

FAIRS による文献検索

情報処理センター 阪上直美

FAIRSとは、文献情報、数値情報、テキスト情報など、多くの情報分野をサポートする情報検索システムであり、富士通が開発したものである。

情報をFAIRSの管理のもとにディスク上のファイルに構築し、データベースを形成する。そしてそのデータベースを検索する。

FAIRSには、文献など、文字情報を対象としたFAIRS I、人事情報など数値情報を対象としたFAIRS IIがある。ここでは、FAIRS Iの概要、利用について説明する。

FAIRSには、次の三つのシステムからなる。

1. FAIRSの運用管理
2. データベースの情報蓄積、更新
3. データベースの情報検索

FAIRSでは、システムの利用者を次の三つに分類する。

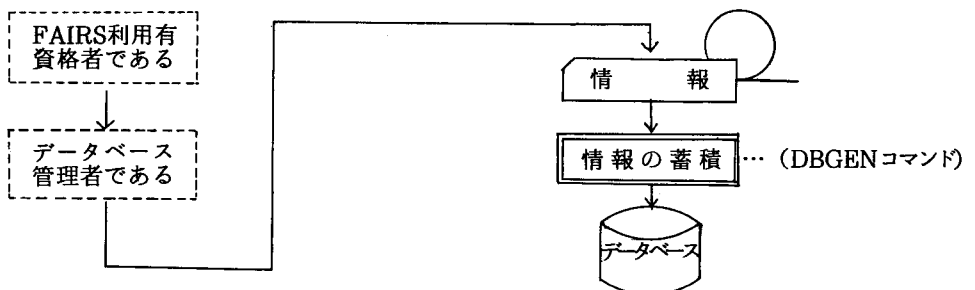
1. システム管理者
2. データベース管理者
3. 一般利用者

システム管理者とは、FAIRSの運用について管理する者であって、主にFAIRS利用者の登録FAIRSシステムの生成を行う。通常、計算機システム管理者がこれに当たる。データベース管理者とは、データベースの登録（情報蓄積）者のことであり、データベースの更新、データベース利用者の登録を行う。一般利用者とは、主にデータベースを検索する者である。

これらは、その管理者としての資格、利用者としての資格があるかどうか検査され、許可された範囲内で、FAIRSの利用、データベースの利用権が与えられ、機密保護は十分に保たれる。

FAIRSの利用は、バッチ処理でも、TSS処理でも可能であり、FAIRSシステムを動かす事により、TSSコマンドと同じ様なFAIRS用のコマンドを利用することによって、情報の蓄積、検索ができる。

1. データベース情報の蓄積について



(1) 情報処理センター「TOSHOデータベース」の情報蓄積、登録について

情報処理センター図書のうち、計算機に関する書物を選び、1冊毎に、下記の様な項目にまとめ、その項目に対する内容（項目値）が、単語形式（Character属性）か文章形式（TEXT属性）か、又、何文字（バイト）格納するか決めて、「TOSHOデータベース」として、ディスク上へ1ファイルとして登録した。

○ 「TOSHOデータベース」項目について

	項目（略）	文字数（バイト）	属性		生起回数
1.	整理番号（ABN）	4	CHARACTER	右詰	
2.	標 題（TI）	100	TEXT		
3.	要 約（AB）	300	TEXT		
4.	分類コード（CL）	12	CHARACTER		
5.	キーワード（I）	100	CHARACTER		10
6.	著著名（AU）	200	CHARACTER		10
7.	編集者名（E）	200	CHARACTER		10
8.	翻訳者名（T）	200	CHARACTER		10
9.	出版社名（P）	80	CHARACTER		
10.	発行年月日（PD）	10	CHARACTER	右詰	
11.	ページ数（PN）	4	CHARACTER	右詰	

ファイルの中へは、項目名は英字だが、項目値はカナ文字使用にした。これは今日、日本語処理が注目されている故、今後何らかのデータになればよいと思い、テスト的にカナ文字にて登録した。

文字数については、固定的に指定文字数だけ領域をとる場合と、可變的に（余った部分をつめる）領域をとる場合とがあり、整理番号、分類コード、ページ数は、固定の文字数分とった。

標題は本の題名、要約とは内容を簡単にまとめたもの（目次）、分類コードとは附属図書館での分類番号、キーワードとはその本の内容を示す言葉、いくつかのキーワードの事である。

（文章形式項目とは、一つ以上の単語（文字列の前後が、英数字又は、カナ文以外のものからなる）からなり、一つ以上の文（. ? ; ! の後ろに二つ以上の空白）からなる。

○ 「TOSHOデータベース」登録について

前述の項目別にまとめが終り、コーディングも済むと、次は、実際に、ファイルを作成するためのデータ作成である。TSS端末機で、登録用コマンドを用いて、作成できるが、新規に創成するファイルのデータであり、もし、壊れた場合、又、属性を変更して、もう一度ファイルを作り直したりする場合のため、最終的に、カードにデータをパンチし、バッチ処理で「TOSHO データベース」を創成した。

技術解説

次に、順にデータベースの作成を示す。

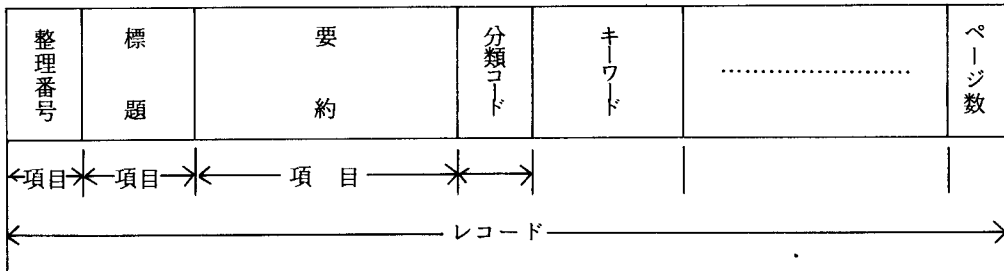
データの蓄積を行う前に、蓄積すべきデータの論理構造を定義する。

① データベースの全般的な定義（説明情報や、扱う文字種別）

データベースの定義情報を格納する論理構造ファイル（LRSFILE）と、実際のデータを、蓄積すべきサブファイル（SUBFILE）の領域の開設と初期化。

② マスタレコードの論理構造定義

個々の情報（図書文献情報における1文献）をマスターレコード（レコード）という単位で編集する。そのレコードの枠組みの定義である。



• ELEMENTコマンドにより

項目、及び、そのデータ長（文字数）生起回数、右詰めか左詰めかなどの定義

• APPENDIXコマンドにより

項目に対して、説明文、項目見出し文字列の定義

• INDEXコマンドにより

項目のうち、特に、検索項目（インデックスインバーテッド）として使用するものを定義

.....キーワード項目を定義

• KEYWORDコマンドにより

文章形式項目であるものについて、各項目値（文章）の中から、抽出される単語を検索の対象とするキーワードインバーテッド項目を定義

• CLASSIFYコマンドにより

「TOSHOデータベース」の検索時に参照不可項目を定義

③ データの蓄積

②で定義したデータベースの論理構造定義に基づき、実際に、図書に関するデータをINPUTコマンドを用いて入力する。

これにより、サブファイルに、マスタレコードが蓄積される。

④ インバーテッドファイルの作成

検索を高速に行うために、INVERTコマンドにより、インバーテッドファイルを作成する。

次に登録時のソースリストを示す。

```

① //FAIRSCAT EXEC FAIRSCAT
② //SYSPRINT DD SYSOUT=A
③ //SYSOUT DD SYSOUT=A
④ //SYSIN DD *
⑤ FAIRS U(F0004) P(1111 CLASS1 4)
⑥ SCRATCH TOSHO SUBFILE(*) P(1234 ELPMAS 5)
⑦ INITIALIZE TOSHO U(F0004)
⑧ LRSFILE SPACE(1) UNIT(PUB)
⑨ SUBFILE SUB1 SPACE(1) UNIT(PUB)
⑩ CLASSIFY ELPMAS PASSWORD('A+D')
⑪ END
⑫ DBGEN
⑬ SELECT TOSHO S(SUB1) P(1234 ELPMAS 5)
⑭ DATABASE OWNER('NAOMI.SAKAUE') /
⑮ UNIT('DOCUMENT(S)') /
⑯ TITLE('=== TOSHO ===') /
⑰ EXPLANATION('TOSHO DATABASE INCLUDES 10 DOCUMENTS') /
⑱ SUMMARY('TOSHO DATABASE') /
⑲ SUBFILE SUB1 /
⑳ LEVEL(5) /
㉑ EXPLANATION('TOSHO DATABASE INCLUDES 10 DOCUMENTS') /
㉒ SUMMARY('TOSHO SUBFILE(SUBFILE-1)') /
㉓ ATTRIB LETTER(KANA)
㉔ KEY DOCUMENT_NUMBER ALIAS(DN) LENGTH(6) FIXED
㉕ ELEMENT ABSTRACTNO ALIAS(ABN) LENGTH(4) FIXED CHARACTER /
㉖ ADJUSTMENT(RIGHT)
㉗ ELEMENT TITLE ALIAS(TI) LENGTH(100) TEXT
㉘ ELEMENT ABSTRACT ALIAS(AB) LENGTH(300) TEXT
㉙ ELEMENT CLASS ALIAS(CL) LENGTH(12) FIXED CHARACTER /
㉚ ADJUSTMENT(LEFT)
㉛ ELEMENT INDEX ALIAS(I) LENGTH(100) VARIED CHARACTER /
㉜ DC(10)
㉝ ELEMENT AUTHOR ALIAS(AU) LENGTH(200) VARIED CHARACTER /
㉞ DC(10)
㉟ ELEMENT EDITOR ALIAS(E) LENGTH(200) VARIED CHARACTER /
㊱ DC(10)
㊲ ELEMENT TRASLATOR ALIAS(T) LENGTH(200) VARIED CHARACTER /
㊳ DC(10)
㊴ ELEMENT PUBLISHR ALIAS(P) LENGTH(80) VARIED CHARACTER
㊵ ELEMENT PUBDATA ALIAS(PD) LENGTH(10) FIXED CHARACTER /
㊶ ADJUSTMENT(RIGHT)
㊷ ELEMENT PAGENO ALIAS(PN) LENGTH(4) FIXED CHARACTER /
㊸ ADJUSTMENT(RIGHT)
㊹ APPENDIX DN HEADER('DOCUMENT NUM')
㊺ APPENDIX AU HEADER('AUTHOR')
㊻ APPENDIX CL HEADER('CLASS')
㊼ INDEX ABN
㊽ INDEX PN
㊾ INDEX I
㊿ KEYWORD KEYWORD ALIAS(KW)
① CLASSIFY CLASSX NOSEE(PAGENO) PASSWORD('A+D')
MAP
② INPUT LIST
③ ADD SAM001
④ -ABN 1
⑤ -TI コンピユ-タ システム-シヨク
⑥ -AB コンピユ-タ システム-シヨク シツケン ノ ケイカク。 ランスウ ハツセイ ノ ホウホウ。 マチキヨクシツ。 スケジムノ
⑦ イケリシツ システム ノ コンピユ-タ モデル。
⑧ -CL
⑨ -I システム-シヨク
⑩ -AU T.H.ネイラ-/J.L.ハリアンフ/D.S.ハ-チイツク/K.チユ-/
⑪ -E
⑫ -T ナカノ イキオ/コハマシ ヨシオ/
⑬ -P ハイフクカク
⑭ -PD 1971.06.10
⑮ -PN 350

```

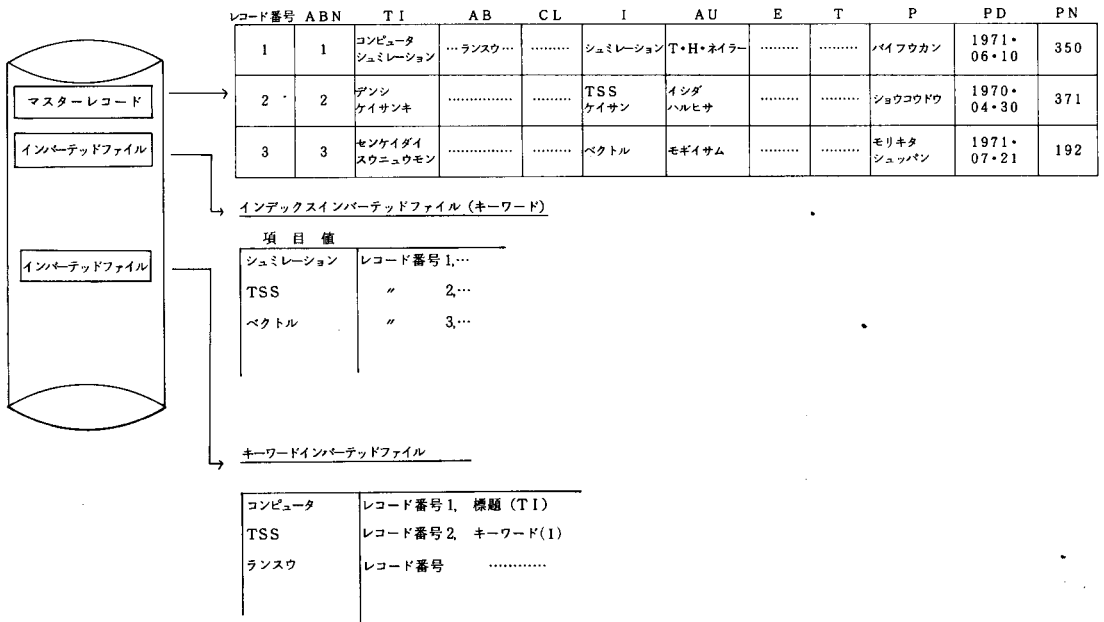
技術解説

```

⑥ ADD SAM050
⑦ -ABN 50
⑧ -AB ヘクトル カイトキ。 セキフツ。 キヨウレツシキ。 キヨウレツ。 レンリツホウテイシキ。 フーリエ キヨウスウ。
⑨ -CL 411.3.4C
⑩ -I スウチケイサン/ヘクトル/
⑪ -AU E.クライツイク
⑫ -E
⑬ -T タシマ イチロウ/コトウ シロウ/
⑭ -P ハイフウカン
⑮ -PD 1970.09.15
⑯ -PN 267
⑰ END
⑱ INVERT
⑲ END
⑳ END

```

「TOSHOデータベース」のサブファイルの構造



ソースリストの説明

- ①～④ FAIRSを起動させる制御文
- ⑤ FAIRSの利用に当たり、有資格の定義、コマンドクラスを指定
- ⑥～⑨ データベースの作成の前のファイル領域の開設と初期化
- ⑩ FAIRS利用者の特権クラス名、パスワードを指定
- ⑫ 情報の蓄積、及び更新を行うDBGENサブシステムの起動
- ⑬～⑲ 蓄積するデータベースの定義
- ⑳ マスターレコードを構成する項目のうち、レコードキー項目についての定義
- ㉑～㉓ レコードキー項目以外の項目に対する属性定義
- ㉔～㉖ 項目に対する説明文、項目見出し文字列の定義
- ㉗～㉙ 特に検索項目 (インデックスインバーテッド) として使用するものの定義

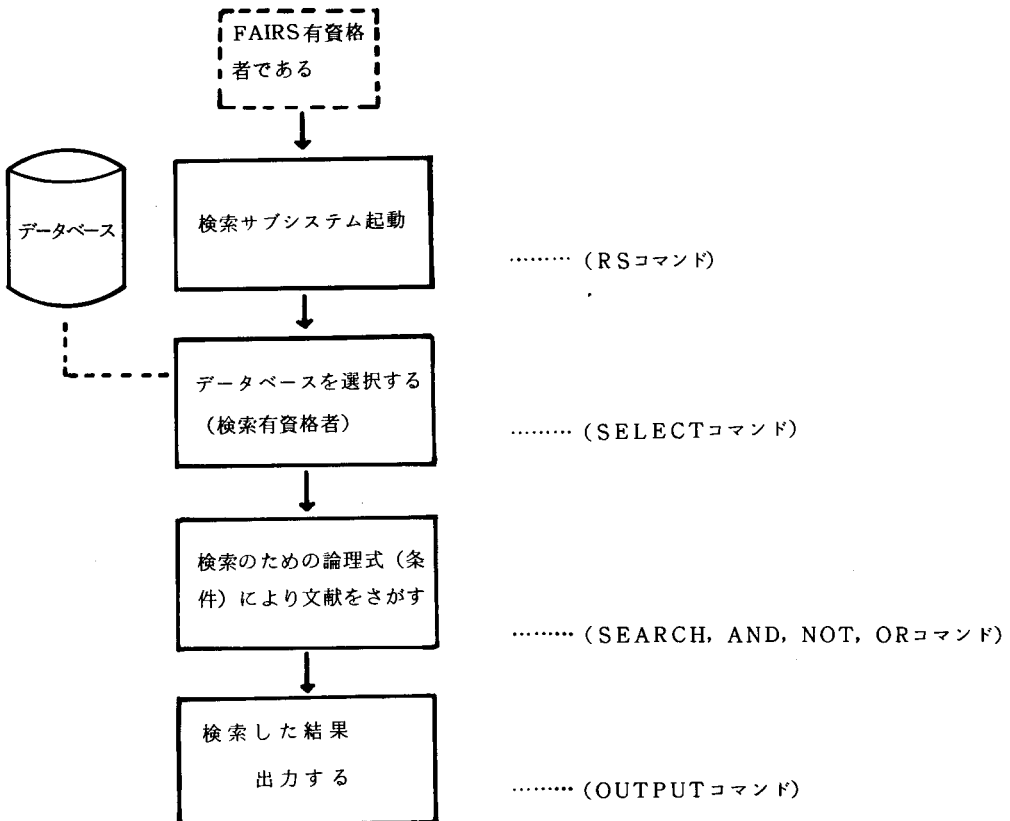
- ⑤⑩ 文章形式項目であるものについて、その中から抽出される単語を検索の対象とする
(キーワードインバーテッド) 項目の定義
- ⑤⑪ マスターレコードを構成する項目ごとに機密保護を行う
- ⑤⑫ マスターレコードの作成, 追加, 更新, 削除を行う
- ⑤⑬ ⑤⑭ マスターレコードを, サブファイルに登録する
- ⑤⑮~⑤⑯ } 項目名と項目値を入力
- ⑤⑰~⑤⑲ } 項目名と項目値を入力
- ⑦⑦ INPUTコマンドの終了
- ⑦⑧ インバーテッドファイルを作成する
- ⑦⑨ DBGEN終了
- ⑧⑩ FAIRS終了

2. データベースの検索について

TSS端末機でコマンド (FAIRS用) を用いることによって, データベースの検索ができる。

データベースを検索するには, そのデータベースの検索有資格者とならなければ, 検索はできない。ここでは, その説明は, 省く。

(1) 検索する手順について



技術解説

(2) 検索コマンドについて (SEACH, AND, NOT, OR コマンド)

データベースに蓄積されている情報の中から、必要とする情報を検索するために、条件を与える。条件は、項目名と、表1：関係演算子と表2：文字式とで表わす。※

AND, NOT, OR コマンドは、直前の検索結果の集合に、更に条件を付加して、検索させるものである。

これらのコマンドの実行が終了すると、条件を満足したレコード件数が表示される。この検索結果は、OUTPUTコマンドで、内容を端末に表示する。

※条件式を、次の様に用いる。

(例) 項目値KW (キーワード) が, "Computer" であるものをさがす。

SEACH KW EQ COMPUTER
 項目名 論理演算子 文字式

論理演算子の種類：優先順位

(論理差) NOT 1
 (論理積) AND 2
 (論理和) OR 3

表1 関係演算子の種類

関係演算子	省略形	意 味	項目の形式		
			左詰	右詰	文章
EQ	=	項目の値が文字式と等しい	○	○	○
NE	≠	項目の値が文字式と等しくない	○	○	○
GT	>	項目の値が文字式より大きい		○	
GE	>=	項目の値が文字式より大きいか等しい		○	
LT	<	項目の値が文字式より小さい		○	
LE	<=	項目の値が文字式より小さいか等しい		○	
HAS	なし	項目の値が文字式の語を含んでいる			○*
HASNT	なし	項目の値が文字式の語を含んでいない			○*

* HAS/HASNTはキーワードインバーテッドファイルがあるときだけ使用できる。

** 項目の属性に従って○印のものが使用可能である。

表2 文字式の指定方法

種 類	形 式	使用可能 関係演算子	項目の属性		
			左結	右結	文章
完全一致	文字列又は語番号	* すべて使用可能	○	○	○
前方一致	文字列@又は 語番号@	* EQ, NE, HAS, HASNT	○		○
後方一致	@文字列又は @語番号	EQ, NE	○		○
両端一致	文字列@文字列, 文字列@語番号 語番号@文字列又は 語番号@語番号	* EQ, NE, HAS, HASNT	○		○
任意一致	@文字列@又は @語番号@	EQ, NE	○		○
範囲指定	文字列:文字列, 文字列:語番号, 語番号:文字列又は 語番号:語番号	EQ, NE		○	

* 関係演算子がHAS/HASNTの場合にはキーワードインパードファイルがなければならない。

** 項目の属性に従って○印のものだけが使用可能である。

次に、「TOSHOデータベース」の検索例を示す。

…… 端末入力を示す。

```

READY
FAIRS
+FCA000I F A I R S (V02/L02) .STARTED
+FAR001A ENTER COMMAND CLASS NAME-
+FAR003A ENTER PASSWORD FOR 4100-
FAIRS> RS
RS> SELECT TOSHO
+FAR002A ENTER SECURITY CLASS NAME-
+FAR003A ENTER PASSWORD FOR 2221-
RS> SEARCH I EQ シュミレーション …… (キーワードで「シュミレーション」を持つものをさがす)
+FRS100I 1 DOCUMENT (S) FOUND
RS> OUTPUT
=== TOSHO ===
#1
DOCUMENT NUM SAM033
ABSTRACTONO      33
TITEL            ランスウ ノ チシキ
ABSTRACT         イチヨウ ランスウ ノ ツクリカタ. ランスウレツ ノ ケンテイ. イチヨウ
                  ランスウ レツ カラ ノ ヘンカン. ランスウレツ ト シュミレーション.
```


技術解説

CLASS 417.LO.A
INDEX ランスウ
シユミレーション
AUTHOR ワキモト カズマサ
PUBLISHR モリキタシユツパンK.K
PUBDATA 1970.08.20

RS> SEARCH TI HAS ケイサン (標題で「ケイサン」を持つものをさがす。)

+FRS100I 3 DOCUMENT (S) FOUND

RS> AND I EQ ケイサン (かつ、キーワードが「ケイサン」であるものをさがす。)

+FRS100I 1 DOCUMENT (S) FOUND

RS> OUTPUT

=== TOSHO ===

#1

DOCUMENT NUM SAM008

ABSTRACTONO 8

TITLE ケイサン ノ リロン

ABSTRACT ケイサン カノウ セイ ノ リロン. イツパンリロン ノ オウヨウ.
イツパンリロン ノ テンカイ.

CLASS 418.2

INDEX ケイサン

AUTHOR M.デーヴィス

TRASLATOR ワタナベ シゲル

PUBLISHR セキ セツヤ

イワナミシヨテン

PUBDATA 1966.10.28

RS> END

FAIRS> END

+FCA100I FAIRS ENDED

READY

参考文献

FACOM OSIV/F4 FAIRS-I使用手引書