

## 電子ジャーナル管理データベースのデータベースモデルと入力仕様

### はじめに

NACSIS-CAT は、「全国の学術情報資源を集約したい」との考えで発展してきた総合目録データベースで、電子媒体である電子ジャーナルにおいても「どのタイトルをどの参加組織で閲覧可能か？」は何らかの形で確認できることが望ましいと考えられてきた。しかし、電子ジャーナル担当者からは、NACSIS-CAT の電子ジャーナルの扱いの規則が実際の業務に適合しない、という意見が多数寄せられていることから、今回の研修では、理想的なデータベースモデルの作成を試みる。

電子ジャーナルは NACSIS-CAT での基本的な書誌作成単位である物理実体がなく、従来の情報源からの書誌作成という考え方では限界があるように思われた。電子媒体、とりわけインターネットの世界では様々な状況の変化が桁違いに速く、近年は電子ジャーナルを様々なパッケージとして購入するという傾向があることから、従来の物理単位という発想を変え、「パッケージ中のタイトルを年間契約でアクセス権を購入するもの」という観点からデータベースモデルを作成し考察した。

考察の進め方として、第一に「電子ジャーナル担当者が求める理想のデータベースモデル」の提示、次に「理想実現のための具体的なシステム内容」、続けて「現在ある NACSIS-CAT データベース資源で理想をどのように表現するか」、最後に「NACSIS-Webcat のデータ更新・移行の工夫」の4段階での考察として、極力具体的にモデルを考察・作成したつもりである。今回はデータベース実務研修内での限られた時間内、それも雑誌受入・目録担当者の4名のグループという限られた人員での検討となったため、様々な事例に対応できない部分も多数見受けられると思われるが、今後の電子媒体の取り扱いについて具体的で実用的なモデルを限られた条件の中で提案することにした。

### 電子ジャーナル担当者が求める理想のデータベースモデル

データベースモデルはインターネット世界の中に無数あり、常に現在進行であるので、理想を論ずるにも戸惑うような部分がある。しかし、「全国の学術情報資源を集約する」、もしくは「自分の求めている情報を早急に探せる」ものを作る、といった限られた目的を実現する場合には、ある程度具体的な手段を決定してから考察を進めるべきだ、と思われる。よって、目的と手段を明確にする作業から始めたい。

自分の目的とする情報を取得するには、まず適切なデータを作成してそれが「自分の求めている情報と一致する」ことで、初めて情報を手に入れる第一段階となる。現在の進化した検索システム・データベース検索では、検索システムの優劣よりも、むしろ内容の部分、データの中身の真価が問われるといつてよい。よって、データベースの項目内容が洗練されていない<sup>1)</sup>と、全く役に立た

ないものを作ってしまった、という状態に陥る危険性は十分に考えられる。今後もコンピュータを利用してデータを検索するという作業は継続される、と決定づけられるので、項目は少なくとも情報を検索するという目的を確実に実現できる具体的モデルを考案した。

今回の考察では電子媒体全般の中の「電子ジャーナル」を中核として取り上げ、データベースモデルを考案した。電子ジャーナルは物理的実体がないことが前提であるが、基本的なモデルとしては次の状態を表現できるようにした。

ある年に、  
あるパッケージで、  
あるタイトルを、  
ある接続先を利用して、  
どの図書館が、  
どれだけのモノを、(いつからいつまで) 提供されているか(アクセス権があるか)、

また、ホームページでのリンクとしては、パッケージのメインページから検索してタイトルページを表示、そこから各号の論文に到達するという基本モデルを元に、書誌階層としてパッケージのメインページからタイトルページへのリンクを親書誌・子書誌の関係で表現した。

パッケージのメインページ      タイトルページ      各号論文

上記のイメージをデータベースモデルで表現するために、従来の「物理単位での目録作業」、  
「雑誌なので初号を基準に継続を重視する」という発想を変え、「パッケージ中のタイトルを年間契約でアクセス権を購入するもの」という観点からデータベースモデルを作成した。

図書館閲覧担当者の理想としては「google のような www ページ検索エンジンと、NACSIS-Webcat の書誌詳細を検索できるツールの融合」で、google のように検索目的語の入力部分は1つになっていることが条件である。1つの検索キーワードでありとあらゆる情報を得ようとするわけであるが、NIIの現在運用中のシステムでは、GeNii がそれに相当する。ただし、GeNii は、データベースを横断検索するツールと位置づけられ、個別のデータベースの理想形とは考えられない。よって、電子媒体の「全国の学術情報資源を集約する」という目的からすると、横断検索だけでは不十分で、実際の個々のデータベース、横断検索される側の個別のデータベースがうまく機能していない状態では、十分活用できるとはいいがたい。次章で、具体的なデータベースモデルを提案する。

. 理想実現のための具体的なシステム内容 - 年ファイルの存在意義 -

具体的なデータベースモデルとして、次の項目とリンク構造が必要である、との結論に至った。

## 目的として

- ・「全国の学術情報資源を集約する」ものであること。
- ・実務を担う図書館電子ジャーナル担当者に負担がかからないもの。

以上2点を同時に実現できるモデルを考案した。運用は次のとおり。

## 運用モデル

書誌データは電子ジャーナル出版社に協力してもらい、翌年の契約予定タイトルの書誌をNIIで一括登録する<sup>2)</sup>(8～9月頃)。ここでの登録は購入ベースのもので、パッケージ契約によるものが多くなるとされる。

書誌データの入力完了後、一括登録した全タイトルのリストを各参加組織館でNIIサーバーからダウンロードする、などの方法で配付する。

参加組織館では、電子ジャーナルのアクセス権を購入するタイトルに FANO などの図書館の識別コード、ILLが可能な場合はその図書館からアクセス可能な年次・巻次を記入してデータを作成する。電子ジャーナルの場合、遡及入力などで、契約期間内に閲覧可能年次・巻号が変更される場合があるので、年次・巻次は特定できない。よって入力必須事項としない。

参加組織館から、アクセス権保有データを一括で登録する、もしくはそのままNIIにデータを送りNIIが登録する。この作業は11月～12月中に行い、1月からの運用(4月からも可)に備える。

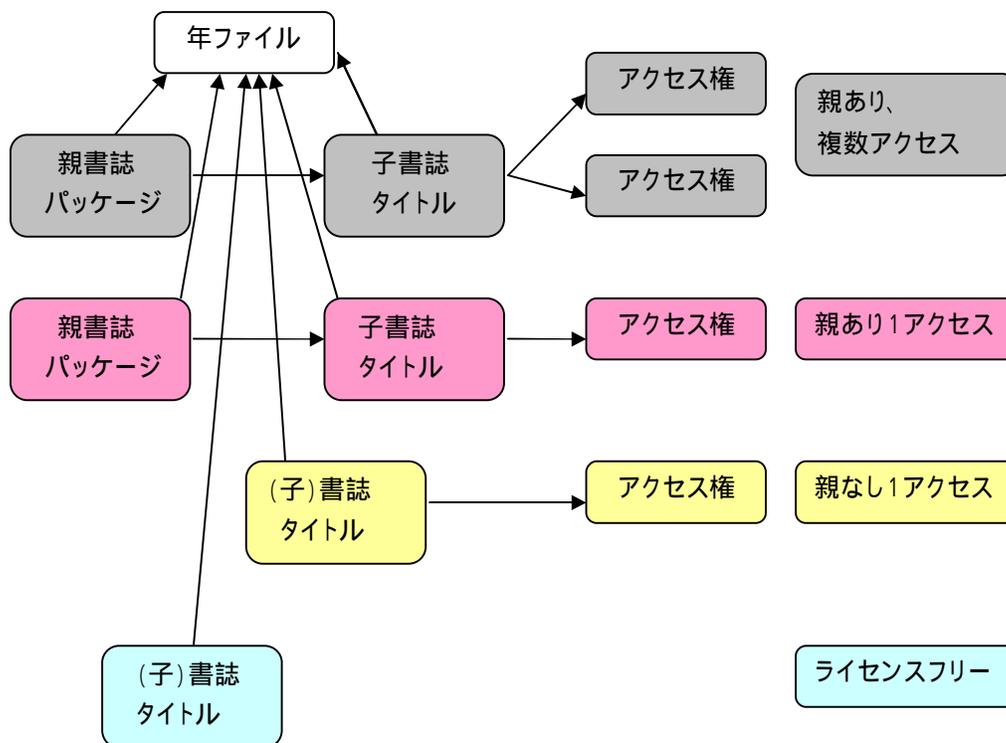
ライセンスフリーの電子ジャーナルも資源共有という意味合いから、書誌データのみを登録して、検索対象とする。

書誌データ、一括登録したアクセス権保有データのメンテナンスは基本的に所定のNIIホームページから行う。

検索は「Google」や「全国漢籍データベース」<sup>3)</sup>のように検索項目を入力する部分は1箇所にする。購入年を指定するためにもう一つ入力箇所を設け、利用可(現在(今年)利用できる電子ジャーナル、購入年(今年)のもの)と9999を抽出)を検索できるようにチェックボックスも設ける。



リンク構成図 書誌ファイルはすべて年ファイルとリンクする。



参加組織ファイルはアクセス権にそれぞれリンクされる。

### 検索表示(詳細)例

<[AA00000111](#)>

契約年: 2005

TR: Journal of thermal stresses (1521074X)

PUB: Taylor and Francis

PTBL: Taylor and Francis 2005<[AA00000004](#)>

URL: <http://journalonline.tandf.co.uk/openurl.asp?genre=journal&eissn=1521-074X>

URL: <http://www.ingentaconnect.com/content/tandf/uths>

所蔵図書館 3

大 図

大

大 電子 130-159<1998-2005> <http://www.ingentaconnect.com/content/tandf/uths/111>

以上が理想の形であり、この形であると、GeNii などからの横断検索も可能と思われる。

#### ファイル構造

##### 年ファイル

年ファイルは「ある年」を表現する。事実上、契約年を表現する。すべての書誌ファイルがリンクする。

YID=\*\*\*\*\*

YEAR=2005

年ファイルは購入年を表現する。書誌データ内に入力してもよいが、リンクすることにより、購入年での検索時に高速化が期待できる。また、購入年次を限定するためにはリンクさせておくことが望ましい。項目中のゴジック体はリンク項目としても利用する。ライセンスフリーのものは、「いつまでも利用できる」という意味で、「9999」を入力する。

##### 参加組織ファイル

参加組織ファイルは所属する図書館の名称などを表示する。アクセス権ファイルにリンクする。

FID=+++++

FANO=FA000000

NAME= 大学図書館

##### 書誌ファイル

##### 親書誌

親書誌は基本的に「あるパッケージ」を表現する。

BID=BA00000002

YEAR=2005

GMD=w

SMD=r

TTL=eng

TXTL=eng

ISSN=

XISSN=

TR=Springer Link

PUB=  
IDENT=<http://link.springer.de/link/service/journals/>  
NOTE=Access: via WWW  
PTBL=

#### 子書誌

子書誌は基本的にパッケージ内のタイトル、「あるタイトルを」「ある接続先を利用して」を表現する。「ある接続先」は IDENT に入力し、複数の接続先に対応するとともに所蔵館ごとの複数の接続も表現する。また、パッケージではなく雑誌タイトル単体で購入する場合は、子書誌のみで書誌情報を構成する。

BID=BA00000001  
GMD=w  
SMD=r  
YEAR=2005  
TTLL=eng  
TXTL=eng  
ISSN=14240661  
XISSN=14240660 ……繰返可  
TR=Annales Henri Poincare  
PUB=Basel, Switzerland : Birkhauser ……繰返可  
IDENT=<http://link.springer.de/link/service/journals/00023/index.htm> ……繰返可  
NOTE=Access: via WWW ……繰返可  
PTBL=Springer Link<BA00000002> ……繰返可

以上のようなレイアウトで、書誌情報を表現する。ライセンスフリーの電子ジャーナルは(無料のもの)はアクセス権が発生しないので、書誌のみ登録し、「このDBで検索し、書誌情報を確認、URLがわかる」、という具合になる。親書誌と子書誌の関係で書誌は NACSIS-CAT 図書DBのように2階層として表現し、親書誌にはパッケージの代表URLを記録する。有料無料の区分は YEAR が「9999」になっているものを基本的に無料として解釈する。逆に YEAR が「9999」になっているものは、年度更新を行わないものと解釈されるので、アクセス先などの変更時には手作業でのデータ修正が考えられる。

#### アクセス権保有ファイル

アクセス権保有ファイルは「ある接続先を利用して」、「どの図書館が」、「どれだけのモノを(いつからいつまで)」、提供されているか表現する。年次・巻号は契約途中からアクセスできるタイトルが変更される場合もあり、加えてILLの利用不可の場合も多いと考えられるので、年次・巻号は入力必須としない。

ID=CD00000001  
BID=BA00000001  
IDENT=<http://link.springer.de/link/service/journals/00023/index.htm>  
LOC=図  
FANO=FA000000  
RGTN=11111            ……空値可、繰返可  
HLYR=2000-2005       ……空値可、繰返可  
HLV=45-50            ……空値可、繰返可  
CLN=請求記号         ……空値可、繰返可  
CPYNT=ILL可          ……空値可、繰返可  
LDF=図書館定義  
LTR=ローカルトレーシング

アクセス権保有ファイルは基本的に書誌ファイルとは BID と IDENT でリンクすることになるが、個々の参加組織の接続先の URL が、書誌の URL とおなじものでないといけない、という仕様にせず複数持つ仕様とし、各館で異なるものでも登録できるようにしたい。その場合は LDF フィールドなどを利用する。しかし、大多数は書誌の URL で問題ないと思われるので、IDENT でもリンクが可能と判断した。また、複数の接続先で同じものを利用できる場合は、複数のアクセス可能な URL を登録できるようにファイル構成を工夫した。

図書館でアクセスできるタイトルに、ILL が可能な場合はその図書館からアクセス可能な年次・巻次を記入してデータを作成する。電子ジャーナルの場合、遡及入力などで、契約期間内にアクセス可能年次・巻号が変更される場合があるので、年次・巻次は特定できない。

リンクモデルとしては、アクセス権保有ファイルに URL を書誌とのリンク項目として利用するところが特長であるが、出版社の都合で急に接続先が変わるといった状況も考えられる。アクセス先は適切に接続しているかを自動的に点検する仕組みを作り、変更があったものは書誌ファイルを自動的に修正する。また、書誌の IDENT を修正し、データを確定した時点で、「その書誌に所属するアクセス権所有ファイルの IDENT も同時に修正後のものに更新する」、というバッチプログラムが走るように設計する。この処理で、リンク不完全なアクセス権保有データがなくなると思われる。

年更新時(11月から12月)には各図書館からデータを提出する場合、次の項目の Excel シート(CSVファイル)に入力する。項目例を次に挙げる。

BID=BA00000001  
YEAR=2005  
GMD=w  
SMD=r

TLL=eng  
TXTL=eng  
ISSN=14240661  
XISSN=14240660  
TR=Annales Henri Poincare  
PUB=Basel, Switzerland : Birkhauser  
IDENT=<http://link.springer.de/link/service/journals/00023/index.htm>  
PTBL=Springer Link<BA00000002>  
LOC=  
FANO=  
RGTN=  
HLYR=  
HLV=  
CLN=  
CPYNT=  
LDF=  
LTR=

書式の関係で縦に記述したが、この項目順で横展開のデータとなっており、ゴジック体の部分に参加組織館はデータを入力することになる。ちなみに、参加組織内で閲覧可能なタイトルに FANO を入力すると、このタイトルに対して所定の 1 年間は「アクセス権あり」というデータになる。Excel のデータなので、購入し(購入予定で)、アクセス可能なタイトルデータのみ自館の FANO をコピーして保存する。最後にファイルを NII に提出という手順である。

・現在ある NACSIS-CAT のデータベース資源で理想をどのように表現するか  
- NACSIS-CAT 図書 DB での表現 -

理想モデルは一つの理想であるが、現実的に、現在の NACSIS-CAT のデータベースレイアウトや、NACSIS-Webcat データベースで、できる限り理想に近づけるには、発想の転換が不可欠であろうと思われる。現在のシステムが、「どのような目的で作成されたか」を考えるよりも、発想が異なっても、「今ある DB が何に使えるのか」、を突き詰めることにした。提案としては次のようになる。

電子ジャーナルの書誌情報は、原則的に「最新号を基準として書誌を作成する」<sup>4)</sup>という概念が浸透し、定着しつつあると思われる。最新号を基準にすることは、現在の NACSIS-CAT 雑誌 DB の原則である「初号主義」における「初号からの継続性」を優先するものとは異なり、電子ジャーナルの「最新号優先」の考え方にそぐわない。また、「パッケージ中のタイトルを年間契約でアクセス権を購入するもの」という発想では、NACSIS-CAT 雑誌 DB に書誌階層、親書誌・子書誌という概念がないことから、パッケージというものを表現できないと考えた。よって、NACSIS-CAT 雑誌 DB に電子

ジャーナル書誌・所蔵データを強引にデータ登録して「全国の学術情報資源を集約する」ことは一度あきらめ、違う手段を考えることになった。

次の考え方としては、統計書などの「年鑑」「白書」などの年刊発行の継続図書というイメージでデータを作成できないか、という考察に入った。この場合、「パッケージ中のタイトルを年間契約でアクセス権を購入するもの」の書誌階層のモデルはうまくいくが、書誌を新たに毎年作り直さなければならない、といった手間が発生する。しかし、「最新号を基準とする」という基準からいうと、毎年書誌を作成することは致し方ないのではないか？むしろ「年間契約中の最新の書誌」を毎年作らないと、最新の情報を維持できないのではないか？とも考えられた。そこで、NACSIS-CAT 図書DBでデータベースモデルの表現を試みた。加えて、NACSIS-CAT システムのデータを NACSIS-Webcat のデータとして利用していることを考えると、NACSIS-CAT 図書DBで表現できるという長所は大きいと思われる。

また、前章 . . . で説明した理想のデータベースモデルは、具体的には年ファイルでデータをコントロールするように設計したが、NACSIS-CAT の現状ファイル構成ではそのようなファイル、リンク項目がないので、親書誌と子書誌にある YEAR 項目の年を流用してデータを作成する。また、親書誌は ED の版項目に年を入力し、前年とは版が異なるという状態を表現した。

#### 書誌ファイル

##### 親書誌

親書誌は「あるパッケージ」と「ある年」を表現する。

ID=BA00000001

YEAR=2005

GMD=w

SMD=r

VOL=<http://link.springer.de/link/service/journals/> (IDENT)

TR=Springer Link

ED=2005

PUB=

PTBL=

##### 子書誌

子書誌はパッケージ内のタイトル、「あるタイトルを」「ある接続先を利用して」を表現する。「ある接続先」は図書書誌の VOL 部分を利用して入力し、複数の接続先に対応するとともに、アクセス権保有館ごとの複数の接続も表現する。パッケージではなく、単体で購入する場合は子書誌のみで書誌情報を構成する。

ID=BA00000002  
YEAR=2005  
GMD=w  
SMD=r  
VOL=<http://link.springer.de/link/service/journals/00023/index.htm> (IDENT)・・・繰返可  
ISSN=14240661  
TR=Annales Henri Poincare  
PUB=Basel, Switzerland : Birkhauser  
NOTE=Access: via WWW ;  
PTBL=Springer Link <BA00000001>

ID:BA00000003  
GMD=w  
SMD=r  
YEAR: 2005  
VOL=<http://link.springer.de/link/service/journals/00134/index.htm> (IDENT)・・・繰返可  
VOL=<http://link.springer.de/link/service/journals/00135/index.htm> (IDENT)・・・繰返可  
ISSN=14321238  
TR=Intensive care medicine  
PUB=Berlin : Springer  
VT=VT:Intensive care med ; Intensive care medicine  
NOTE=Access: via WWW  
NOTE=Title from table of contents  
NOTE=Organ of: European Society of Intensive Care Medicine  
NOTE=ISSN for printed issues: 03424642  
PTBL=Springer Link <BA00000001>

#### 所蔵ファイル

アクセス権保有データは、NACSIS-CAT 図書DB、所蔵ファイル上で VOL(IDENT)の繰り返しで表現し、複数の接続先を表すとともに、CLN(請求記号)にアクセス可能な年次・巻号を入力し、「ある接続先を利用して」、「どの図書館が」、「どれだけのモノを(いつからいつまで)」提供されているか、を表現する。年次・巻号は契約途中からアクセスできるタイトルが変更される場合もあり、加えてILLの利用不可の場合も考えられるので、年次・巻号は入力必須としない。

ID=CC000000001  
BID=BA00000003  
FANO=FA000001



また、考えを変えると、NACSIS-Webcat 内で他のデータベースとの横断検索を行うことも解決策の一つであると考えられる。具体的には検索項目から NACSIS-Webcat の本体と、先ほど提案した「電子ジャーナルアクセス権保有データベース(仮称)」、「全国漢籍データベース」などを横断検索する、などの手法も今後の検討課題であると考えられる。NIIからすると、「横断検索は GeNii があるじゃないか」ということになるが、個々のデータベースをまず見直すことが、地味な作業であるが「全国の学術情報資源を集約する」という目的を実現できる唯一の方法ではなからうか。

## まとめ

電子ジャーナルのILLでの利用は、「根本的にできるのか?」、「著作権上許されるのか?」、「ホントに大丈夫なのか?」という問題点が冊子体ではない電子資料という理由が常につきまどう。電子ジャーナルの概念を表現すると、適合する用語が見当たらなかったが、通勤時に利用する「定期券」のような感覚か?と思われた。「定期券」は人に貸すことはできないので、ILLのような相互利用はできない、となるのだろうか? また、論文単位でバラ売りしているというジャーナルも見受けられることから、ILLよりはバラの論文を購入するほうが有効なのではという意見も存在する。加えて、著作権などの法律整備の理解も必要だが、現在のところグレーな部分であることは間違いない。NACSIS-CAT の参加組織の中には、「ILLに提供できるものを NACSIS-CAT に登録する(ILL利用できないものは登録しない)」という考え方の参加組織もあるようで、「全国の学術情報資源を集約したい」というNIIの基本となる考え方とは異なってきた部分も見受けられる。電子媒体については冊子体の資料とは異なり、変化のスピードが極端に早く、図書館としても対処できないでいる、のが現実ではなからうか。

これまで、「全国の学術情報資源を集約する」という目的を実現する手段としていくつかの方法を考案した。データベースモデルもさることながら、実際の業務担当者の理想を具体的に形にしてみたのであるが、今後の NACSIS-CAT の展開に協力できることになれば幸いである。

## 注および文献

- 1) システム設計速習読本 : 南條優のSEパワーアップ / 南條優著. - 東京 : 学習研究社 , 1992.7. - 154p ; 26cm P80-88 業務を効率化するという観点からコンピュータを使うことについて適切に説明がなされている。
- 2) NACSIS-CAT/ILL ニュースレター / 国立情報学研究所 [編] 14号 (2004.6.28) p16 「総合目録データベースにおける電子ジャーナルの取扱いについて」 NACSIS でも書誌・所蔵データの一括更新は検討されている。

- 3) 「全国漢籍データベース」(<http://kanji.zinbun.kyoto-u.ac.jp/kanseki/>) 全国漢籍データベース協議会 「漢籍」については伝統的な四部分類を基礎とする別個のデータベースが必要とされている。(全国漢籍データベース協議会 HP 抜粋)  
<http://kanji.zinbun.kyoto-u.ac.jp/kansekiyogikai/>
- 4) NACSIS-CAT 目録システムコーディングマニュアル 「6.0.4 電子ジャーナルの書誌記述」  
「17.0.1 電子ジャーナルの所蔵記述」
- 5) 「漢籍担当職員講習会 初級・中級」京都大学人文科学研究所 附属 漢字情報研究センター  
初級コースでは宮澤彰教授が「NII 総合目録データベースと全国漢籍データベース」という演題で講師をされている。

このグループ演習では、電子ジャーナルの実情・希望・理想を導入経験者にリアルに語っていただき、別の視点から、電子ジャーナルを導入していない大学図書館担当者がデータベースモデルを考案、レポートを作成、再度運用手順モデルなどを確認いただきました。単にモデルというだけで、具体例を添付することが難しかったが、ある程度具体的に、レポートの形にまとめることができたと思います。発表もこのレポートを基準に行えたのは幸運でした。これもグループ演習の私以外の3人の適切なアドバイスと、お世話いただいた研修係をはじめ夜遅くまで夜景のすばらしい施設を利用させていただいた皆様のご協力があったからこそ、と思っております。心よりお礼申し上げます。誠にありがとうございました。

(平成17年度総合目録データベース実務研修 雑誌班 磯野記)