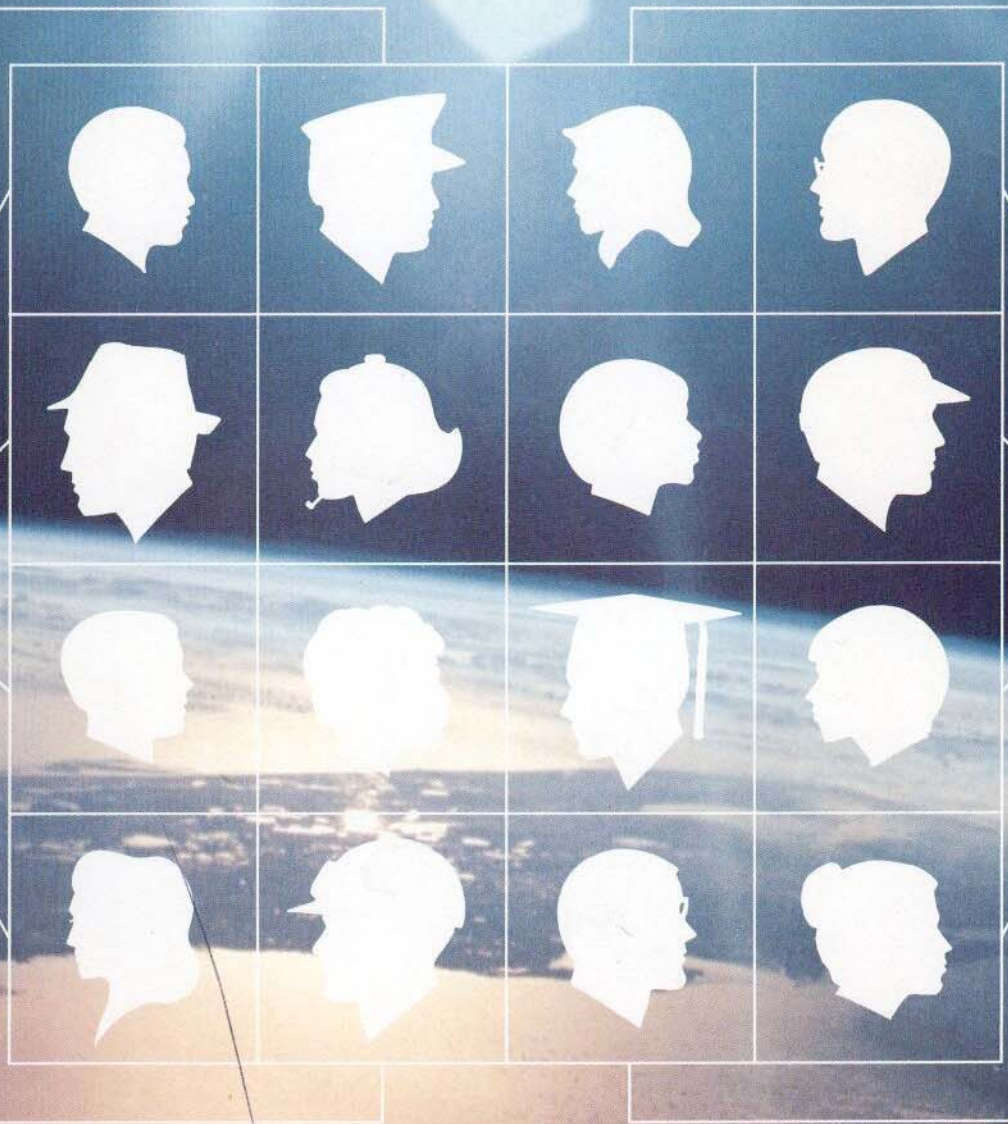


LITWIN

# Tutorial: End User Facilities in the 1980's

James A. Larson



IEEE CATALOG NO. EHO 198-2  
LIBRARY OF CONGRESS CATALOG CARD NO. 82-83256  
ORDER NO. 449

IEEE  
**COMPUTER  
SOCIETY  
PRESS** 

 THE INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, INC.

## MESSIDOR SYSTEM

Catherine MOULINOX (1)  
Jean-Claude FAURE (1)  
Witold LITWIN (2)

### Abstract

MESSIDOR is an interactive information retrieval system. It differs from current systems in that it allows the simultaneous search of several bibliographic databases. The databases may be on different sites and may use different query languages (MISTRAL, QUEST, ...). These local languages are invisible to the users of MESSIDOR. They are all translated to a single language. We describe MESSIDOR's goals, system architecture, user language and some details of the implementation. The system is implemented on a MICRAL 80-30 micro-computer.

### 1- INTRODUCTION

A bibliographic database (BD) is a set of bibliographic references. The recorded information generally includes the authors and the title of the document, the publishing date, details of its source, index terms and abstract. The documents may be serial articles, books, reports, conference announcements, etc.

BDs are now implemented on many computers. A user can access many of them with his terminal through telecommunication lines or through computer networks [ZIM 79] (EURONET, THYMESHARE, etc...). For instance a user of EURONET may actually access several hundreds of BDs [1]

A BD usually contains information about only one kind of document. For instance one BD may deal with books while another deals with technical reports. However several BDs are relevant to the same domain. For instance on EURONET there are 18 BDs relevant to computer science.

- .....
- (1) STERIA, 26 av. de l'Europe, 78140  
Vélizy, FRANCE
  - (2) INRIA, BP 105, 78153 Le Chesnay  
Cedex, FRANCE

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

It then follows that typically a user has to access several BDs. On the one hand, he must deal with various connection procedures.

On the other hand he must frequently deal with different languages, since one BD may be provided with MISTRAL while another one is provided with QUEST. Finally, he must switch from one BD to another, come back to the first one, etc. This renders an exhaustive retrieval very annoying and frequently, in fact, impossible in practice.

In order to make the task of a user easier, the methodology of databases has introduced the concept of a distributed database (DD). One may consider that a set of some databases is a DD when it has all the properties of a database [LIT 80]. A user of a DD may therefore, for instance, issue one query which is "automatically" processed against all databases (Local databases (LDs)) composing the DD. In particular, he formulates his queries in one language, even if the LDs are provided with different ones. The design of various kinds of DDs is the object of the French pilote project called SIRIUS [LEB 81].

A distributed bibliographic database (DBD) is a DD whose elements are bibliographic databases. For instance some BDs all relevant to computer science may lead to a DBD. A user of a DBD should be free of the annoying problems which face a user of the same BDs when he must deal with each of them separately. The goal of the MESSIDOR system is to provide for its users a DBD.

In what follows we describe the MESSIDOR system. Section 2 presents the general characteristics of MESSIDOR. Section 3 discusses the MESSIDOR language. Section 4 concludes the paper.

### 2 - MESSIDOR SYSTEM

#### 2.1. Overview

MESSIDOR is an interactive system. It allows the user to search simultaneously several bibliographic databases. These databases which constitute the LDs of the MESSIDOR DBD may be located on the same computer or on different geographically distant computers. LDs may be searched by various query languages : MISTRAL, QUEST, etc. However, the user converses with the system MESSIDOR using a single query language. This language, which we call the MESSIDOR language, is issued from the Standard Command Set [NEG 77]. Command Set is recommended by European Communities as the standard language for information



retrieval systems connected to EURONET.

MESSIDOR automatically performs the connection dialogue to LDs. A query formulated in the MESSIDOR language is automatically translated and sent to each LD. Then, MESSIDOR displays the set of the bibliographic references retrieved from all LDs. The user may interactively modify his query, receive new references, etc. At the end of the search, the user may edit complete information on every document selected from each LD using a single command. This makes retrieval search on several bibliographic databases possible and easy. Actually MESSIDOR is the only system known to the authors able to offer such facilities.

MESSIDOR is implemented on a micro computer MICRAL 80-30. This feature, awarding a low cost (approximately 60,000 french francs (\$ 12,000)), should allow libraries and information centers to easily purchase it. The normal dial-telephone and an acoustic 300 Bd modem are used to connect the system to telecommunication networks such as THYMSHARE, TRANSPAC, EURONET, etc. (see fig. 1).

The telecommunication protocol of MESSIDOR permits the simultaneous connection of several computers through different networks. Each connected computer sees MESSIDOR as an ordinary terminal (see fig. 2).

The MESSIDOR system is composed of three sub-systems :

- (1) The coordinator which manages the user-dialogue and the other sub-systems (2) and (3).
- (2) The interface with the databases, i.e. translators into LD languages (MISTRAL, QUEST, etc.)
- (3) The telecommunication driver which is the interface between MESSIDOR and communication networks.

The sub-systems (1) and (2) are written in the BASIC language. The sub-system (3) is written in Z80 assembler. More details on MESSIDOR software are given in [MOU 81]. [FAU 80].

## 2.2. Work with MESSIDOR

MESSIDOR supplies to the user a set of bibliographic databases composing the DBD. These databases are selected by the information center which owns MESSIDOR according to the needs of the center.

A user of MESSIDOR works with the system typically as follows :

- At the beginning of each search session, MESSIDOR shows all the databases composing the DBD, i.e. all databases selected by the information center. The user then indicates the LDs which he knows are pertinent to the search ; this facility avoids paying for access to LDs which the user knows to be useless for this search. For each selected LD MESSIDOR then generates the connection procedure and informs the user of their result (success or failure).

-Once all LDs are on line, the user does his search. The search is specified through MESSIDOR commands. The commands are processed as follows :

- (1) Each command is translated into every query language of the connected LDs.

(2) The translated commands are sent to the system managing the LDs.

(3) The results are received and stored

(4) The results are reformatted in order to become the DBD result and are displayed.

(5) The user is then allowed to issue a new command.

- At the end of the search, the user issues the command asking for the DBD disconnection. MESSIDOR then performs all the corresponding disconnection procedures.

## MESSIDOR LANGUAGE

### Main features

The MESSIDOR language guides the user in his interaction. MESSIDOR presents to the user menus or multiple choice questions (see fig. 3). Obvious error messages are displayed, when needed, so it is possible to make immediate corrections.

The displayed results clearly show how each connected LD contributes to the global result so that the user can know the usefulness of each LD. He can ask for the disconnection of LD(s) with a low contribution. The following searches will be more rapid and their cost will be lower.

### Example

Suppose the user is searching a DBD made of the three databases NTIS, INSPEC and PASCAL. He wants to know how many documents are indexed by the term : ANALYSER. After translating the command, searching each database, and reading their different results, MESSIDOR displays the global result :

2760	DOCUMENT(S)	REPARTI(S)	COMME	SUIT :
	NTIS	51		
	INSPEC	2708		
	PASCAL	1		

The result means that NTIS has 51 such documents, INSPEC has 2708 such documents and PASCAL has only 1 such document. So if for every query PASCAL has a very small contribution, the user can ask for its disconnection. The search will continue with just NTIS and INSPEC.

The user selects a set of documents by means of an appropriate command at each search step. Details of these documents can be either viewed or printed or both thanks to a special command. The user obtains all documents details for every online databases.

### 3.2. - Commands

#### 3.2.1. BASE command

The first thing the user must do after loading

the MESSIDOR system and typing his password is to choose the LDs for which the search is to be performed. All LDs of the DBD may be displayed using the command named BASE.

Example :

?? BASE

This command may lead to the following result :

BASE	SITE	LOGICIEL	RESEAU
1. BASE002	PROG. INSTITUTE	MISTRAL	CYCLADES
2. BASE002	PROG. INSTITUTE	MISTRAL	CYCLADES
3. NTIS	ESA	QUEST	THYMSHARE
4. INSPEC	ESA	QUEST	THYMSHARE
5. PASCAL	ESA	QUEST	THYMSHARE

MESSIDOR gives the LD names preceded by a code. Further, it gives for the user's information where the LDs are located (SITE), their data base management system (LOGICIEL), and the network through which they can be accessed. The user chooses LDs for the search through their codes. This is performed using the command :

?? BASE + 3, + 4, + 5

### 3.2.2. - LIST command

This command allows the user to look for search terms. A search term can be a key word (KW) or an author's name (AU). The search terms are sorted into a dictionary in each database. Search terms can be looked up in the dictionary by using the LIST command. This command should be followed by a particular search term.

Example :

?? LIST KW = ANALYSE

The result of this command is a display of parts of the dictionaries showing a list of 15 terms alphabetically close to ANALYSE. It may lead to the following result :

	NOMBRE DE DOCUMENTS :		
	NTIS	INSPEC	PASCAL
1. ANALYERS	2	1	
2. ANALYSIS	3	15	
3. ANALYSISIS		2	5
4. ANALYSABLE		2	1
5. ANALYSATOR		2	1
6. ANALYSE	228	350	300
7. ANALYSED	3	105	10
8. ANALYSER	51	2 708	1
etc ...			

MESSIDOR gives the number of indexed documents in each database. The period means the term is unknown in this database e.g., ANALYERS in PASCAL. Further, each term is given a code e.g., ANALYSE is given 6.

### 3.2.3. - SELECT command

It is possible to select terms from the LIST

list using the SELECT command. The code is used to avoid typing long terms e.g.,

SELECT 4, 8

This requires the selection of the terms ANALYSABLE, ANALYSER. Terms will be combined using a logical sum i.e., the command (see example below) will give the number of documents which are indexed by either ANALYSABLE or by ANALYSER or by both. Further, each search step, for which the result is a number of documents (i.e. a search step which select a set of documents) is given a number. This makes it possible to combine these sets using this number (see § 3.2.5)

Example

```

ETAPE 1
?? SELECT 4, 8
2763 DOCUMENT(S) REPARTI(S) COMME SUIT :
NTIS 31
INSPEC 2710
PASCAL 2
  
```

\* meaning : 2763 DOCUMENTS DISTRIBUTED AS FOLLOWS.

### 3.2.4. - FIND command

It is also possible to select sets of documents by directly typing dictionary terms using the FIND command. This command does not require the use of the LIST command. Terms - Keywords (KW) or Author (AU) - may be combined using Boolean logical operators. There are three logical operators :

- (1) Logical sum (represented by +)
- (2) Logical product (represented by \*)
- (3) Logical complement (represented by /)

Examples :

```

ETAPE 2
?? FIND KW = NUMERICAL CONTROL + NUMERICAL CONTROL SYSTEMS
398 DOCUMENT(S) REPARTI(S) COMME SUIT :
NTIS 288
INSPEC 100
PASCAL 10
ETAPE 3
?? FIND AU = MARTIN (D.) * SMITH (JOHN)
2 DOCUMENT(S) REPARTI(S) COMME SUIT :
NTIS 1
INSPEC 1
PASCAL 0
ETAPE 4
?? FIND KW = COMPUTER/COMPUTER AIDED DESIGN
30915 DOCUMENT(S) REPARTI(S) COMME SUIT :
NTIS 10100
INSPEC 11005
PASCAL 9810
  
```

In step number 2 (ETAPE 2) MESSIDOR selects a set containing the documents indexed by either NUMERICAL CONTROL or NUMERICAL CONTROL SYSTEMS or both.  
 In step number 3 (ETAPE 3) MESSIDOR selects a set containing the documents written by the both authors MARTIN (D.) and SMITH (John).  
 In step number 4 (ETAPE 4) MESSIDOR selects a set containing the documents indexed by COMPUTER but not by COMPUTER AIDED DESIGN.  
 The FIND command can also restrict the search output by the use of a publishing data criterion

Example :

```
ETAPE 3
?? FIND DA ET = 4/80 - 81
310 DOCUMENT(S) REPARTI(S) COMME SUIV :
  NTIS      102
  INSPEC    105
  PASCAL    103
```

- \* DA for DATE
- ET for ETAPE (meaning : step)

In this step MESSIDOR extracts, from the set selected in step number 4, a subset containing the documents published in either 1980 or 1981.

3.2.5 - COMBINE command

Selected sets may be combined using Boolean logical operators by means of the COMBINE command. In this case the user indicates the step number in which the set was selected. The logical operators are the same as the FIND command ones (see § 3.2.4.)

Example :

```
ETAPE 6
?? COMBINE 2 * 3
1 DOCUMENT(S) REPARTI(S) COMME SUIV :
  NTIS      1
  INSPEC    0
  PASCAL    0
```

In this step MESSIDOR extracts, from the set selected in step number 2, a subset containing the documents written by MARTIN (D.) and SMITH (John).

3.2.6. - DISPLAY command

More details on selected documents can be viewed or printed by means of the DISPLAY command. Reference, title, publishing date, author(s), key words, and abstract may be displayed for each document.  
 The documents to be displayed must have been selected by a FIND, COMBINE or SELECT command.

Figure 4 shows extracts of printed edition of documents. MESSIDOR gives for each document the name of the LD which contains it.

3.2.7 - HISTORY command

The history of a search may be viewed by means of the HISTORY command. The result indicates the step number, the Boolean expression, and the global result.

Example :

?? HISTOIRE		
HISTORIQUE DE LA RECHERCHE		
ET	EXPRESSION	RESULTAT
1	ANALYSTE = ANALYSER .....	2 763
2	NUMERICAL CONTROL - NUMERICAL SYSTEMS .....	398
3	MARTIN (D.) = SMITH (JOHN) .....	2
4		
5		
6		
7		
8		

3.2.8. - LOCAL and GLOBAL command

During a search session, the user may search one of the LDs using its own query language. He does this by entering the command LOCAL followed by the LD's code. He may ask for the disconnection of the other LDs if he wishes. The user may again search the DBD with the MESSIDOR language simple by entering the GLOBAL command.

Example :

?? LOCAL 3

The user will now search the LD NTIS with the QUEST language.

?? GLOBAL

The user now may again search the DBD with the MESSIDOR language.

3.2.9. - STOP command

The last thing the user has to do when he has ended his search is to ask for the disconnection of each online database. The command that does this is the STOP command. After entering STOP, MESSIDOR generates all the appropriate disconnection procedures.

3.2.10. - HELP command

It is possible to get full explanation of every command's use by entering the HELP command. The explanations of each command includes examples of its usage. The command may be used either during a search or when the system MESSIDOR is off line.

3.3. - Locating documents

A user generally likes to know where he can get the documents he has selected by his search.

The MESSIDOR system offers commands allowing documents to be located. The location of documents is performed by the REVUE system to which MESSIDOR connects automatically [BOS 80].

#### 4. CONCLUSION

We have presented the MESSIDOR system. We have discussed the design and presented the user interface, i.e. the MESSIDOR language. This language is very simple and similar to the standard of EURONET.

A prototype of the system was developed in INRIA. The following bibliographic databases may be accessed :

- (1) NTIS, INSPEC, NASA and PASCAL. These databases are managed by the QUEST system and implemented by the ESA (European Space Agency).
- (2) Databases of the INRIA information center managed by the MISTRAL system.

The networks used are, respectively, THYMSHARE and RHIN. Also, it is planned to offer in a near future the access to QUESTEL DBs. QUESTEL offers about twenty BDs managed by MISTRAL and accessible through TRANSPAC.

MESSIDOR is the only information retrieval system that allows a user to work in a single query language with several bibliographic databases each using a different query language. A bibliographic database using any query language may be easily added provided that corresponding translators could be built.

It should then be easy to use MESSIDOR with EURONET BDs, DIALOG BDs etc. The features of MESSIDOR, in conjunction with its small cost, should render MESSIDOR a useful tool for any information retrieval center.

#### BIBLIOGRAPHY :

- [ZIM 79] H. ZIMMERMANN : Developments in network architecture and protocols standards, 17 th Infotech State of Art Conf., Paris, Apr 80.
- [1] Guide Euronet Diane 1981 : "Toute L'information pour l'Europe".
- [LIT 80] W. LITWIN : "A model for a Distributed Data Base", Proc. of the ACM 2nd Annual Louisiana Computer Exposition, Feb. 1980, SIRIUS report (MOD-I-029).
- [LEB 81] J. LE BIHAN : "A french nation wide project on distributed data bases", Proc. of the 6th international Conference on Very Large Data Bases, Oct. 80, Montréal.
- [NEG 77] A.E. NEGUS : "Standard Commands for retrieval systems" Interim report on a study being carried out for the Commission of the European Communities - May 77.
- [FAU 80] J.C. FAURE : Cahier des Charges du système MESSIDOR, SIRIUS report.
- [MOU 81] C. MOULINOUX : Le Logiciel de Commandes du langage MESSIDOR, SIRIUS report.
- [BOS 80] P. BOSC : Query processing in REVUE : a DBMS - oriented system, SIRIUS report.