

FAIRS-I/JEFによるJICST科学技術文献ファイル(情報工学)の検索 : 日本語文献データの検索

二村, 祥一
九州大学大型計算機センター研究開発部

平野, 広幸
九州大学大型計算機センター業務掛

入江, 啓一
九州大学大型計算機センター業務掛

高木, 利久
九州大学大型計算機センター研究開発部

他

<https://doi.org/10.15017/1474925>

出版情報 : 九州大学大型計算機センター広報. 15 (1), pp.13-20, 1982-03-10. 九州大学大型計算機センター

バージョン :

権利関係 :

FAIRS-I/JEFによるJICST科学技術文献ファイル (情報工学)の検索—日本語文献データの検索

二村祥一*, 平野広幸**, 入江啓一**, 高木利久*, 鬼塚千代子*, 古城久美子*, 松尾文碩*

1. はじめに

富士通が開発した会話型情報検索システムFAIRS-Iは、現在、センターではINSPECの検索に使用され、そのほか2, 3の利用者によって個人用文献データの管理に用いられている。FAIRS-Iは、日本語機能をもつFAIRS-I/JEFが—昨年(1980年)夏から使えるようになった。センターでは、昨年(1981年)秋から日本科学技術情報センター(JICST)の科学技術文献ファイルを使ってFAIRSの日本語機能のテストを行ってきた。この度、利用者の方々にも、FAIRSによる日本語文献データの検索を体験していただくために、このシステムを公開することにした。これは、分野が限られていること(情報工学関係だけ); データ量が少いこと(1981年4月以降のテープのデータ); 使用する端末に限られること(JEF用FACOM6652ディスプレイサブシステムだけ)、更に検索できる者は九大の利用者に限られること(ただし、九大以外の利用者でもセンターの端末で使うのは差支えない)などにより、実験システムの域を出るものではないが、これにより利用者間に日本語データベース開発の機運が芽生えれば、センターの目的は達せられたことになる。

JICST科学技術文献ファイルは、JICSTが刊行している抄録誌「科学技術文献速報」の機械可読版で、周知のようにJICST自身がJOISによってこのファイルのオンライン検索サービスを行っている。今回、本センターでサービスを始めたデータベースには、管理・システム技術編と電気工学編から抽出した情報工学関係だけを入れている。現在の収録量は、昭和56年度からの分であり、月に1回データの追加を行っている。

この文献ファイルを検索できる者は、JICSTとの契約により、上述のように基本的には九大の利用者だけであるが、九大以外の利用者でもセンターにおいて利用することは許されている。

2. JICST科学技術文献ファイルのデータベース定義情報

JICST科学技術文献ファイルのデータベース名は、JICSTJである。JICSTJの各レコードは次の情報を含んでいる。

{ 抄録番号, 著者名, 所属機関名, 原文標題, 標題, 雑誌名, 抄録, キーワード, 発行年,
 { ISBN, CODEN

JICSTJのデータの例を図1に示す。抄録番号は、各文献を一意に識別するためにJICSTが付けた番号で、これは文献複写をJICSTに依頼する場合の文献識別番号にもなるものである。著者名、所属機関名、標題、雑誌名、キーワードは、それぞれ漢字表現と、フリガナ表現を持つ。漢字表現はJEF漢字コード、フリガナ表現はEBCDIC(カナ)である。原文標題、抄録は漢字表現のみ持つ

* 九州大学大型計算機センター研究開発部

** 同 業務掛

抄録番号 1009929
 著者 LAVERGNAT J, PORTES J, SYLVAIN M
 所属機関 Lab. Geophysique Externe, France
 標題(言語) 正弦波信号のパラメータ決定新アルゴリズム (EN)
 原文標題 A new algorithm to determine the parameters of a sinusoidal signal.
 雑誌 J Comput Phys, VOL. 36, NO. 3, P. 347-365, 1980, USA
 抄録 非線形関数の最小二乗法による推定的主問題は反復過程に対する出発値を見出すことである。スピン安定宇宙機による測定値など、正弦波信号の場合に初期値推定制約のないアルゴリズムを示した。パラメータが既知のデータ、実験データに応用し、既存の方法と比較した正弦波、パラメータ推定、最小二乗法、数値計算、計算機アルゴリズム、計測
 キーワード LAVERGNAT J
 著者 PORTES J
 SYLVAIN M
 所属機関 LAB. GEOPHYSIQUE EXTERNE, FRANCE
 標題 セイクンハシコンコウ ノ ハ°ラメータケツテイシナルコリスム
 雑誌 J COMPUT PHYS
 キーワード セイクンハ
 ハ°ラメータスイテイ
 サイシヨウニシヨウネウ
 スウチケイサン
 ケイサンキアルコリスム
 ケイソク
 発行年 1980
 CODEN JCTPA

図1. JICSTJ データベースのデータ例

いる。その他の発行年, ISBN, CODENはEBCDICで表わされる。

図1は英語の文献なので、原文標題と(和文)標題を持っている。日本語の文献の場合には原文標題はない。標題の後の言語は原文の記述言語であり、雑誌の最後は雑誌の発行国である。言語、及び発行国のコードは付録1にまとめる。キーワードは、文献の主題を表わす日本語の用語で、これはJICSTシソーラスにより統制されている。図1の文献は、ISBNの値は持っていない。

JICSTの項目定義を表1に示す。項目名は、EBCDICのものは2字、漢字のものは「J」で終

表1. JICSTJ データベースの項目定義

内容	項目名	属性		コード
抄録番号	AN	KEY	右詰	EBCDIC
キーワードフリガナ	KW	INDEX	左詰	EBCDIC
著者名フリガナ	AU	INDEX	左詰	EBCDIC
所属機関名フリガナ	AA	INDEX	左詰	EBCDIC
雑誌名フリガナ	JN	INDEX	左詰	EBCDIC
発行年	PY	INDEX	右詰	EBCDIC
標題フリガナ	TI	TEXT	左詰	EBCDIC
著者名	AUJ		左詰	出力 JEF漢字
所属機関名	AAJ		左詰	出力 JEF漢字
標題	TIJ		左詰	出力 JEF漢字
原文標題	OTJ		左詰	出力 JEF漢字
雑誌	JNJ		左詰	出力 JEF漢字
抄録文	ABJ		左詰	出力 JEF漢字
ISBN	BN		左詰	EBCDIC
CODEN	CO		左詰	EBCDIC
キーワード	KWJ		左詰	JEF漢字

る3字を付けている。FAIRS-I/JEFには、FDMSのようなカナ漢字変換の機能がないので、FACOM6652ディスプレイ端末からの標準的な検索方法は、EBCDIC(カナ)で検索条件を設定し、検索結果を漢字で表示することになる。そのため項目KW, AU, AA, JN, PYをINDEX項目に、

項目TIをTEXT項目に、項目AUJ,AAJ,TIJ,OTJ,JNJ,ABJを標準出力項目にしている。
項目AN,PYは、右詰項目なので、数値の大小比較が可能である。

JICSTJデータベースは、現在、ただ1つのサブファイルからなる。サブファイル名の形式は次のとおりである。

JICJxxyy { xx:西暦年度の下2けた。
 yy:通常なし。年度の途中でサブファイルを新規に作成した場合、その月を与える。

3. 検索例

JICSTJの検索例を示す。下線部が利用者の入力部分である。

- ① LOGON TSS F1234/PASSWORD P(FAIRS) S(768)*
KDS700011 F1234 LAST ACCESS AT 15:28:45 ON 82.008
KEQ564551 F1234 LOGON IN PROGRESS AT 15:31:28 ON JANUARY 8, 1982
JOB NO = TSU4449 CN(01)
KEQ569511 NO BROADCAST MESSAGES
INFORMATION RETRIEVAL SERVICE OF INSPEC DATABASE IS NOT AVAILABLE
ON THE FOURTH THURSDAY,AND THE FOURTH FRIDAY OF ALL MONTHS.
READY
- ② FAIRS LINESIZE(90) POINT(9)
+FCA0011 FAIRS-I/JEF(V02/L06B)
- ③ FAIRS> RS
- ④ RS> SHOW DB
+FRS2731 あなたが利用できるデータベースは次の通りです。
INSPECC INSPECB INSPECA INSPECJ INSPECCP INSPECBP JICSTJ
- ⑤ RS> SHOW DB(JICSTJ)
+FRS2721 指定されたデータベースは次の通りです。
DB名 件数
JICSTJ 14278
 概略説明: JICST-J DATABASE, 1981
(JICJ81) 14278
 概略説明: JICST-J (INFORMATION SCI.) 1981
- ⑥ RS> SEL JICSTJ
- ⑦ RS> SHOW EL
+FRS2741 利用できる項目名は全部で17個です。
KY AN AUJ AU AAJ AA TIJ TI OTJ JNJ JN ABJ BN CO PY KWJ KW
- ⑧ RS> SHOW EL(BN)
+FRS2811 指定された項目に関する情報は次の通りです。
BN 国際標準図書番号. International Standard Book Numberの略。
HEADER('ISBN')
CHARACTER; VARIED; ADJ(LEFT); NOOUTPUT; OCC(1); LENGTH(30);
- ⑨ RS> SHOW INDEX
+FRS2741 利用できる項目名は全部で6個です。
AN AU AA JN PY KW
- ⑩ RS> SHOW TEXT
+FRS2741 利用できる項目名は全部で2個です。
KY TI

* 4月以降は、TSSの使用仮想記憶域の標準値が1024キロバイトとなるので、S(768)は省略できる。

使用例の説明

- ① TSSセッションを開始する。専用のログオンプロシジャと使用仮想記憶域を、P(FAIRS), S(768)と指定する必要がある。
- ② FAIRSを起動する。OUTPUTコマンドで検索結果を日本語ラインプリンタに出力する場合の1行あたりの文字数、及び日本語文字列の文字の大きさ(ポイント数)を、LINESIZE(90), POINT(9)と指定している。オペランドの省略値は、LINESIZE(136), POINT(12)である。FAIRSの初期設定が完了するとモードメッセージFAIRS>が出力される。
- ③ 検索サブシステムRSを起動する。RSの初期設定が完了するとモードメッセージRS>が出力される。
- ④ SHOWコマンドで利用可能なデータベースを表示する。
- ⑤ データベースJICSTJの内容を表示する。現在のレコード数が14278件であることと、サブファイルとしてJICJ81のみがあることがわかる。
- ⑥ SELECTコマンドにより、検索の対象としてJICSTJを選択する。
- ⑦ JICSTJに定義されている項目名の全部を表示する。
- ⑧ 項目BNの説明と定義情報を表示する。
- ⑨ INDEX項目を表示する。
- ⑩ TEXT項目を表示する。項目KYは、キーワードインバーテッド項目であるが、この場合TEXT項目は1項目のみであるので、KYとTIは同じである。

⑪ RS> SEA KW EQ シンコウチノウ
+FRS100I 195 件みつかりました。

⑫ RS> AND KW EQ ハーターニンシキ
+FRS100I 13 件みつかりました。

⑬ RS> OUT
** JICST-J **

#1	著者 所属機関 標題(言語) 原文標題 雑誌 抄録	SCHWARTZ R D Singer Co. 人工知能における再焦点化と再生(EN) Refocus and resurgence in artificial intelligence. Proc IEEE Natl Aerosp Electron Conf, VOL. 1980, NO. 1/3, P. 26-32, 1980, USA 人工知能の分野はひどい失敗の結果、その開発はほとんど全く休止してしまった。しかし、現在起りつつある変化は強力な機械知能をめざす仕事は達成する見込みがあるという楽観視をうみつつある。新研究は機械よりも実現される知能行動のモデルに焦点が置かれている。本文は早期の失敗の因となったモデル化の誤りを示し、もっと高度に練られた計画によってもたらされる成功の本質を指摘した。すなわち人工知能と認識シミュレーションの失敗を述べ、Pribram氏の類語に言及し、ソビエトの研究法、西欧の研究法等について論及した
#2	著者 所属機関 標題(言語) 雑誌 抄録	津川定ゆき 機械技術自動車安全公署 知能自動車(JA) 計測技術, VOL. 8, NO. 12, P. 63-69, 1980, JPN この知能自動車は「人工の目」ともいふべき道路パターン認識装置と「人工頭脳」ともいふべき問題解決装置から構成され無人で走行する自動車であるとして、知能自動車の構成について述べ、ついで、通路パターン認識手法、問題解決手法、走行実験について説明
#3	著者 所属機関 標題(言語) 雑誌 抄録	飯島純一、油田信一、金山裕 電通大、筑波大 自立ロボット「山彦」の画像認識(JA) 情報処理学会全国大会講演論文集, VOL. 21ST, P. 793-794, 1980, JPN ロボットが移動する際必要となる基本的情報は、障害物の有無と障害物までの距離である。このため、超音波を利用してこれらの情報を得る機構を作成した。この方法は、単純な機構で済み、消費電力や重量が小さいという利点があるが、詳細な情報を得るには限界がある。そこで、テレビカメラを利用して画像認識の実験を実施

#4 著者 宮川達夫
所属機関 富士通研情報処理研究部
標題(言語) 特集:新しい計算機システム技術とその応用 III 計算機システム基礎技術の動向 第4章 パターン認識と人工知能(JA)
雑誌 電気学会雑誌, VOL. 100, NO. 11, P. 1016-1021, 1980, JPN

↳

⑭ RS> SORT KEY(AU)

⑮ RS> OUT EL(AUJ TIJ KWJ)

** JICST-J **

#1 著者 飯島純一, 油田信一, 金山裕
標題(言語) 自立ロボット「山彦」の画像認識(JA)
キーワード 人工知能, 知能ロボット, 画像処理, ハードウェア, テレビジョンカメラ, パターン認識

#2 著者 木島洋一, 安居院猛
標題(言語) 種々の知識データを用いたパターン分類法(JA)
キーワード パターン分類, 人工知能, 知識, パターン認識, 計算機シミュレーション, 特徴抽出

#3 著者 坂井利之, 稲垣耕作, 加藤俊一
標題(言語) CPU:Z8001 パターン理解マルチマイクロプロセッサシステム(JA)
キーワード パターン認識, 人工知能, 並列処理, マイクロプロセッサ, 画像処理, 画像, 文書画像

#4 著者 津川定ゆき
標題(言語) 知能自動車(JA)
キーワード 人工知能, 自動化, パターン認識, 自動車, 走行試験, 道路, 最適化, 操作, 自動操縦, 通路

↳

⑯ RS> OUT SY(K)
+FRS330I 出力番号は 0001

⑰ RS> HISTORY 10

*SHOW EL
*SHOW EL(BN)
*SHOW INDEX
*SHOW TEXT
*SEA KW EQ ショウコウチノウ
.FRS100I 195 件見つかりました。
*AND KW EQ ハターニンシキ
.FRS100I 13 件見つかりました。
*OUT
*SORT KEY(AU)
*OUT EL(AUJ TIJ KWJ)
*OUT SY(K)
.FRS330I 出力番号は 0001

使用例の説明

- ⑪ キーワードとして "人工知能" を持つレコードを検索する。195件見つかった。
- ⑫ ⑪で見つかった195件で、更にキーワードとして、"パターン認識" を持つレコードを検索する。13件見つかった。
- ⑬ OUTPUTコマンドにより、検索結果を表示する。標準出力項目が表示される。
- ⑭ 検索結果を、項目名AU(著者名フリガナ)について昇順にソートする。
- ⑮ OUTPUTコマンドにより、検索結果を項目名AUJ(著者名), TIJ(標題), KWJ(キーワード)についてのみ表示する。ソーティングの効果がみられる。
- ⑯ OUTPUTコマンドにより、検索結果をラインプリンタに出力する。出力は漢字およびカナを含むので出力クラスKを指定する。1行あたりの文字数、日本語文字列の文字の大きさは、②のFAIRSコマンドで指定した値がとられる。なお、これらの値はOUTPUTコマンドでも指定できる。

<例> OUT SY(K) LINESIZE(80) POINT(12)

⑰ HISTORY コマンドにより、検索処理経過を表示する。この場合、10個前のコマンドから直前までのコマンドが表示される。

⑱ RS> BROWSE KW EQ ショウホウケンサク

```
+FRS200I 「KW」項目のインバーテッドファイルより
  語番号      件数  項目値
#00001      2 ショウホウカツトウ
#00002      2 ショウホウキカン
#00003      1 ショウホウキョシユツ
#00004      1 ショウホウキョシユツシヤ
#00005      1 ショウホウクウカン
* #00006     69 ショウホウケンサク
#00007     59 ショウホウケンサクシステム
#00008      5 ショウホウケツン
#00009      5 ショウホウケツンフコウカ
#00010      2 ショウホウケンモテナル
#00011      1 ショウホウコウカイセイト
```

⑲ RS> SEA KW EQ #6 OR #7
+FRS100I 127 件みつかりました。

⑳ RS> SEA TI HAS ニホンゴ
+FRS100I 12 件みつかりました。

㉑ RS> BROWSE JN EQ JOHO

```
+FRS200I 「JN」項目のインバーテッドファイルより
  語番号      件数  項目値
(*)#00012      3 JOHO KANRI
#00013      96 JOHO SHORI
#00014      20 JOHO SHORI GAKKAI BUNSAN SHORI SHISUTEMU KENKYUKAI
#00015      6 JOHO SHORI GAKKAI DENSHI SOCHI SEKKEI GIJUTSU KENK
#00016      11 JOHO SHORI GAKKAI DETA BESU KANRI SHISUTEMU KENKYU
#00017      25 JOHO SHORI GAKKAI IRYO JOHOGAKU KENKYUKAI SHIRYO
#00018      12 JOHO SHORI GAKKAI JINKO CHINO TO TAIWA GIHO KENKYU
#00019      17 JOHO SHORI GAKKAI KEISAN GENGOGAKU KENKYUKAI SHIRY
#00020      11 JOHO SHORI GAKKAI KEISANKI AKITEKUCHA KENKYUKAI SH
#00021      20 JOHO SHORI GAKKAI KEISANKI SHISUTEMU NO KAISEKI TO
#00022      36 JOHO SHORI GAKKAI KONPYUTA, BIJON KENKYUKAI SHIRYO
#00023      14 JOHO SHORI GAKKAI MAIKURO KONPYUTA KENKYUKAI SHIRY
#00024      42 JOHO SHORI GAKKAI RONBUNSHI
#00025      20 JOHO SHORI GAKKAI SOFUTO WEA KOGAKU KENKYUKAI SHIR
#00026     863 JOHO SHORI GAKKAI ZENKOKU TAIKAI KOEN RONBUNSHU
```

㉒ RS> AND JN EQ #26
+FRS100I 4 件みつかりました。

㉓ RS> OUT

** JICST-J **

<p>#1 著者 所属機関 標題 (言語) 雑誌 抄録</p>	<p>西田豊明, 榊原洋三, 堂下修司 京大工 日本語の述語の解析と英訳について (JA) 情報処理学会全国大会講演論文集, VOL. 21ST, P. 825-826, 1980, JPN 日本語において重要な役割を果たすと考えられる述語に焦点を当て解析と意味抽出を行ない、簡単な英文を生成する実験を行なった。パーザとしてLINGOL系のパーザを用い、意味を高階述語論理式で表わす。述語の解析は動詞、アスペクト形式素、「せる」「させる」「れる」「られる」等の助動詞及び終助詞の解析に分けて実施</p>
<p>#2 著者 所属機関 標題 (言語) 雑誌 抄録</p>	<p>畑田のぶ子, 相沢輝昭 NHK総技研 日本語を混成させたデータベース照合言語とその使用実験 (JA) 情報処理学会全国大会講演論文集, VOL. 21ST, P. 1017-1018, 1980, JPN 「集合と写像」を記述する数学流の記法をほぼそのまま借用した、エンドユーザ用データベース照合言語SMLの一部に、日本語を併用することによって、SMLと日本語と両者の長所をともに生かす試み</p>

#3	著者 所属機関 標題(言語) 雑誌 抄録	茂垣真人 東大 日本語の二字組のひん度分布(JA) 情報処理学会全国大会講演論文集, VOL. 21ST, P. 1037-1038, 1980, JPN 日本語文のタッチタイプ式入力における打けんコード確定のため, 日本語の特性について漢テレコード化されたマシンリダブルテキストを用いて解析, そのさいにえられた日本語の二字組のひん度分布と一字のひん度分布との関係について報告
#4	著者 所属機関 標題(言語) 雑誌 抄録	星野雅美 国文学研究資料館 日本語の索引作成システム 検索処理について(JA) 情報処理学会全国大会講演論文集, VOL. 22nd, P. 507-508, 1981, JPN 索引作成システムの蓄積処理は多単位に分割され, 複数の属性が付加された1文を1レコードとして入力し分割単位・属性を1単位とするセルを持ち, 各レコードはこのセルの集まりで表現する方式を採る。本報告では, この蓄積ファイルに対し, セルのリンクに従って順に探索する方式を採り, 単純なものから複雑な検索要求までを処理可能にした

- ②4 RS> END
- ②5 FAIRS> END
+FCA100I FAIRS 終了。
- ②6 READY
LOGOFF
RETURN CODE : 0000
KEQ56470I F1234 LOGGED OFF AT 15:45:03 ON JANUARY 8, 1982+
KEQ54100I SESSION ENDED

使用例の説明

- ⑱ BROWSEコマンドにより, 項目名KW(キーワード)のインバーテッドファイルの索引語を, "情報検索"の前後5つを表示する。語番号, その索引語を含む文献数, 及び索引語が表示される。
- ⑲ キーワードとして, "情報検索", あるいは"情報検索システム"を持つレコードを語番号を用いて検索する。127件見つかった。
- ⑳ 標題に, "日本語"を索引語として含むレコードを検索する。12件見つかった。項目名TI(標題)はTEXT項目なので関係演算子HASを用いる。
- ㉑ BROWSEコマンドにより, 項目名JN(雑誌名)のインバーテッドファイルの索引語で, "情報@ (前方一致)"と等しいものを表示する。語番号, その索引語を含む文献数, 及び索引語が表示される。
- ㉒ ㉑で見つかった12件で, 更に雑誌名として"情報処理学会全国大会講演論文集"を持つレコードを語番号を用いて検索する。4件見つかった。
- ㉓ OUTPUTコマンドにより, 検索結果を表示している。
- ㉔ RSを終了する。
- ㉕ FAIRSを終了する。
- ㉖ TSSを終了する。

4. おわりに

初めに述べたように, この文献データベースのサービスは実験的性格が強い。JICSTのデータそのものが研究用に別の目的で購入したものである。従って, ディスクの事情や予算の関係でいつまでサービスを続けられるかは確約できない。センターとしては, これを契機に, 利用者の日本語データベースへの関心が高まることを期待したい。

参考文献

1. 二村, 鬼塚, 高木, 松尾 FAIRSによる文献検索, 九大大型計算機センター広報, 12, 4, 1979, 319-350.
2. 計算機マニュアル FACOM OSM FAIRS-I コマンド文法書, 情報検索編(70AR-3181-1), 富士通株

付録1. 言語コード, 発行国コード

言語コード	言 語	発行国コード	発 行 国
DE	ドイツ語	AUS	オーストラリア
EN	英語	CAN	カナダ
JA	日本語	DEU	西ドイツ
RU	ロシア語	FRA	フランス
		GBR	イギリス
		JPN	日本
		NLD	オランダ
		SUN	ソビエト連邦
		USA	アメリカ合衆国