

## 国立教育政策研究所教育研究情報データベースの 概要と検索システムの特徴

Summary of the Database and Characteristics of the Information Retrieval System  
- The Database of Educational Resources of the National Institute for Educational Policy Research -

吉岡 亮衛\*  
YOSHIOKA Ryoei

### Abstract

In the National Institute for Educational Policy Research, we have discontinued a database information retrieval system operating with the legacy general purpose computer system and started a new information retrieval service using a server-typed system.

First opened on October 20, 2006, the service was renovated in February, 2007 in reforms that made the system even simpler and easier to use.

Here, I comment on the characteristics of a new information retrieval system and summarize five databases in order to provide and introduce usage.

### 1. はじめに

旧国立教育研究所では、我が国における教育情報システムの構築に関する研究の一環として、昭和59年度に当研究所に導入されたコンピュータ<sup>①</sup>を利用して教育文献情報データベースを構築した。そして全国の都道府県・政令指定都市教育委員会、公・私立の教育研究所・教育センター、国・公・私立大学・短期大学・高等専門学校、海外日本人学校・補習授業校、公立図書館等の中から当研究所のデータベースの利用を希望する機関<sup>②</sup>に対して、オンラインによる検索サービスを開始し、実際の情報利用に伴う文献情報データベース形成上の問題点の把握とその解決策の探究に努めてきた。

その後はコンピュータと情報通信システムの発達に合わせておよそ4年毎にシステムを更新<sup>③</sup>してきた。また、画像や音声のデータベース化についても研究を進めながら、様々な種類のデータベース<sup>④</sup>を開発してきた。

当研究所では当初より教育情報データベースに関する研究の意味と、データベースの普及・啓蒙の意味から、汎用コンピュータ上のデータベース管理システムによるデータベースの構築及び検索サービスを行ってきた。しかしながら、昨今のインターネットとウェブの急速な発展に対応し、時代に則したデータベース提供サービスが望まれてきた。そこで平成16年度のコンピュータシステムの更新を期に、サーバ中心のシステム構成へと変更することとした。当面汎用コンピュータシステムによる検索サービスを継続しつつ、新規に導入したサーバ型のデータベース管理システム<sup>⑤</sup>とデータベース検索システム<sup>⑥</sup>を連携した新システムを開発して、データベース検索サービスの移行完了とともに旧システムを廃止するという手順を取った。約1年半程の期間をかけて新システムを開発し、既存のデータベースを順次新システムへ移行した。また、データベースに蓄積された源情報の

\* 教育研究情報センター総括研究官

提供機関に対しては、これまでとは異なる検索システムにより機関ではなく個人を対象にサービスを行うことについての同意を求める手続きを行った。一連の準備を経て平成18年10月20日に教育研究情報データベースとしてオープンしたことは、国立教育政策研究所広報第151号<sup>(7)</sup>に紹介した通りである。

その後平成19年2月に検索サービスの見直しを行い、教育研究情報データベースをリニューアルオープンした。ここでは、現行のシステムの特徴と、現在検索サービス中の5つのデータベースの概要について解説するとともに、教育研究情報データベースを利用する方法について紹介する。

## 2. システムの特徴

システムの最も大きな特徴は、市中のインターネット検索エンジン、例えば Yahoo! や Google と同様の検索が可能となったことである。これが意味することの一つは、インターネットに繋がるパソコンがあり、そのパソコンにブラウザが入っていれば、誰でも、いつでも、どこからでも<sup>(8)</sup>当研究所が検索サービスをする教育研究情報データベース<sup>(9)</sup>を利用できるということである。もう一つの意味は、本システムは Yahoo! や Google と同じ検索方式<sup>(10)</sup>を取っているため、Yahoo! や Google を利用したことがあれば、初めて利用する者でもその経験を基に、容易に情報検索ができるということである。

次に、検索機能の特徴について説明する。一つは、通常の簡易検索機能の他に、詳細な検索設定ができるオプション検索メニューを用意したことである。オプション検索では、特定の項目だけを検索対象とする項目設定が可能である。また、AND/OR/NOT の論理検索式を組み合わせた検索を可能としている。

二つ目は絞り込み検索機能である。あるキーワードで検索した結果に対して、さらにキーワードを追加して絞り込み検索を行う場合に、最初の検索結果を保持し、絞り込みのキーワードを次々と変えて検索し直すことができる。例えば「総合的な学習の時間」で検索した結果に対して、更に、国語に関するレコードは何件か、理科に関するレコードは何件かを知りたいときに、「総合的な学習の時間」の結果をそのままにして「国語」「理科」というように絞り込みのキーワードを次々差し替えて検索することを可能にしている。

三つ目は N-gram 方式<sup>(11)</sup>を用いた検索処理を行っている点である。例えば「総合的な学習の時間」を検索する場合に、通常はこのような長いキーワードを辞書に持っていないので、検索システムは自動的に「総合的」「学習」「時間」のように単語分割して検索し、それらの文字列が含まれる雑多なレコードを検索結果として返してくる。一方、N-gram 方式は、「総合的な学習の時間」という文字列が含まれるレコードだけを検索結果として返してくれるので、検索精度が高く検索漏れが少ないシステムと言える。

四つ目として横断検索機能がある。これまでは、個別のデータベースをひとつずつ検索する必要があったが、本システムでは、5つの教育研究情報データベースのすべてを一度に検索することが可能である。もちろん検索対象のデータベースを取捨選択し関心のあるデータベースのみを検索するような設定もできる。また、横断検索を実行した場合には、検索結果をすべて一度に表示することも可能であるが、データベース毎に順に検索結果を表示させることもできる。

次に検索結果を表示する際の工夫について3点説明する。一つは検索結果一覧画面である。各データベースの特徴的な項目のみを1レコード当たり2～3行で表示することとした。これにより一画

面に多くのレコードの情報を表示できるので、検索結果を効率よく閲覧することができ、利用者が知りたい真の情報が検索結果の中に含まれているかを容易に判断できる。さらに、目的のレコードについては、一覧をクリックすることで詳細情報画面に全情報を表示できる。

2点目は簡易ソートである。検索結果として得られたレコード数がある程度多い場合には、レコードを並べ替えて表示することができれば便利である。簡易ソート機能により、検索結果一覧に表示されているいずれか一項目について、文字コードの昇順または降順<sup>13)</sup>に並べ替えることができる。さらにより複雑なソート条件の設定は、オプション検索画面で設定が可能である。

3点目は検索結果のダウンロード機能である。利用者の情報活用の便を考えて、検索結果をパソコンにCSV形式のファイルとしてダウンロードできる機能を備えている。検索結果一覧画面からは、結果として得られたすべてのレコードが1つのファイルにダウンロードできる。また、詳細情報画面からはそこに表示された個別レコードの情報のみをダウンロードできる。ダウンロードされたファイルは、マイクロソフトのエクセル等で開くことができ、文献情報の整理・活用が行える。

さらに新システムでは、これまでは試験的に構築し蓄積してきた全文テキスト画像<sup>13)</sup>を書誌情報の検索結果とリンクさせて閲覧することを実現した。これにより、以前から利用可能であった書誌情報に加えて、一次情報も即時に利用できるようになった。

新旧システムで大きく異なる点としてユーザ管理がある。旧システムでは、利用者は申請書による利用者登録申請を行い、管理者側で承認と登録の手続きが必要であった。一方、新システムでは、利用者は、インターネットを通して自分自身で利用者登録を行い、ユーザIDとパスワード<sup>14)</sup>を取得できる。これにより、利用者登録の手間・時間・経費が不要となり、利用者は利用したいときにシステムの利用を開始できるようになった。

利用者登録時の個人情報に対するセキュリティの確保は重要な課題である。そこで本システムと利用者との間の通信には、SSLによる暗号化通信<sup>15)</sup>を使用した情報の保護対策を行っている。また、そのことの証明として、日本ベリサイン株式会社のサーバ証明書「グローバル・サーバID<sup>16)</sup>」を取得している。したがって利用者には利用者登録時に安心して氏名やメールアドレス等の個人情報を入力することができる。また、検索のやりとりについても通信は暗号化されているのでネットワークに対してプライバシーは保証されている。

最後に、今回のリニューアルでは、オープン当初設けていた一般利用者と教育関係利用者という2つの利用者区分を撤廃し、一般利用者区分でだれでも自由にデータベースを利用できるようになったこと、また、全文テキスト画像についてもダウンロードして利用することが可能となったことを付記する。

### 3. 教育研究情報データベース

新システムでは、当面、旧システムで検索サービスを行っていたデータベースの中から、当研究所が情報の収集からデータベースの構築までのすべてを担っている5つのデータベースについてデータを移行し、検索サービスを行うこととした。次にそれら5つのデータベースについて、内容等について解説し、収録項目一覧を示す。なお、データベースに収録されている情報は、基本的に2006年5月までに収集蓄積した情報になるが、実際に利用できるレコード数は、高校入試問題データベースを除いて、蓄積されている件数よりも少ないことをお断りしておく。これは、本システムによる公開の同意が、すべての情報提供機関から得られていないためである。それぞれのデータベースに

ついて、現時点で何件のレコードが利用できるのかは、検索キーワードを入力する欄に空白を入力して検索ボタンをクリックすることで知ることができる。

また、検索結果一覧で項目名のみ表示され、項目値が空白の場合がある。データベースに収録された全レコードについて、すべての項目に値が入っているとは限らず、情報収集時に欠落していた項目もあれば、論文の著者名等について情報提供機関から非公開を求められた場合等にも項目の値は空白となっていることをお断りする。

次節以降の項目一覧表では、A欄は検索項目として使用できる項目に が付き、B欄は、検索結果一覧で表示される項目に が付いている。また、詳細情報画面にはすべての項目値を持つ項目を表示する。

### 3. 1 教育研究所・教育センター刊行論文（略称：CENTER）

本データベースは、昭和56年度から平成16年度までに、全国の教育研究所・教育センターで刊行された研究紀要や報告書に掲載された研究論文や調査報告等の文献情報データベースである。情報収集は、当研究所が全国教育研究所連盟加盟機関を含む全国600余の各教育センター・教育研究所を対象として毎年度実施している教育情報調査に寄っている。毎年の更新レコード数は、約1,500件である。平成11年度分以降の情報で、当研究所の教育図書館に研究紀要等が寄贈されている場合には、掲載された論文等についてPDFによる本文画像情報を検索結果と連動してオンラインで閲覧できる形にしている。現在までの蓄積件数は、39,357件（内PDFファイル有りは、6,307件）である。なお、本文画像情報に写真が含まれる場合、その写真に人物が写っているものは黒いマスクをかけて個人を特定できないよう加工処理を施していることを御理解いただきたい。

	項目名	A	B	項目内容説明
1	文献番号	○		レコードに付与された8桁の個別番号
2	年度	○		データ収集年度（西暦）
3	標題	○	○	研究論文や調査報告等のタイトル
4	著者名	○		著者名あるいは著者グループ名
5	掲載誌名	○	○	論文等が収録されている紀要等の名称
6	掲載ページ			論文等の掲載ページ（開始ページ～終了ページ）
7	発表年月	○	○	紀要等が発刊された年月
8	索引語	○	○	論文等に付与されたキーワード
9	研究対象	○	○	研究対象となる学校段階
10	研究方法	○	○	実験研究、調査研究、実践研究、開発研究等の研究方法の区別
11	文献の種類	○	○	研究論文、指導資料、実践記録、調査報告等の文献内容の区別
12	内容の要約	○		200字程度の概要
13	本文画像	○		本文画像ファイル（PDFファイル）がある場合のリンク
14	本文画像存在	○	○	本文画像ファイル（PDFファイル）がある場合に、Yを入力
15	機関番号	○		報提供機関の番号（国立教育政策研究所で定めた4桁の数字）
16	資料保存場所	○		掲載誌が保存されている機関（国立教育政策研究所図書館に収蔵される場合、国立教育政策研究所と表記する）
17	請求番号			国立教育政策研究所図書館に収蔵する場合の図書請求番号

### 3. 2 初等中等教育諸学校における実践的教育研究主題 (GAKKOU)

本データベースは、昭和56年度から平成15年度までに、全国の初等中等教育諸学校において、研究指定等を受けるなどして実践された教育研究についての研究主題情報データベースである。(途中昭和57年度、59年度、61年度～平成2年度は、諸事情により欠落している。) 基本的な情報源は、各都道府県及び政令指定都市教育委員会が把握する実践的教育研究主題に関する情報であり、多くの教育委員会では、例えば「特色ある学校一覧」というタイトルの小冊子にこれらの研究情報をとりまとめしており、それらが基となっている。年間の更新レコード数は、約10,000件である。現在までの蓄積件数は、201,277件である。毎年情報を収集しデータベースに登録しているので、たとえば研究指定を受けた期間が複数年にわたるレコードは、それぞれの年に登録されている点は、注意を要する。

	項目名	A	B	項目内容説明
1	文献番号	○		レコードに付与された7桁の個別番号
2	年度	○		データ収集年度(西暦)
3	学校名	○	○	研究が行われた学校/園名
4	主題/特色	○	○	研究の主題あるいは研究の特色
5	分野/領域	○	○	研究の行われた分野あるいは領域
6	校/園長名			研究が行われた学校/園の代表者名
7	学校種別	○	○	学校段階の種別
8	学校所在地	○		学校の主たる所在地の住所
9	交通機関	○		学校への最も一般的な交通機関を使ったルート
10	学級数	○		普通学級と特殊学級を合わせた総学級数と特殊学級数を括弧内に示す
11	研究指定	○	○	研究指定の名称又は種類
12	指定年度	○	○	研究指定を受けた年度(開始年度～終了年度)
13	都道府県名	○	○	源情報提供教育委員会の都道府県名または政令指定都市名
14	学校番号	○		学校基本調査で用いられる6桁の数字
15	報告書の有無	○		報告書が出されている場合に「有」と入力
16	備考	○		その他の特記事項がある場合

### 3. 3 地方教育センター等における教職員研修講座 (KENSHU)

本データベースは、昭和60年度から平成17年度までに、全国の教育研究所・教育センター等で実施されている教職員研修講座に関する情報のデータベースである。(昭和61年度、62年度は未調査のためデータが欠落している。) 情報収集は、教育研究所・教育センター刊行論文の情報収集と同じ対象機関に同時に行っている。年間の更新レコード数は約8,000件であり、現在までの蓄積件数は、144,094件である。

	項目名	A	B	項目内容説明
1	講座番号	○		レコードに付与された10桁の個別番号
2	年度	○	○	講座実施年度(西暦)
3	講座名	○	○	研修講座タイトル
4	機関名	○	○	情報提供機関名
5	機関番号	○		情報提供機関の番号(国立教育政策研究所で定めた4桁の数字)
6	主催者	○		講座の主催者
7	開催場所	○		講座の開催場所
8	受講者	○	○	講座の受講対象者の別
9	校種	○	○	講座の受講対象者の校種
10	受講義務	○		講座の受講対象者の要件
11	受講定員	○		講座の総受講定員
12	実施形態	○		講座が実施される形態。例えば毎年開催、隔年開催等
13	実施回数	○		年間に開催される講座の実施回数
14	実施日数	○		1回の講座が実施される日数
15	開催月	○		講座が実施される月
16	テキスト有無	○		使用するテキストの有無
17	テキスト種類	○		使用するテキストが市販のものか、自作のものかの別
18	使用機器	○		講座で使用する機器類
19	索引語	○		講座の内容を示すキーワード
20	研修方法	○	○	講座で用いられる研修の方法
21	講師	○		講師に関する身分や肩書の情報
22	講座内容	○		400字程度の講座内容説明

### 3.4 高校入試問題(MONDAI)

本データベースは、平成3年度から平成12年度までの、各都道府県教育委員会が行った公立高校の入試問題に関する情報のデータベースである。現在のところ解答情報について、正答例、正答率等の情報を提供しており、問題文については問題文画像の形で正答についても正答画像を公開できるよう準備を進めているところである。また、平成13年度以降の情報についても、現在データの整理中である。平成12年度までの蓄積件数は、問の数で76,269件である。年間の更新レコード数は、問の数で約8,000件である。

	項目名	A	B	項目内容説明
1	問題ID	○		レコードに付与された8桁の個別番号
2	年度	○	○	入試年度(西暦)
3	都道府県	○	○	都道府県名
4	教科	○	○	国語・社会・数学・理科・英語のうちのいずれかひとつ
5	大問番号	○	○	問題文中の大問番号
6	小問番号	○	○	問題文中の小問番号
7	正答率	○	○	百分率で表示する
8	配点割合	○	○	全体に対する当該設問の配点を示す。例えば3/60
9	難易度	○	○	正答率に基づく難易度を1(易)～5(難)で表示する
10	問題番号	○	○	当該問題に付与された問題文中の文字表記
11	正答例	○	○	正答として示された例
12	備考	○	○	

### 3. 5 教育学関係博士・修士学位論文題目 (JEDI)

本データベースは、新制大学発足以降に我が国の大学院研究科において学位を授与された博士及び修士の学位論文の中で、次のものを収録の対象とする。

- ・教育学及び教科教育学の研究。
- ・国語・国文学、外国語・外国文学、数学、物理学、地球科学、化学、生物学等の理学、体育学、家政学、音楽学、美術学、情報科学、福祉・看護学等の研究のうち、教育に関する研究または学校教育及び教科教育の研究に資する研究をテーマとしている研究。

これらについて、その論文題目等の情報と抄録誌に収録された抄録の全文画像からなるデータベースである。

当初は、教育学の学位を授与された論文題目のみを収集してきたが、平成8年から学校教育及び教科教育の研究に資する研究テーマまで情報収集の範囲を広げた。また、平成17年度には抄録誌の全文画像情報を追加しデータベースの拡充を図ってきている。年間の更新レコード数は、約6,000件である。現在までの蓄積件数は、95,664件（内 PDF ファイル有りは、35,142件）である。

	項目名	A	B	項目内容説明
1	レコード番号	○		レコードに付与された8桁の個別番号
2	ID			12桁のレコード管理情報
3	研究科番号	○		数字4桁の研究科コード
4	大学院名	○	○	大学院の名称
5	研究科名	○	○	研究科の名称
6	学位授与年(元号)	○		元号による学位授与年月
7	学位授与年(西暦)	○	○	西暦による学位授与年月。例：198903
8	学位名称	○	○	学位の名称。例：修士(教育学)
9	氏名	○		学位を授与された者の氏名
10	論文区分	○	○	修士の場合は論文が主論/副論の別。博士の場合は課程博/論博の別
11	論文題目	○		学位論文のタイトル
12	専攻名	○		研究科における専攻名
13	分野	○		研究分野
14	抄録	○		抄録のPDFファイルへのリンク
15	抄録有無	○	○	抄録ファイルが存在する場合にYを入力

## 4. 利用方法

本システムへのアクセスについて説明する。当研究所のホームページ (<http://www.nier.go.jp>) のメニューにある教育研究情報データベースのパナーをクリックするか、または、直接ブラウザのアドレス欄に、<http://www.nier.go.jp/database/>と入力することで、教育研究情報データベースのトップページへとジャンプする。教育研究情報データベースの利用に関する情報はすべてここにまとめているので、ブックマークにはこのページを登録することが望ましい。利用規程及びその他の規定等、利用するためのパソコン環境、利用マニュアル、提供データベースの解説もすべてこのページから参照できる形となっている。

ここでは、最もシンプルな利用方法を説明し、更に詳しい利用方法については上述のページに譲る。データベースの検索はいたって簡単である。教育研究情報データベースのページにあるリンクからログイン画面を呼び出し、ユーザIDとパスワードを入力する<sup>7)</sup>。簡易検索画面が表示される

ので、検索したいデータベースを選択し、キーワード入力欄に検索対象となるキーワードを入力して検索ボタンをクリックする。基本操作はこれだけである。検索し終わった後は、画面右上のログアウトボタンをクリックして終了である。

データベースの中から関係する情報を漏れなく検索したい場合、あるいは目的の情報をピンポイントで検索したい場合などは、オプション検索を活用して行うことができる。詳しくは、ホームページの利用方法のページを参照のこと。

ログインからログアウトまでの操作中にひとつだけ守っていただきたい注意事項がある。それはデータベースの検索を行っている間は、ブラウザの戻るボタンを使用しないということである。ブラウザの戻るボタンが使用されると、サーバ側での利用者のセッション管理に不具合が生じ、正しく検索ができなくなることがある。そのためこの点を守って、快適に検索を行っていただきたい。

教育研究情報データベース並びに検索システムについての御意見・御要望等は、E-mail: dbadmin@nier.go.jp まで御連絡願いたい。

## 注

- (1) コンピュータはファコム・ハイタック社製、FACOM M340S。データベース管理システムは、FACOM FAIRS-I。
- (2) 教育関係機関に限り、利用申請を受け付けて、利用者 ID とパスワードの発行を行う形を取った。
- (3) 過去のシステム更新は、昭和63年度、平成2年度、平成7年度、平成11年度、平成16年度（現行システム）に行った。
- (4) 財教科書研究センターとの共同開発によるもの、岐阜大学教育学部附属カリキュラム開発研究センター及び国立大学教育工学センター協議会によって作成され当研究所に移植されたもの、科学研究費補助金等の研究グループ及びデータベース作成委員会により作成されたもの等、最大で18タイトルのデータベースの提供サービスを行っていた。それら以外に、統計データのデータベース化も行った。
- (5) Solaris 9版の日本オラクル社製、Oracle 10g。
- (6) Windows 版のアクセラテクノロジー社製、Bizsearch V.3をカスタマイズ。
- (7) 国立教育政策研究所広報第151号、P.7、平成18年10月発行。
- (8) 旧システムでは、利用時間はウィークデーの午前10時～午後5時に限定し、利用者は事前に郵送による利用申請を行うことが必要であった。また、利用者は、教育関係機関・部局に所属する者あるいは所長が認めた者に限られた。
- (9) 新システムへのリニューアルを期に、「教育情報データベース」を「教育研究情報データベース」と名称変更した。
- (10) ここでは、入力ボックスに思いつく言葉を入力して検索実行するという方式のことを指す。旧システムは、コマンドラインに検索コマンドと検索子と検索する文字列を入力するコマンド形式を採用していた。
- (11) 自然語解析処理によるこれまでの単語一致方式とは異なり、N-gram 方式は、入力された文字列の長さに基づいて検索処理を行う方式と言える。
- (12) 文字コードの値が小さいものから大きいものへ並べる場合を昇順、大きいものから小さいものへと並べる場合を降順という。仮名は漢字よりもコードの値は小さく、ひらがなの中では「あ」から「ん」までの順に値が大きくなる。
- (13) 現在情報提供サービスを行うデータベースのうち、教育研究所・教育センター刊行論文及び教育学関係博士・修士学位論文題目の2つについて、前者は紀要等の本文を、後者は抄録全文を提供者から公開の許諾を得られたものについて利用可能となっている。



- (14) 利用者が設定するユーザ ID とパスワード、及び、姓名、メールアドレスを入力することにより、ユーザ ID とパスワードが登録される仕組みとなっている。
- (15) SSL (Secure Socket Layer) は、Netscape Communications 社が開発した、インターネット上で情報を暗号化して送受信するプロトコル。現在インターネットで広く使われている WWW や FTP などのデータを暗号化し、プライバシーに関わる情報やクレジットカード番号、企業秘密などを安全に送受信することができる。
- (16) 日本ベリサイン株式会社は、ウェブサイトの信頼性をアピールする証明書の発行機関としてのリーディングカンパニーであり、研究所のサーバはこの証明書を取得している。
- (17) はじめての場合にはユーザ登録を先に行う必要がある。