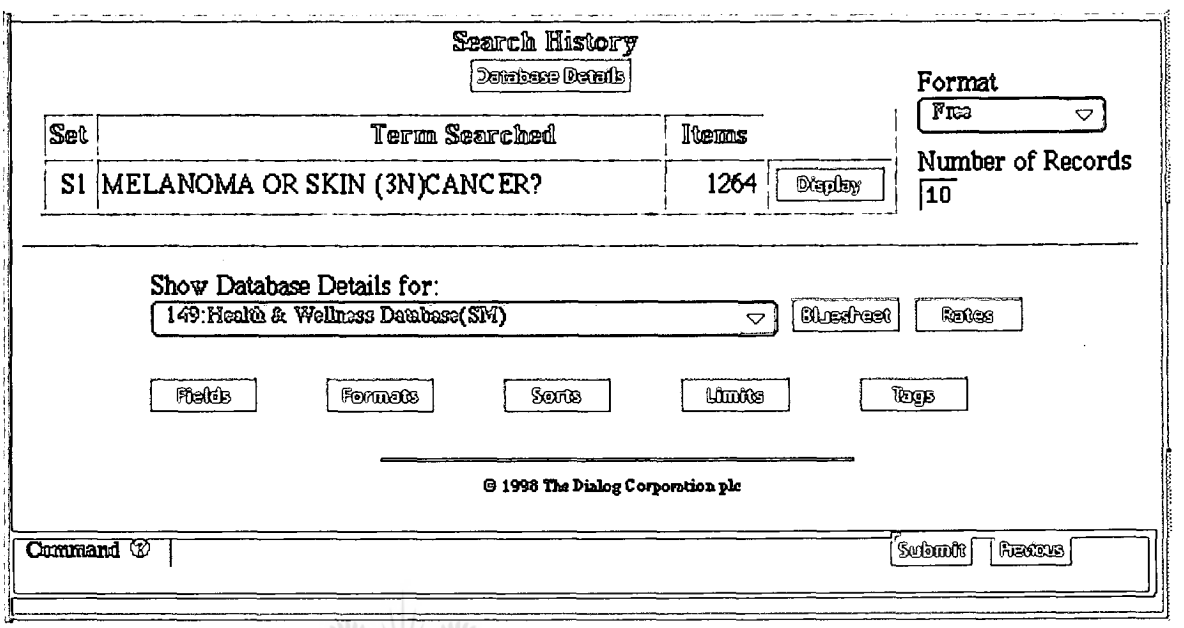

FIGURE 5.3: Link to *Blue Sheets*

5.3.19 Duplicate removal

This feature removes the documents which appear to be duplicates in the search result.
In DialogClassic duplicates are removed by selecting the "*Set commands*", whereas in DialogWeb the searcher has to click "*remove duplicates*" feature on the check box. That is, both DialogClassic and DialogWeb support this feature.

5.4 SUMMARY

The purpose of this chapter was to discuss the methodology used and the results to determine how a certain group of commands perform in both a Web-based online search service and a traditional front-end. A description of the search processes at the basic, mediate and complex search levels were included in the methodology phase.

The data analysis phase on the other hand, was covered by discussing the output of the search processes. The results were also augmented by establishing relevancy of the search output in all the three search levels. These discussions provided a framework for the comparison of the online interface features of the two Dialog interfaces.

This framework outlined interface features which were supported in both DialogClassic and DialogWeb. On the other hand, certain interface features which were not supported in DialogClassic or DialogWeb were also outlined. The results of such comparison of these interface features could form a basis for future interface development. From the empirical searches as well as the comparisons in the discussions under 5.2 and 5.3 respectively, a summary of results could be created (see Table 5.7).

Features	DialogClassic	DialogWeb
Use of different techniques	X	X
Database selection	X	X
Automatic logon	X	X
Uploading strategy	X	
Searching and relevancy establishment	X	X
Storing	X	X
Logoff	X	X
Search hold	X	X
Session capture	X	
Postings	X	
Capturing of costs	X	X
Help features	X	X
Stacking	X	
Offline facilities	X	X
Ranking	X	X
One search	X	X
Check box		X
Link to Blue Sheets	X	X
Duplicate removal	X	X

TABLE 5.7: Summary of features available in DialogClassic and DialogWeb

It can be seen from Table 5.7 that there were only four areas in DialogWeb which were not available in DialogClassic. Although these features were not yet supported by DialogWeb, it seems possible that they could be added in the near future. Therefore, searchers should search on both DialogClassic and DialogWeb or could also use DialogClassic on the Web for effective results.

In conclusion, the resulting comparison of interface features which were summarised in Table 5.7 could form a basis for future interface development.

CHAPTER 6

SUMMARY AND CONCLUSION

6.1 SUMMARY

The introduction of the Web not only provided new facilities to online searching, but has also enhanced the quality of an interface to online databases. In this context, the purpose of this research was to investigate the advantages and disadvantages of Web-based access to commercial database vendors. It has been argued that the introduction of the Web has brought some new interface features into the online industry. It was found that some typical front-end features were incorporated during the development of Web interfacing, whereas other features were ignored. On the other hand, some features were improved for example, linking to database information files such as *Blue Sheets* in DialogWeb.

This investigation was influenced by the fact that the Web is changing the impact of the online industry and is already being accepted as a vital information tool throughout the world. As a result, database producers and database vendors are increasingly adopting the Web as a digital information publication medium.

The focus of the research was to identify the typical interface features for end-user information seeking. The first part of the research was concerned with interface transformation, from the conventional online services to those accessible via the Web. Conventional online systems were previously dominated by the use of commands and menus. The new Web interfaces also use a combination of these features in addition to new features outlined in this report (see paragraph 5.3).

A general description of interface features was included in the first part of the research. As a result, DialogClassic via DialogLink and DialogWeb were selected and compared by the researcher to demonstrate different facets of online interfacing. Examples illustrating various interface features of a typical front-end and a Web-based service have been provided.

The development of front-end software for online systems was also discussed. Commands, menus and graphical user interfaces (GUIs) were at the core of front-end software development. These developments marked the improvement of interface features in conventional online systems.

The introduction of the Web brought further improvements in the interface area. Web interfacing made it possible for database producers and database vendors to make their databases more accessible and easier to use for end-user searching. Consequently, examples of database producers and database vendors who took advantage of the Web have also been outlined.

The report also concentrates on comparing typical front-end and Web interface features. Interface features on both DialogClassic and DialogWeb have been evaluated. DialogClassic served as an example of a command-driven online service, whereas DialogWeb represented Web interfacing. The researcher selected DialogLink (DialogClassic) and DialogWeb due to the fact that The Dialog Corporation has been the standard in the online industry for many years.

Additional features such as uploading strategy, session capture, postings, stacking and check box available in either DialogClassic or DialogWeb have also been discussed. These features proved to be basic to online searching regardless of whether they are available on DialogClassic or DialogWeb. Features such as database selection, storing, capturing of costs, link to Blue Sheets and duplicate removal are available in both DialogClassic and DialogWeb, but they are applied differently during online searching.

These features perform the same tasks although their approaches to online searching are different.

Lastly, a framework of interface features which could form a basis for future interface development was developed by this research project.

It was found that interface features such as uploading strategy, session capture, postings and stacking were not available in DialogWeb whereas, DialogClassic does not have check box features. The popularity of Web interfacing would prove DialogWeb the more used online search service than DialogClassic. It is therefore necessary that the inclusion of interface features which are absent in DialogWeb should be considered in an effort to improve the search results obtained by online searching techniques.

6.2 CONCLUSION

An empirical study was launched to determine how a certain group of similar commands and different levels of search queries were executed via both a traditional front-end and a Web-based online search service. The results of the comparison indicated that it was difficult for end-users to conduct searches on command-driven online services. Despite the high costs of conducting searches, serious end-users should be able to use search commands to enable them to get better results.

The introduction of Web interfacing brought improvements in online searching and consequently reduced the difficulty experienced in conducting searches. There were, however, disadvantages associated with the Web. One of the shortcomings was that some of the useful front-end features were not incorporated into Web interfacing yet. These features included uploading strategy, session capture, postings and stacking.

The fact that these features in the Web interface were excluded, might prompt users to conduct their searches on conventional online systems as well, to improve search results. Because of this, conventional online and Web interfaced systems complement each other.

Finally, it is recommended that the typical interface features which were excluded during the development of the Web and all the newly introduced features should be incorporated into the next generation of user interfacing. These would include features such as uploading strategy, session capture, postings, stacking and check box available in conventional online systems as well as Web-based services.

6.3 FURTHER RESEARCH

Web-based searching for information brought improvements in online searching, despite the disadvantages associated with it. Further research is needed to investigate the implications and applications of the suggested interface features. This was proved by the findings of this research, which were based on actual searching on DialogClassic and DialogWeb. The following are suggested for further research topics:

- A study investigating the implications and applications of the suggested interface features. The implications and applications of these features would determine their role in online searching.

- A study investigating the possibilities of further improving current online interface features. Such a study should focus on improving the use of these features.

BIBLIOGRAPHY

Arnold, S.E. 1987. End-users: dreams or dollars. *Online,* 11(1):71-81.

Ashton, J. 1998. Development of the British Library's OPAC 97: the value of a user-centred approach. *Program,* 32(1):1-24.

Barker, A.L. 1995. Non-Boolean searching on commercial online systems: optimising use of Dialog TARGET and ESA/IRS QUESTQUORUM. *Proceedings of the 19th International Online Information Meeting,* London, 5-9 December 1995. London : Online Information, 71-78.

Bates, M.E. 1997. The Internet: part of a professional searcher's toolkit. *Online,* 21(1):47-50.

Blake, R. & Wells, J. 1996. War of the Web. *Inform,* 10(5):20-28.

Borah, E.G. 1997. A virtual "hello": a Web-based orientation to the library. *Journal of Business and Financial Librarianship,* 3(2):3-18.

Bruce, H. 1998. User satisfaction with information seeking on the Internet. *Journal of the American Society for Information Science,* 49(6):541-556.

Catarci, T. 1996. Databases and the Web: new requirements for an easy access. [*Online*]. Available WWW: *http://www.acm.org/surveys/1996/CatarciDatabases.htm.*

Cox, K. & Walker, D. 1993. *User interface design.* Hertfordshire : Prentice-Hall.

Dalton, G.M.E. 1991. *User satisfaction as a measure of subject reference performance in an academic library.* (MLS thesis). Pretoria: University of South Africa (unpublished).

Davies, A. & Hepworth, M. 1993. PC-based interfaces to online information services and the development of FT Profile's freeway. *Program,* 27(2):149-163.

Dole, W.V. 1987. Libraries, publishers and vendors : looking for Mr Goodbuy. *Library Acquisitions : practice and theory*, 11 : 125-134.

Eaton, J. 1997. The future of Web services: changes in the use of the Web have been a mixed blessing. *Information World Review,* 125:13.

Efthimiadis, E.N. 1990. Online searching aids: a review of front-ends, gateways and other interfaces. *Journal of Documentation,* 46(3):218-262.

Ellsworth, J.H. & Ellsworth, M.V. 1995. *Marketing on the Internet: multimedia strategies for the World Wide Web.* New York : John Wiley.

Estabrook, L. 1981. Productivity, profits, and libraries. *Library Journal,* 106(13):1377-1379.

Feldman, S. 1998. Web search services in 1998: trends and challenges. [*Seacher*]. Available WWW: http://www.infotoday.com/searcher/jun/story2.htm.

Furner, J. 1997. IR on the Web: an overview. *Vine,* 104:3-13.

Galt, J.S. 1997. Does the future of the online industry lie in individually customised services?: how online hosts are adapting to the future. *Proceedings of the 21st*

International Online Information Meeting, London, 9-11 December 1997. London : Online Information, 13-18.

Halperin, M. & Pagell, R.A. 1983. Searchers' perceptions of online database vendors. *Special Libraries,* 74(1):119-126.

Harris, R. 1988. The database industry: looking into the future. *Database,* 11(5):42-46.

Harter, S.P. 1986. *Online information retrieval: concepts, principles, and techniques.* Orlando : Harcourt Brace Jovanovich.

Hawkins, D.T. & Levy, L.R. 1985. Front end software for online database searching, part 1: definitions, system features, and evaluation. *Online*, 9(6):30-36.

Head, A.J. 1997. A question of interface design: how do online service GUIs measure up? *Online,* 21(3):20-29.

Kassel, A. 1997. Here they come!: database producers on the Web. [*Searcher*]. Available WWW: *http://www.infotoday.com/searcher/jul/story2.htm.*

Khan, K., & Locatic, C. 1998. Searching through cyberspace: the effects of link display and link density on information retrieval from hypertext on the World-Wide Web. *Journal of the American Society for Information Science,* 49(2):176-182.

Kleivane, T. 1987. The emergence of smart gateways: the implication for the online industry. *Proceedings of the 11th International Online Information Meeting*, London, 8-10 December 1987, Online Information, London, 157-166.

Klingener, A. 1997. DIALOG Web: finally it's arrived. *Online,* 21(5):18-25.

Klopper, S.M. 1997. Testing, stretching, pushing, and pulling the DIALOG Web. *Online,* 21(5):26-33.

Krumenaker, L. 1997. Vendors past: Andrew Elston on NewsNet, "the canary in the mine?". [*Searcher*]. Available WWW: *http://www.infotoday.com/searcher/nov/story2.htm.*

Kwan, J., Watson, J.L. & Deeney, K. 1987. DIALOG medical connection: an evaluation. *Online,* 11(6):32-38.

Laurel, B. (Ed) 1990. *The art of human-computer interface design.* Reading, Massachusetts : Addison-Wesley.

Lemay, L. 1995. *Teach yourself Web publishing with HTML in a week.* Indianapolis, Indiana : Sams.

Letshela, P.Z. 1995. *End-user training of post-graduate students in the use of CD-ROM databases with special reference to the University of the North.* (MLS thesis). Johannesburg: Rand Afrikaans University (unpublished).

Lexis-Nexis expands services. 1998. [*Information Today*]. Available WWW: *http://www.infotoday.com/it/feb/news/htm.*

Livonen, M. & Sonnenwald, D.H. 1998. From translation to navigation of different discourses: a model of search term selection during the pre-online stage of the search process. *Journal of the American Society for Information Science,* 49(4):312-326.

McGuire, M. et al. 1997. *The Internet handbook for writers, researchers and journalists.* Ontario : Trifolium.

Nicholas, D. et al. 1987. Online: views on costs and cost-effectiveness. *Journal of Information Science*, 13(2):109-115.

Notess, G.R. 1995. Comparing commercial WWW browsers. *Online,* 19(3):43-49.

Notess, G.R. 1997. Measuring the size of Internet databases. *Database, 20(5)*:69-72.

Notess, G.R. 1997. Internet search techniques and strategies. [*Online*]. Available WWW: *http://www.onlineinc.com/onlinemag/JulOL97/net7.html*

Ochsner, H. & Thomas, K. 1996. The Internet: the demise of traditional online. *Celebrating 20 years of Online Information*, London, 3-5 December 1996. London : Online Information, 477-482.

O'Leary, M. 1988. EasyNet revisited: pushing the online frontier. *Online,* 2(5):22-30.

O'Leary, M. 1997. Wall Street Research Net embodies Web information model. [*Online*]. Available WWW: *http://w3.searchbank.com/infotrac/session/133/998/16145865w3/17!mltview&amk.*

Oxborrow, E. 1986. *Databases and database systems: concepts and issues.* Sweden : Chartwell-Bratt.

Powell, J. 1994. Adventures with World-Wide Web : creating a hypertext library information system. *Database*, 17(1):59-136.

Quint, B. 1998. It's a vision thing. [*Searcher*]. Available WWW: *http://www.infotoday.com/searcher/apr/voice.htm.*

Ratnakumar, P., & Haravu, L.J. 1994. Design and development of a user interface for the library's database at ICRISAT. *Program,* 28(1):15-27.

Rehkop, B.L. 1994. Cypress: a GUI interface to Dow Jones News/Retrieval. *Online,* 18(1):72-75.

Ross, C. 1984. Lessons of the pilots. *Bulletin (Office automation pilot projects),* 1(3):1-11.

Schoch-Gruebler, U. 1995. Future user environment: a view from industry. *Information Services and Use,* 15:31-37.

Sebastiani, F. 1998. Trends in...a critical review on the role of logic in information retrieval. *Information Processing and Management,* 34(1):1-18.

Shneiderman, B. 1992. *Designing the user interface: strategies for effective human-computer interaction,* 2nd ed. Edinburgh : Addison-Wesley.

Steele, I. & Tseng, G. 1992. End user training for CD-ROM medline: a survey of U.K. medical school libraries. *Program,* 26(1):55-61.

Tenopir, C. 1998. Will online vendors survive? *Library Journal,* 123(2):35-36.

Tenopir, C. 1997. Reading vendor literature. *Library Journal,* 122(10):35-37.

Tenopir, C. & Barry, J. 1997. The data dealers. *Library Journal,* 122(9):28-36.

Van Brakel, P.A. 1987. Criteria for the evaluation of database access software. *Proceedings of the 11th International Online Information Meeting,* London, 8-10 December 1987. London : Online Information, 111-122.

Van Brakel, P.A. 1988. Evaluating an intelligent gateway: a methodology. *South African Journal of Library and Information Science,* 56(4):277-290.

Van Brakel, P.A. 1997. Online database vendors: will they transform to pull technology? *South African Journal of Library and Information Science,* 65(4):234-242.

Venditto, G. 1996. IW lab tests four brand new Internet connection and navigation tools. *Internet World,* 7(1):53-58.

White, M. 1991. Doing business in Europe: the information industry. *Aslib Proceedings,* 43(2/3):43-53.

Williams, M.E. 1986. Transparent information systems through gateways, front-ends, intermediaries, and interfaces. *Journal of the American Society for information Science,* 37(4):204-214.