

PATOLIS-INPADOC オンラインシステム

斎藤 誠, 中池伸一 (日本特許情報センター)

1. はじめに

PATOLIS (PATIENT-ONLINE-INFORMATION-SYSTEM) - INPADOC (INTERNATIONAL PATENT DOCUMENTATION CENTER) オンラインシステムは、アメリカ・西ドイツ・日本・イギリス・フランス・ソビエト等の49カ国とヨーロッパ特許庁 (EPD) ・世界的所有権機構の2機関の特許情報を、日本国内にサービスしている特許情報検索システムであります。

INPADOCにおける世界各国の特許の累積件数は、昭和58年8月現在一千万件の大台を突破して1006万件におよび、年間70万件余の追加が発生しています。本システムは、このデータを汎用DBMSであるADABASを用いてデータベース化し、昭和57年12月からサービスを開始しました。

本システムの概要を、

- 1) PATOLISシステム概要
- 2) PATOLIS-INPADOC概要
- 3) データベース構成
- 4) 検索システム

の順で紹介する。

2. PATOLISシステム概要

2.1 システム概要

PATOLISは、日本特奥・日本意匠・日本商標・INPADOCの4データベースを、特定回線漢字端末・公衆漢字端末・公衆カナ端末を通じて、日本全国のユーザーに提供しております。ユーザー(登録パスワード)数は、9百余におよび特定回線漢字端末62台、公衆漢字端末460台、公衆カナ端末390台の内訳となっています。図1にシステム概念を示す。

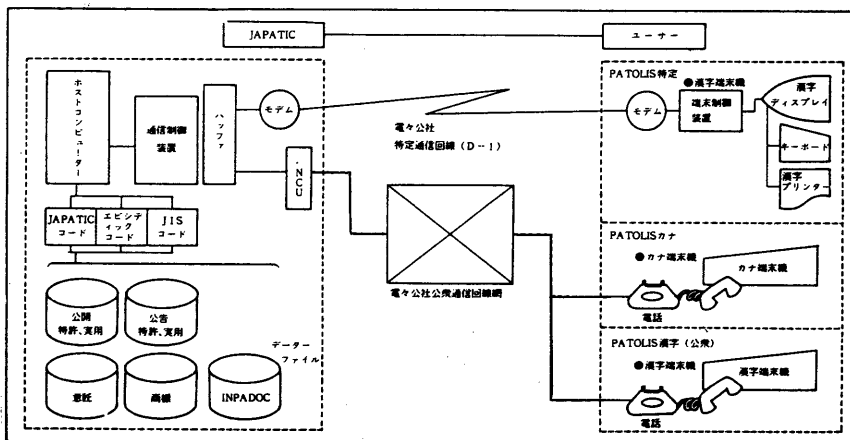


図1 システム概念図 =1=

昭和54年4月に日本で初めての漢字オンラインシステムとして産声を上げ、日本特許・アメリカ特許の情報検索サービスを開始しました。昭和55年10月には、公衆カナ端末のサポートを行う為にオンラインモードからTSSモードにシステムを全面更改し、合わせてOSをEDOS/MSDからVDS3に変更し、さらに日本意匠・日本商標・パテントファミリー(同一発明を各国に出願している場合、その出願国を検索する)のデータベースを追加しサービスを拡充をしました。昭和58年1月には公衆漢字端末によるサービスも開始しました。

2.2 ハードウェア構成

当センターのハードウェア機器構成を図2に示す。ホストコンピュータは、HITAC・M200H(10MB)、M180(6MB)、M240H(6MB)の3CPU、使用オペレーティングシステムは、VDS3、VDS3-SPI/TSS3およびEDOS/MSDを使用しています。本年12月には、M180をM280H(16MB)にリプレースする予定となっています。

DISCは、H-8589-11(200MB)×16、H-8576(635MB)×14、H-8598(1260MB)×16、およびH-8523-11(MSS 34.8GB)を使用しています。

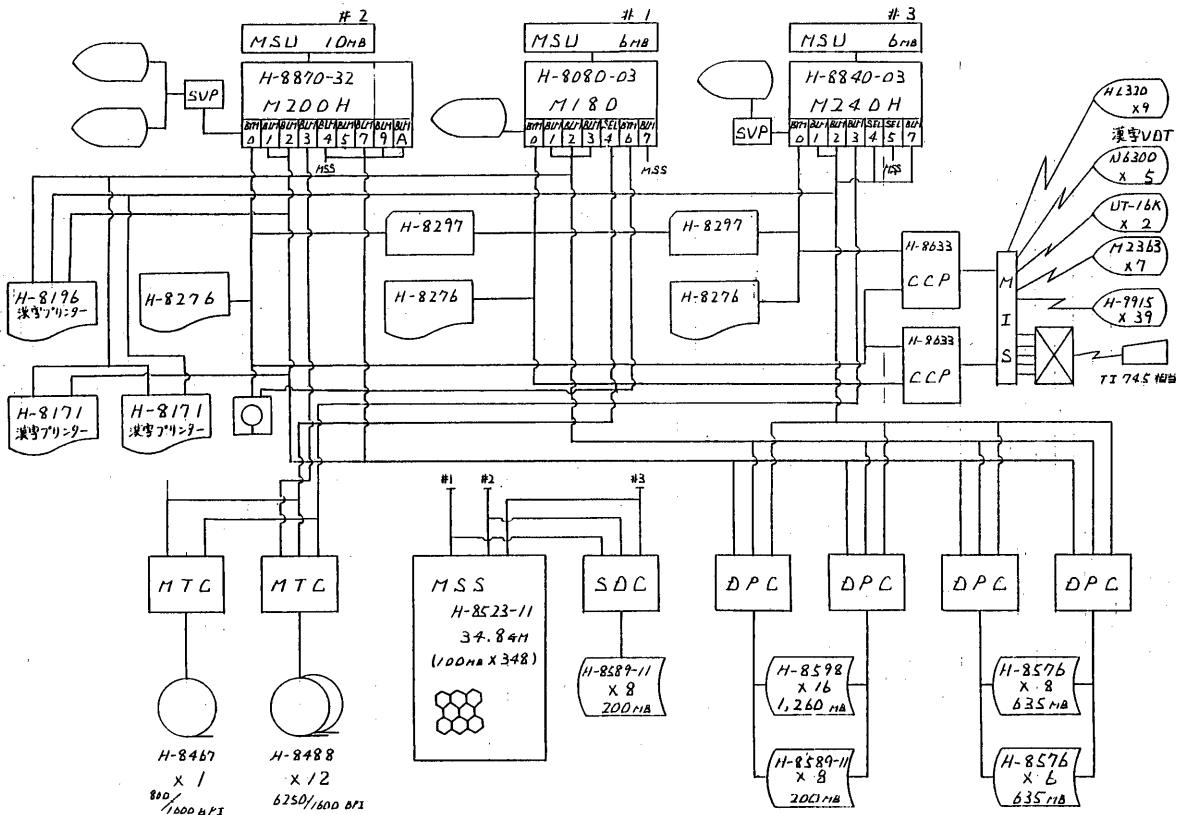


図2 ハードウェア構成

通信制御処理装置は、H-8b33X2台。漢字プリンターは、H-8196-20, H-8171X2台です。

PATOLISは、MZDDHをメイン機にM180をバックアップ機として運用しています。データファイルはH-8598X6, H-8576X6, CCPは1台を使用し、オフライン出力として夜間バッチ処理により漢字オフライン回答にH-8196-20, 図形オフライン回答にH-8171を使用しています。

2.3 ソフトウェア構成

PATOLISのソフトウェア構成を図3に示す。オペレーティングシステムはVDS3-SPI/JSS3, 通信制御にBTAM・BCP, TSSはTIOP IIを使用しています。

各プログラムは、TSSのCP(コマンドプロセッサ)として動作し、日本特奥・INPADOCプログラムはノモリ効率上からリエントラントコーディングとなっておりMPLPA(MODIFIED LINK PACKAGE AREA)に常駐させています。

データファイルは、日本特奥・日本意匠・日本商標マスタ、インバーテッドファイル、各番号テーブルはVSAMファイル、INPADOCファイルはADABASとなっています。

本年11月には、日本商標がADABASファイル化されます。

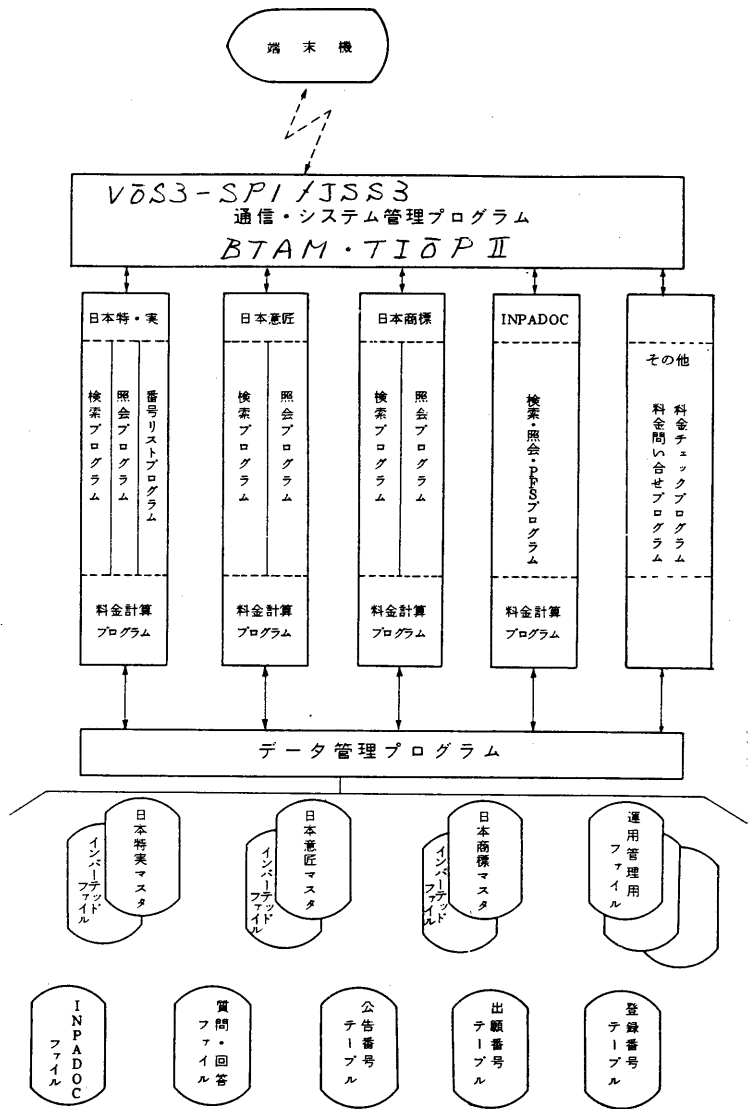


図3 ソフトウェア構成

2.4 蓄積データ件数

8月現在における蓄積データ件数は、特許公開1,764,095件、特許公告981,180件、実用新案公開1,811,370件、実用新案公告988,260件、発明430,749件、商標1,031,564件、INPADOC10,063,707件、合計17,052,925件となっています。

3. PATOLIS-INPADOC概要

3.1 システム概要

PATOLIS-INPADOCは、日本特許情報センターとINPADOCの協定が成立し、INPADOCの保有する全データの入手が可能となり、アメリカ特許・パテントファミリーの両データベースを包含し、なおかつ各国特許の検索・照会・パテントファミリーサーチを行うシステムとして昭和57年4月に開発に着手し、同年12月にサービスを開始しました。

開発期間・ソフトウェアメンテナンス・データ量等の諸条件の検討を行い、従来データベースはVSAMによるインバーテッドインデックスを用いた検索システムを当センターで開発し使用しているが、個々のデータベース毎に存在し、かつその規模も膨大である点を考慮し商用DBMSの適用に踏み切り、かねてよりベンチマークテストを行っていたADABASの導入を決定しました。

本システムは、英語ベースの会話形式コマンドによる情報検索システムで、

検索条件を満足する集合を作成するFINDコマンド、集合と集合の論理演算を行なうBOLDOLEコマンド、集合を解放するRELEASEコマンド、端末に検索結果を出力するLISTコマンド、オフライン出力を行うPRINTコマンド、対応特許調査（パテントファミリーサーチ）を行うPFSコマンド

| 項番 | コマンド種別 | コマンド | | | 機能概要 |
|----|--------|---------|-----|----------------|---|
| | | コマンド名 | 省略形 | サブコマンド | |
| 1 | 検索 | FIND | F | | 検索条件を満足する集合を作成する。 集合と集合の論理演算を行い新しい集合を作成する。 集合のソートを行う。 集合の解放を行う。 端末に出力を行う。 オフラインプリントを行う。 対応特許調査（パテントファミリー照会）を行い、端末に出力する。 検索処理を終了する。 |
| | | BOOLE | B | | |
| | | SORT | - | | |
| | | RELEASE | REL | | |
| | | LIST | L | | |
| | | PRINT | - | | |
| | | PFS | - | | |
| 2 | サポート | STOP | - | | |
| | | DISPLAY | D | CLUSTER (省略形C) | 集合の件数、作成したコマンドの表示を行う。 |
| | | | | FIELD (省略形F) | フィールドネーム、フィールド名称の表示を行う。 |
| | | | | TIME (省略形T) | 日付、時間、コネクトタイム、処理時間の表示を行う。 |
| | | NEWS | - | | メッセージを表示する。 |
| | | CHARGE | - | | 当該年度の各月使用料金を表示する。 |

表1 コマンド一覧

等のコマンド群より構成されています。表1にコマンド一覧を示す。

サービスしている国は、出願人・発明の名称のフリーワード(文字)検索が可能な特定国—アメリカ、イギリス、西ドイツ、フランス、ソビエト、PCT、EPD、韓国、の6カ国2機関—と、その他の一般国に大別されます。

3.2 データ項目

INPADOCのデータは、各国特許庁が送付した磁気テープもしくは各国特許庁公報から入力したデータをINPADOCが収集し、毎週当センターに航空便により送付してきます。そのデータは書誌と呼ばれる公報発行国・公報種別・公報番号・公報発行日・出願番号・発明の名称・出願人等の項目より構成されています。

INPADOCデータに関しては、INPADOCはデータベースプロデューサー(作成者)、当センターはディストリビューター(提供者)であるという考えに基づき、送付されたデータを用いて回若千の出願人名称統一処理を行った後、重複削除・レコード追加のデータベース更新処理を実行しています。

14万件の内訳は、日本2ノクろ件・西ドイツノ68万件・アメリカノ04万件・フランス82万件・イギリス58万件・ソビエト55万件となっており、この6カ国で684万件約68%、特定国においては48ノ万件約48%となっています。検索サービス面では考えますと特定国で全世界の半分をカバーし、日本特許サービスと合わせて利用す

| | フィールド名 | 検索式指定項目 | 検索項目 | 前方一致\$マーク指定 | 条件記号 | リスト項目 | 最大指定長 |
|----|--------|------------------------|------|-------------|-----------------|-------|-------|
| 1 | CC | 公報発行国 | ○ | × | = | ○ | 2 |
| 2 | KD | 公報の種類 | ○ | × | = | ○ | 2 |
| 3 | DN | 公報番号 | ○ | × | = | ○ | 8 |
| 4 | PD | 公報発行日 | ○ | × | =, >, >=, <, <= | ○ | 6 |
| 5 | AN | 出願番号 | ○ | × | = | ○ | 12 |
| 6 | AD | 出願日 | × | / | / | ○ | / |
| 7 | KA | 出願の種類 | × | / | / | ○ | / |
| 8 | IC | 国際特許分類(IPC) | ○ | ○ | = | ○ | 13 |
| 9 | NC | 自国分類(ナショナルクラス) | ○ | ○ | = | ○ | 13 |
| 10 | RC | 優先権主張国 | × | / | / | ○ | / |
| 11 | RN | 優先権主張番号 | × | / | / | ○ | / |
| 12 | RD | 優先権主張日 | × | / | / | ○ | / |
| 13 | KR | 優先権主張種別 | × | / | / | ○ | / |
| 14 | TI | 発明の名称 | ○ | ○ | = | ○ | 30 |
| 15 | AP | 出願人 | ○ | ○ | = | ○ | 30 |
| 16 | IN | 発明者 | × | / | / | ○ | / |
| 17 | CA | ケミカルアブストラクト番号 | × | / | / | ○ | / |
| 18 | DC | PCT指定国 | × | / | / | ○ | / |
| 19 | ZM | 公報発行国+公報発行年月 | ○ | × | = | × | 6 |
| 20 | ZD | 公報発行国+公報番号 | ○ | × | = | × | 10 |
| 21 | ZA | 公報発行国+出願番号 | ○ | × | = | × | 14 |
| 22 | ZP | 公報発行国+出願人 | ○ | ○ | = | × | 32 |
| 23 | ZI | 公報発行国+IPCサブクラス | ○ | × | = | × | 6 |
| 24 | AE | 出願番号+出願日 | × | / | / | ○ | / |
| 25 | DE | 公報発行国+公報の種類+公報番号+公報発行日 | × | / | / | ○ | / |
| 26 | RE | 優先権主張国+優先権番号+優先権主張日 | × | / | / | ○ | / |

表2 データ項目一覧

る70%以上の特許情報が検索できることとなります。尚、日本特許庁は特許だけを送付し、実用新案は送付していません。日本の特許だけでINPAの70%のデータの20%をしめています。つまり世界の特許の5件に1件は日本の特許となり、実用新案を含めて考えると実に3件に1件は、日本ということになります。

表3にデータ項目一覧を示す。このデータ項目は、検索項目(ディスクリプタ)・リスト項目・SORT項目に大別されます。検索項目はFINDコマンドで検索可能なフィールドで、公報発行国・公報種別・出願人等の通常検索項目と公報発行国+公報発行年月等のパターン化した高速検索項目に別かれます。リスト項目はLISTコマンドで出力可能なフィールドで、通常リスト項目と、AE・DE・REの編集出力項目に別かれます。SORT項目は、公報発行国・公報種別・公報番号・出願番号・公報発行日の5項目となっています。

3.3 コマンド説明

① FINDコマンド

データベースから、検索式を満足する集合を作成する。

FIND @集合名 検索式〔〔演算子 検索式〕...〕。

検索式 :
$$\left\{ \begin{array}{l} \text{フィールドネーム} \\ \left\{ \begin{array}{l} = \text{検索条件} \\ > \\ > = (= >) \\ < \\ < = (= <) \end{array} \right\} \text{値} \end{array} \right\}$$

@集合名

検索条件 :
$$\left\{ \begin{array}{l} \text{値} \\ \text{値}_1 : \text{値}_2 \text{ (値}_1 \text{ から値}_2 \text{ まで)} \\ \text{値} \# \text{ (前方一致)} \\ \text{"空白を含む値"} \\ \text{(値}_1 \text{ 演算子 値}_2 \text{ ...)} \end{array} \right\}$$

② BODLEコマンド

FIND・BODLEコマンドで作成した集合どおしの論理演算を行う時に使用し、新しい集合を作成します。

BODLE @集合名 @集合名 演算子 @集合名
(演算子 @集合名 ...)

③ SORTコマンド

検索された集合に指定したソートキーに従ったソートを実行する。

SORT @集合名 @ソートする集合名 KEY= .

④ RELEASEコマンド

FIND・BDDLE・SORTコマンドで作成した集合を解放します。集合は16個まで保存できます。

RELEASE @集合名 (@集合名) .

⑤ LISTコマンド

作成された集合を端末に出力します。出力項目および件数範囲を指定して、必要な情報のみを出力することができます。

LIST @集合名 (ITEM=フィールド名{ , ... }
 { m₁
 { m₂ - m₃ } } .

⑥ PRINTコマンド

作成された集合をオフライン出力します。

PRINT @集合名 .

⑦ PFSコマンド

出願番号もしくは優先権主張番号と、国コードを指定しパテントファミリー照会をおこない端末に出力します。

PFS NO="YYXX99999999", CC=XX .

⑧ STOPコマンド

サービスを終了します。

STOP .

⑨ DISPLAYコマンド

各種の状態表示をします。

DISPLAY { CLUSTER
 { FIELD
 { TIME } } .

⑩ NEWSコマンド

JAPATICからのメッセージを表示します。

NEWS .

① CHARGEコマンド

当該年度の使用料金を月毎に表示します。

CHARGE.

4. データベース構成

大量データにおいてのデータベース構成としては、単独ファイルか複数のファイルに分割する方法があり、分割ファイルが一般的に推奨されています。本システムは次に示す理由により単独ファイルでデータベースを構成しています。

特許情報検索においては、ほとんど全ての項目が検索項目となり得るといっても過言ではなく、サービス性等から可能な限り検索項目を設定しています。検索においてもユーザーの検索条件(検索項目の組合せ)は千差万別でありパターン化した検索では対応できませんし、ユーザーニーズを満足することはできません。また、ファイルを分割したとしてもユーザーからは見かけ上1ファイルであるかの様にしなければなりません。

表3に単独ファイルと分割ファイルの比較を示す。

| | 単独ファイル | 分割ファイル |
|----------|--|------------------------------------|
| 検索ロジック | DBMS | ユーザー DBMS |
| 中間結果ファイル | DBMS | ユーザー |
| 汎用性 | DBMSに依存し、ファイル毎に異なるシステムを構築する。ある程度の汎用化は可能。 | DB毎に個別に構築し、分割数によっても複雑化する。汎用化はできない。 |
| システム容易度 | 比較的簡単 | 複雑 |
| システム規模 | 小 | 大 |
| エンドユーザー | | 見かけ上1ファイル |
| 検索スピード | 全体検索 早い 収束インポート 遅い | 分割条件検索 遅い 早い |
| DB更新 | 遅い | 早い |
| 運用 | 単純 | 複雑 |

表3 単独・分割ファイル比較

サービス開始後、検索レスポンスタイムに次の特徴が発見された。

- 1) 公告番号・出願番号等の一意性の検索項目、出願人・発明の名称等の分散性の検索項目の場合、大容量でもスピードに問題はない。
- 2) 発行国・公報の種類等の収束性の検索項目の場合、遅い。
- 3) 2)項目と複合した検索では、極端に遅くなる。

そこで、3)項目のパターンで特に使用頻度の高い発行国と他の検索項目の組合せのインポート(スーパーディスクリプタ)を作成し、本年4月から提供しました。

5. 検索システム

本システムの開発時、A²D²A²B²A²Sを使用した新商標オンラインシステムが開発中であり、また将来開発するシステムも極力A²D²A²B²A²Sを適用する方針がありましたので、システム毎に同一の演算子・優先順位を持った検索ロジックを作成することはシステム開発上効率的ではありません。そこでA²D²A²B²A²S汎用検索ルーチンと呼んでいるA²D²A²B²A²Sデータベース共通検索システムを開発しました。

この検索システムは、オンラインシステム毎に異なる端末インターフェイス(コマンド形式、メニュー画面方式等)の検索質問式を、各々のシステムで表4に示す可変長形式のJAPATIC A²D²A²B²A²S検索DMLに変換し、A²D²A²B²A²S汎用検索ルーチンをコールします。この検索DMLから

- 1) 検索DML論理チェック
- 2) 逆ポーランド変換
- 3) 検索論理演算

を行い、結果を返します。

検索論理演算処理において、検索時間の短縮をはかる為

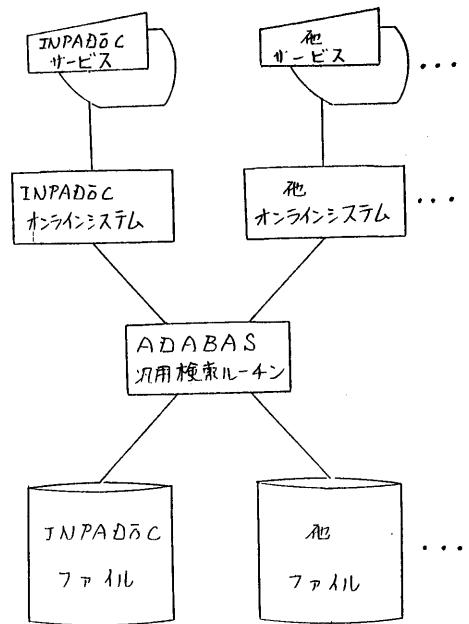


図4 検索システム概念

| 項番 | 名称 | D M L エ レ メ ン ト | | | | | | | | | |
|------|--------|---|---------|-------|-------|---------|-------|-------|---|---|---|
| 1 | ターム | <table border="1" style="width: 100%;"> <tr> <td style="width: 15%;">レングス</td> <td style="width: 10%; text-align: center;">T</td> <td style="width: 10%;">予備</td> <td style="width: 15%;">検索フィールド</td> <td style="width: 10%;">条件フラグ</td> <td style="width: 10%;">値レングス</td> <td style="width: 10%;">値</td> <td style="width: 10%;">値</td> <td style="width: 10%;">値</td> </tr> </table> <p style="text-align: right; margin-right: 20px;">EQ, GT, GE, LT, LE, THRU</p> | レングス | T | 予備 | 検索フィールド | 条件フラグ | 値レングス | 値 | 値 | 値 |
| レングス | T | 予備 | 検索フィールド | 条件フラグ | 値レングス | 値 | 値 | 値 | | | |
| 2 | 演算子 | <table border="1" style="width: 100%;"> <tr> <td style="width: 15%;">レングス</td> <td style="width: 10%; text-align: center;">O</td> <td style="width: 10%;">演算子</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> </table> <p style="text-align: right; margin-right: 20px;">(,), *, +, ≠</p> | レングス | O | 演算子 | | | | | | |
| レングス | O | 演算子 | | | | | | | | | |
| 3 | 中間結果ID | <table border="1" style="width: 100%;"> <tr> <td style="width: 15%;">レングス</td> <td style="width: 10%; text-align: center;">I</td> <td style="width: 10%;">内部後用</td> <td style="width: 10%;">コマンドID</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> </table> | レングス | I | 内部後用 | コマンドID | | | | | |
| レングス | I | 内部後用 | コマンドID | | | | | | | | |
| 4 | 終了サイン | <table border="1" style="width: 100%;"> <tr> <td style="width: 15%;">レングス</td> <td style="width: 10%; text-align: center;">E</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> </table> | レングス | E | | | | | | | |
| レングス | E | | | | | | | | | | |

表4 JAPATIC ADABAS 検索 DML

- 1) AND条件の連続(コマンドIで, THRU条件を含む)
- 2) 同一フィールドのOR条件の連続(THRU条件を含む)
- 3) 1個以上の同一フィールドOR条件後の1個以上のAND条件の場合; 1回の検索コマンド(S1)で検索を実行し最適化を行っています。

6. おわりに

PATROLIS-INPADOCオンラインシステムの概要を紹介した。本システムはサービス開始後2回改造し、現在INDEEXコマンド新設等3期改造中であり、データ整備を含めてシステムの改善に務めサービス向上を目指しております。