IBM World Trade

Application Program

GH19-0012-1

385

ESRI REFERENCE MATERIAL
NOT TO BE REMOVED FROM LIBRARY

IBM System/360
Information Retrieval and Management System (IRMS) (360A-CR-03X)
Application Description Manual

System/360 Information Retrieval and Management System (IRMS) operates under the control of the System/360 Disk Operating System (DOS). It processes information files characterizing documents which can be queried to provide Selective Dissemination of Information (SDI) and Retrospective Search Results.

This manual surveys the operational concepts and capabilities of System/360 IRMS. Routines within the system are written in System/360 Assembler language under DOS.

BIBLIOTHEQUE DU CERIST

IST
54

IBM

Second Edition (January, 1970)

This publication is a major revision of H19-0012-0 and is now renumbered GH19-0012-1. It incorporates changes released in Technical Newsletter N19-0041 dated January 15, 1969. Changes to the text are indicated by a vertical line to the left of the text. Changes to figures are indicated by a • to the left of the figure number.

Any subsequent changes will be reported in further revisions or Technical Newsletters.

BIBLIOTHEQUE DU CERIST

02

Address comments regarding the contents of this publication to:

IBM France
Application Development Center
94-96 Rue Réaumur
75 Paris 2e
France

© International Business Machines Corporation, 1968

CONTENTS

PREFACE iii

INTRODUCTION 1

GENERAL DESCRIPTION OF THE APPLICATION 1
General Aspects of the System 1
Purposes and Objectives of IRMS 2
Extent of Coverage 3
Advantages 4
Advantages of the Systems Approach 4
Advantages Derived from the Use of IRMS 4
Application-Oriented Concept 4
Machine-Oriented Concepts 4
Performance 5
General System Logic 5
Thesaurus File 6
Inverted Search File and Bibliographic File 8
Requests: SDI Profiles and RS Queries 9
General Features of Search 9
Request Structure 11
Request Preparation Rules 11
Request Form Sheet 11
Request Processing 13
Statistical File 13

MACHINE AND SYSTEM CONFIGURATION 13

IRMS INPUT/OUTPUT EXAMPLES 14

APPENDIX I. Analysis Sheet Input 15

APPENDIX II. Input to the Thesaurus and Output from the Thesaurus Updating Program DCT 1 ... 16

APPENDIX III. Query/Profile Input and Printout 19

APPENDIX IV. Thesaurus File Printout 24

APPENDIX V. Glossary 28

APPENDIX VI. Bibliography 30

Index 31

ILLUSTRATIONS

Figure 1 Basic System Layout iv
Figure 2 Interaction of Thesaurus File with System Input and Output 2
Figure 3 Extent of Mechanization 3
Figures 4 and 5 Performance Charts 5
Figure 6 Thesaurus File 6
Figure 7 Thesaurus File (Further Clarification) 6
Figure 8 General Flowchart Showing Preparation and Updating of Thesaurus File 7
Figure 9 General Flowchart Showing Preparation of Inverted Search File and Bibliographic File 8
Figure 10 General Flowchart Showing Updating of Inverted Search File 9
Figure 11 Simple Illustration of Search 10
Figure 12 Search Program -- Input, Output, and Files Consulted 10
Figure 13 Logical Connector Codes 11
Figure 14 Typical Machine Configuration for IRMS ... 14
Figure 15 Analysis Sheet Input 15
Figure 16 Relation Table 16
Figure 17 Semantic Field Table 16
Figure 18 Input for Deleting 17
Figure 19 Output from Thesaurus Updating Program DCT1 18
Figure 20 Request Form Sheet 20
Figure 21 Query/Profile Input and Printout 21
Figure 22 Query/Profile Input and Printout 21
Figure 23 Query/Profile Input and Printout 21
Figure 24 Query/Profile Input and Printout 22
Figure 25 Query/Profile Input and Printout 22
Figure 26 Query/Profile Input and Printout 23
Figure 27 Thesaurus Printout 25
Figure 28 Alphabetic Dictionary 26
Figure 29 Synonym Dictionary 27

BIBLIOTHEQUE DU CERIST

PREFACE

The information explosion is engulfing our organizations and institutions with vastly increasing amounts of specialized literature. Library services are rapidly becoming overburdened at a time when their function grows more essential, and the urgent need for automated document retrieval is becoming more acute day by day.

IBM System/360 Information Retrieval and Management System (IRMS)* provides a set of programs designed to meet this need. It enables a user to prepare and maintain files of information characterizing his documents which can be queried at intervals to provide Selective Dissemination of Information (SDI) and queried for information found by a one-time Retrospective Search (RS).

Since IRMS, in line with conventional document retrieval practice, is based on characterization by means of word-type descriptors and a thesaurus file, it will be fully compatible with future developments in this rapidly expanding field.

Two such developments are source-automation of documents and remote interrogation. It is very probable that specialized literature will be prepared in a format suitable for automatic reading by character-recognition devices or in such a way that machine-readable copies will be obtained as by-products. This material will be transmitted directly to the documentation center in machine-readable form where it will become available even before publishing date. Remote interrogation will enable the user to request information by telephone and receive typewritten or visual-display answers. These new developments are feasible with today's hardware, and documentation centers incorporating them are expected to be in operation in the near future.

This manual describes the functions provided by IRMS and explains how the user goes about taking advantage of them. Technical detail has been kept to a minimum.

*Previously called SAGESSE (French acronym for Système Automatique de GEStion et de SElection de Documentation)

BIBLIOTHEQUE DU CERIST

BIBLIOTHEQUE DU CERIST

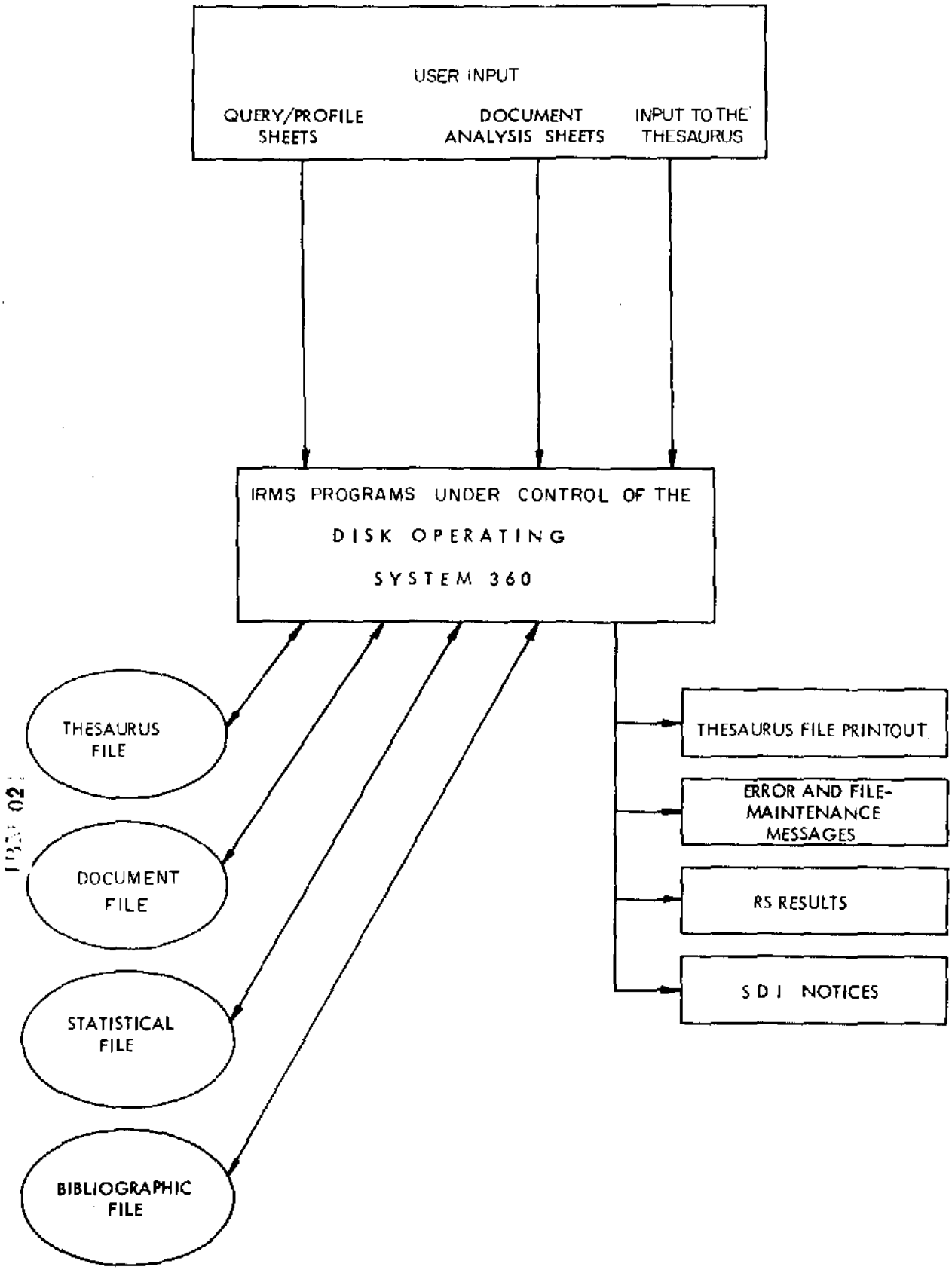

Figure 1. Basic System Layout

BIBLIOTHEQUE DU CERIST

INTRODUCTION

IRMS is designed to provide the user with an efficient document retrieval system that will carry out Selective Dissemination of Information and Retrospective Searches. Briefly, it operates as follows.

The user, who begins with a collection of documents, assigns a serial number to each one of them. He then prepares files of information characterizing his documents in such a way that the files may be queried to obtain information identifying the particular documents he desires. The results obtained consist of document numbers plus certain other bibliographic information and abstracts.

Basically, there is little difference between a request for information and the characterization of a document, since a request is in effect nothing more than a description of an ideal document which the inquirer hopes to find in the file. Both the requests and characterization files contain alphabetic, unformatted indexing information expressed as descriptors. Two types of requests are used: SDI profiles and RS queries.

Other useful output may be obtained from the system in the form of file printouts and statistics. The system is designed to offer the user considerable flexibility.

IRMS is written in System/360 Assembler language to run under DOS.

GENERAL DESCRIPTION OF THE APPLICATION

GENERAL ASPECTS OF THE SYSTEM

The basic layout of the system is illustrated in Figure 1. The three types of input shown in the User Input block are:

- Query/profile sheet — A series of descriptors (and logical connectors) that characterize the information sought by the user
- Analysis sheet — A series of descriptors that characterize the information contained in the user's documents
- Input to the thesaurus — Descriptors Information

IRMS comprises seven programs. Three are dictionary programs for the thesaurus file creation and maintenance: DCT1, DCT3, and DCT4. Three are for the document, bibliographic, and statistical files management: DCT2, INV1, and INV2. The last is the search program SRCH.

The four user-prepared files are shown in the ovals on the left-hand side of Figure 1:

- Thesaurus file — Contains interrelated descriptors
- Document file — Contains Document Accession Numbers (DANs) filed by descriptor number (DNR)
- Statistical file — Created as a by-product of the updating of the inverted search file. Used for utility purposes and preparing management reports
- Bibliographic file — Contains bibliographic data and abstracts filed by DAN

The output blocks are shown on the right-hand side of Figure 1. There are three thesaurus printout options and a number of error and file maintenance messages. RS queries are answered with RS results; SDI profiles are answered with SDI notices.

Figure 2 illustrates other basic aspects of the system.

In this figure, everything to the right of the broken line may be considered to be the mechanized process; the rest may be considered the manual process. Three inputs are shown on the left: the raw document in the author's language which is converted to an analysis sheet by the analyst; updating input taken from existing dictionaries, glossaries, etc., and from the thesaurus file itself (for example, deletion input for superfluous descriptors); and raw requests in the user's language that are converted into query/profile sheets by the analyst. These sheets represent either RS queries or SDI profiles, both of which are processed in exactly the same way, with one exception: the former are obtained by consulting the cumulative files, while the latter are obtained by consulting only recent additions to the files.

The three major files are shown in the middle towards the right-hand side (statistical file is not shown).

The raw document is sent first of all to the analyst, who converts it to an analysis sheet containing DANs, descriptors, and bibliographic data. The descriptors are checked against the thesaurus file. The checked descriptors are assigned descriptor numbers, DNRs, that are entered with DANs in the inverted search file.

The DANs and bibliographic data are entered in