

漢字文献 情報処理研究

第 15 号

漢字文献情報処理研究会 編

好文出版

漢字文献情報処理研究 第15号

目次

論文	4	HTML のルビ標準化の現状と課題	川幡 太一
特集 1	情報化時代における中国学研究基盤を考える	11	
	12	学術情報収集のスキルとインフラ整備	千田 大介
	21	学術データベースの構築と公開	師 茂樹
	30	学術情報共有プラットフォームとしての CiNii	大向 一輝
	36	2014 年度夏期公開シンポジウム 「情報化時代における中国学研究基盤を考える」レポート	佐藤 信弥
特集 2	東洋学論文検索指南	41	
	42	日本の学術論文データベース	小島 浩之
	51	中国の論文を探す	千田 大介
	65	台湾・香港の論文データベース	二階堂善弘
	70	韓国の学術論文の検索方法	師 茂樹
	73	ベトナムの文献を探索する	矢野 正隆
	78	『人文学と著作権問題』余話	漢字文献情報処理研究会
	85	外国語教育ワークショップ「ブチ IT を利用した外国語教育の実践」報告 ～ PC1 台からできる外国語教育における IT 利用	氷野 善寛・紅粉 芳恵
レビュー & リソース紹介	89		
	90	Microsoft Office Mobile & Online	田邊 鉄
	93	一太郎 2014 徹 & ATOK2014	山田 崇仁
	97	BabelPad	師 茂樹

	100	Windows 8.1 環境におけるピンイン処理	千田 大介
	104	When.exe Ruby 版	須賀 隆
	110	中華經典典籍庫	千田 大介
	114	瀚堂典藏	千田 大介
	118	蘇州地方文献データベース	佐藤 仁史
	120	華東師範大学中国文字研究与应用中心	山田 崇仁
	126	中国・台湾の古典籍検索	小島 浩之
	136	「農業古籍」について	村上 陽子
お知らせ	40	漢情研記念論集（仮）原稿募集のお知らせ	
	140	年会費の納入についてのお知らせ	
	141	漢字文献情報処理研究会彙報	
	142	執筆者紹介	
	144	編集後記	

- 本誌記事中のソフトウェア名・プログラム名・会社名などは一般に各社の商標または登録商標です。本文中では、TM・®等のマークは明記していません。
- 本誌記事の記述に基づいて行われた作業の結果生じたあらゆる損害について、編著者・翻訳者および出版社は一切の責任を負いません。
- 本誌記事の内容に関するご意見・ご質問は、漢字文献情報処理研究会 Web サイト（<http://www.jaet.gr.jp/>）のフォームにて受け付けます。書面・電話・FAX によるお問い合わせには応じかねます。

HTML のルビ標準化の 現状と課題

川幡 太一 (かわばた たいち)

回 はじめに

「ルビ」とは本来、文字の脇に付加する、意味や読み方などを示す小さな文字のことである¹⁾。ルビは多くの書籍で見られ、それは近年、急激に普及している電子書籍でも同様である。電子書籍の規格 "EPUB3" では、コンテンツのフォーマットとして HTML5 を採用している。EPUB3 が標準化された当時、HTML5 は正式な勧告となっており、2011 年時点における規格案（ドラフト）が採用された。当時の HTML5 案で規定されたルビは、十分な実装経験による裏付けがなかったため、HTML5 の正式勧告では採用されないリスクがあった。また同案のルビでは、日本語にとって重要な機能の一部が利用できなかった。そのため、日本の電子書籍の普及にあたり、HTML5 において、ルビが日本のニーズを満たす形で標準化されることが必要であった。

ルビは日本では一般には漢字に対する「よみがな」を示すことが多いが、特殊なルビの例として、下図で示すように、日本語と外国語を対応させて表記することもできる。また、ルビは親文字の半分の大きさで表記される場合があるため、大体、ルビと親文字の文字数の比が 1:2 程度ならばバランス良く表示されるが、中にはルビと親文字の文字数のバランスが大きく崩れたものもある。

またルビは「当て字」にも使われ、近年ではラ

editor エディター
編集者 editor

図 1 ルビによる外国語表記の例 (JLREQ より引用)

イトノベルなどを中心に書名にも当て字を用いるものが増えている。

現在、W3C では 2014 年 12 月の勧告を目指し、HTML5 の標準化作業を行っている。本稿では、HTML5 および CSS におけるルビの標準化の現状と課題について述べる。

回 データ構造としてのルビ

テキストは本来、直列的なデータ構造を持つ。一方、ルビは、テキストの特定の連続領域に対して、並列的なデータ構造を表現する。



図 2：テキストの直列構造とルビによる並列構造

テキストのデータ構造の並列化は、例えば視覚障害者に向けた音声などの補助情報や、外国語の文書における対訳翻訳テキストの付加など、様々な応用例を考えることができる。

すなわち、ルビは単なる「読み仮名」としての用途を超えて、様々な応用が考えられ、その HTML5 などのマークアップにおける標準化はテキストに対する新しい可能性を開くことができる。

回 W3C における標準化

◇ HTML5 Ruby と CSS Ruby

「ルビ」の実現は以下の 2 つの技術に分けて考え

ることができる。

- テキストのデータ構造の並列化
- 画面でのルビの配置方法

W3C では前者は HTML5 で標準化され、後者は CSS で標準化される。

✧ W3C の標準化プロセス

W3C における標準化のプロセスは、W3C Process Document^[2] に整理されている。それによると、HTML における標準化は、図 3 に示す手順を踏むことになっている。

W3C の標準化プロセスのうち、困難なのは、「十分な実装経験を蓄積した」と言える Proposed Recommendation へ到達することである。W3C HTML 委員会は、HTML5 を 12 月までに正式な Recommendation (勧告) に到達することを目指している^[3]。そのため、ルビに関するマークアップなどの規定を HTML5 の標準に含めるためには、7 月までに、ルビに関するテストの作成および 2 つ以上のブラウザへの実装を実現する必要があった。



図 3 W3C における標準化プロセス

一方、HTML5 の当初のドラフトでは、ルビは [4] の複雑さを避けるため、<rb> 要素や <rbspan> 属性を使用しない単純なルビモデルが提案された。このルビのモデルは HTML5 と共に EPUB3 で採用され、結果として多くの電子書籍で利用された。

しかし、ルビの利用例などを調査した結果^[5]、この HTML5 のルビのモデルは「熟語ルビ」や「両側ルビ」といった、実際の日本の書籍におけるルビのニーズを満たしきれないことが明らかになった。

この問題を解決するために、W3C の HTML5 の中核メンバーである Robin Berjon が中心となり、ルビの拡張仕様 Ruby Extension Spec^[6] が作成された。この仕様は <rb> 要素や <rbspan> 属性

回 HTML5 における標準化

✧ 規格策定の経緯

W3C におけるルビの標準化プロセスにおける最初の標準化の試みの一つとして、1998 年から 2001 年にかけて標準化作業が行われた、「Ruby Annotations」^[4] (XHTML Rruby) が挙げられる。この規格では、両側ルビや、<rbspan> を用いた複合ルビに対応していたが、実際に実装・普及にはいたらなかった。

「Ruby Annotations」では、複数の親文字とルビ文字を対応させる場合は、<ruby> 要素の直下に <rb> (本文コンテンツ) 要素を入れ、その下に複数の <rb> 要素を入れる。<ruby> 要素には最大、1 つまでの <rb> 要素しか入らない。

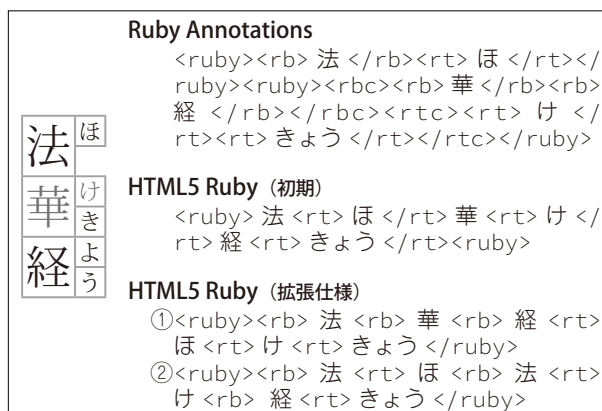


図 4

を避け、EPUB3 で採用された当初のルビ仕様を満たしつつ、「熟語ルビ」などのニーズに対応できるように考慮され `<rb>` 要素も導入された。

図 4 に、「Ruby Annotations」の当初のルビ案、HTML5 の単純ルビ、拡張ルビ案、それぞれで「法華経」に対して「ほけきょう」とルビを入れた場合の表記の違いを示す。

2013 年 10 月に中国・深圳で開催された TPAC 会議において、本拡張案を HTML5 にマージされることが認められ、同年 12 月に HTML Ruby Extension Spec は、HTML5 の規格案に組み入れられた。

※ HTML ルビの特徴

現在の HTML5 のドラフト⁴⁾で規定されているルビのタグには以下の種類がある。

<code><ruby></code>	ルビ全体構造の開始を示すタグ
<code><rb></code>	本文の開始を示すタグ
<code><rt></code>	個々のルビの開始を示すタグ
<code><rtc></code>	ルビ全体の開始を示すタグ
<code><rp></code>	ルビが表示できないブラウザのために、括弧を表示するタグ

HTML5 のルビには以下のような特徴がある。

- ①一部の閉じタグが省略できる、
- ②本文とルビの間に改行を入れることができる、
- ③一つの親文字に対し複数のルビを指定することができる（両側ルビ）、
- ④連続する親文字のグループに対し、対応する複数のルビのグループを指定できる、
- ⑤（両側ルビなどにおいて）複数の親文字のグループに対し、単一のルビを対応させることができる。

その一例を図 5 に示す。

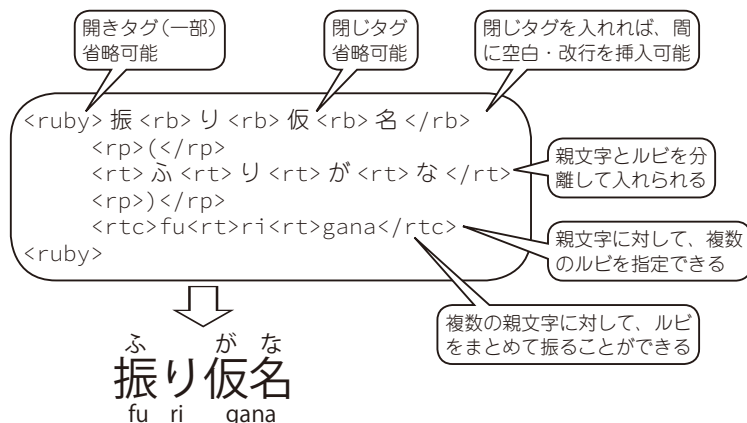


図 5 HTML5 ルビの特徴

※ ルビ要素のパーザ規定

HTML5 におけるルビの規定は、主にパーザ（解釈器）部分とモデル（分割・分類アルゴリズム）部分に別れる。パーザ部分では、HTML5 テキストのルビを解釈して、DOM 木（Document Object Model Tree）を構築する際に、要素を閉じるタグの省略などが規定される。

例えば、

```
<ruby> 東<rb>京<rtc>とう<rt>きょう</ruby>
```

という HTML がある場合、パーザは「京」の直後の `</rb>`、または「きょう」の直後の `</rt>` および `</rtc>` を補完する。

一方、「東」の直前には `<rb>` の開きタグはあってもなくても良い。ない場合は、パーザでは `<rb>` タグは補完されず、そのまま `<ruby>` 要素直下のテキストノードとして解釈され、次に述べる「分割・分類アルゴリズム」によってルビの親文字とみなされる。

※ ルビの分割・分類アルゴリズムとペアリング

パーズされて、DOM 木として構築された HTML5 ルビは、後述の CSS Ruby によって画面に表示される。HTML5 は画面表示に関与しない。しかし、ルビ要素に対して、例えば色を付け

るなどの操作を行う JavaScript などのプログラムは、Ruby 要素の DOM 木のうち、どの部分が「本文」で、どの部分が「ルビ」で、それらがどう対応するかを、自分で見つける必要がある。そのため、HTML5 仕様にはルビの分割・分類 (Segmentation and Categorisation) アルゴリズムが規定されている。

なお、このアルゴリズムはあくまでも、DOM 木に対してルビを解釈したい JavaScript が実装すべきものであり、Web ブラウザは後述の別のアル

❖ テストの作成と実装

ルビを HTML5 で標準化するためには、Ruby パーザのテストスイートの作成およびブラウザへの実装が必要である。ルビに関するテストの作成は、W3C の公式な HTML5 テストライブラリである、"html5lib-tests" への貢献という形で行われた^[8]。W3C の公式なテストは自動的に、各開発対象となるブラウザのテスト・ツールに取り込まれ、テストされる。HTML5 Ruby が実装されていないブラウザでは、これらのテストは失敗する。

このように各ブラウザにおいてテストが失敗することを示すことで、実装の必要性がブラウザの

開発者に客観的に可視化される。テストのコミット後、日本の開発者が中心となり、まず WebKit に対するルビのパーザ実装のパッチを作成してマージを依頼し、受理された^[9]。その後、Gecko (Firefox のエンジン) においてもテストの失敗を根拠とするバグの登録、パッチの作成とコミットを行い、幾つかの質疑応答などがあった後に受理された^[10]。

✧ 現状と今後の課題

現在、HTML5 の標準化団体は、W3C と WHATWG に分かれている。2014 年 9 月 16 日、HTML5 はルビを含めた形で W3C の Proposed Recommendation となった。もう一方の WHATWG においてもルビは W3C の規定と同じ形で HTML5 に取り込まれた。

現在の HTML 5 ルビはパーザ部分は WebKit, Blink および Gecko で実装されているものの、まだ、Internet Explorer への実装は未達成である。そのため、今後もブラウザへの実装を呼びかける必要がある。

回 CSS での標準化

✧ CSS Ruby Level 1

HTML5 の Ruby の規定は、HTML マークアップにおける、ルビのデータ構造を規定する。一方、CSS Ruby^[11] は、一般的な XML 文書におけるルビのデータ構造、およびルビの画面での表示方法を指定する。CSS Ruby の規定は以下の 3 つに大別される。

- ボックスモデル
- レイアウト
- 表示プロパティ

本章ではこれらについて概要を説明する。

✧ ボックスモデル

ボックスモデルは、HTML および XML 文書に

おける、CSS として XML 要素に指定可能なルビに関する display プロパティ値と、DOM 木からレンダリング木への変換方法を規定する。

HTML では、ルビを表現する要素として <ruby> や <rb> などが規定されている。しかし、実際にどの HTML 要素が、ルビの役割を果たすかは、要素名ではなく、各要素の CSS スタイルシートにおける display プロパティ値で最終的に決定される。

また、HTML や XHTML 以外の XML 文書でルビを使用する際、その XML のどの要素がルビになるのかも、CSS では display の値を用いて決定する。

以下の表に、CSS のボックスモデルが規定するルビに関する display の値と、それらの値をデフォルトで持つ HTML5 の要素名の一覧を示す。

display 値	対応する HTML5 要素名
ruby	<ruby>
ruby-base	<rb>
ruby-text	<rt>
ruby-base-container	該当なし
ruby-text-container	<rtc>

上の表で見るように、HTML5 には、CSS のボックスモデルにおける ruby-base-container 値に対応する要素は存在しない。この要素が存在しない場合、CSS では、ボックスモデルで定められた DOM 木からレンダリング木への変換アルゴリズムによって、当該 display 値を持つボックスが自動的に生成・補完される。

HTML5 の DOM 木におけるルビでは、<rb> や <rtc>, <rt> 要素は必要に応じて省略することができる。そのため、DOM 木からレンダリング木を構成する際には、表示に必要なボックスが自動的に補われる。これは無名ボックス生成 (anonymous box generation) と呼ばれる。ルビを画面表示する際のレンダリング木への変換の例を図 7 に示す。

現在の Web では、JavaScript による DOM 木の操作は必要不可欠な技術となっている。JavaScript によって DOM 木が更新されると、ブラウザはそ

例: `<ruby> 東 <rb> 京 <rp> (<rt> とう <rt> きょう <rtc> Tokyo <rp>) </ruby>`

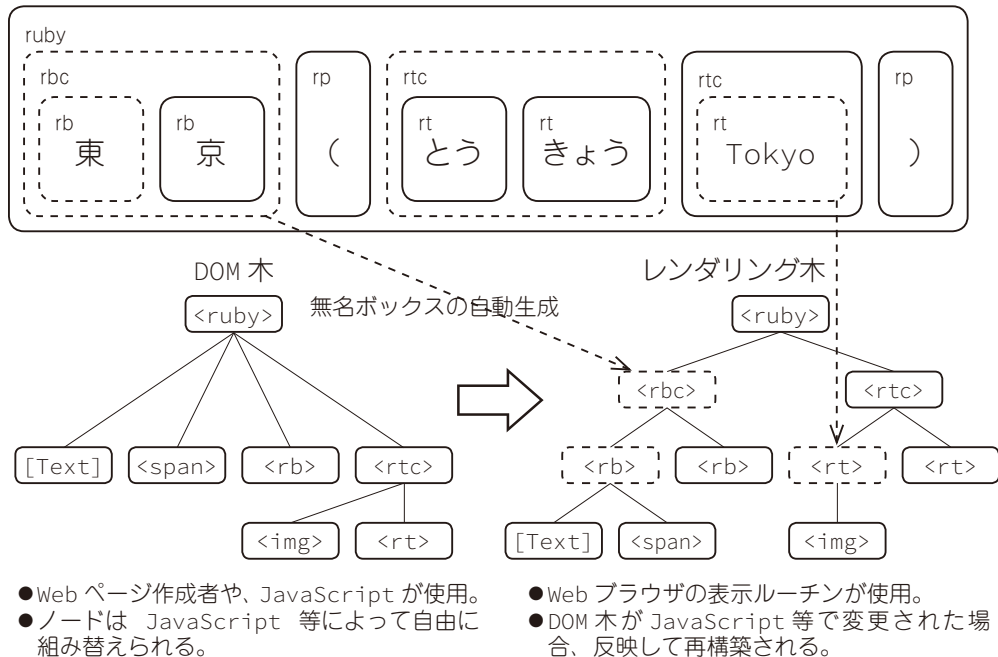


図 7 DOM 木からレンダリング木への交換の例

の都度、画面表示のためにレンダリング木を更新する。

※ ルビのレイアウト

DOM 木からレンダリング木が構成されると、レンダリングエンジンはそれに基づいて画面にルビを配置する。横書きの場合、本文のベースラインは前後の文章と一致するように縦方向に並べられ、ルビまたは本文の横方向の配置は、いずれかの文字列の長さのうち、長い方に合わせて、他方は（日本語では）空白が均等になるように分配する。

HTML と同様、CSS も国際標準であるため、日本語以外の言語でルビが使用される場合の配慮も必要である。

例えば、ルビを行の途中で改行する場合、英語のような単語単位で改行する言語や、本文とルビで書記方向が異なる言語（日本語とアラビア語など）でどのように改行の前後で親文字とルビ文字を配置するか、などに配慮しなければいけないが、これらは JREQ などの規定の対象外であり、不自然でなく、市場のニーズに合う形で新規に規定を

作らなければならない。

※ ルビのプロパティ

CSS Ruby Level 1 では、日本語の電子書籍で必要な最低限のプロパティとして、以下の種類と値を規定している。

プロパティ名	対象	プロパティ値
ruby-position	ルビのコンテナ	over / under / inter-character & right / left
ruby-merge	ルビのコンテナ	separate / collapse / auto
ruby-align	ルビの親文字、ルビおよびそれらのコンテナ	start / center / space-between / space-around

ruby-position は、本文のどちら側にルビをかけるか指定する。例えば "ruby-position: over right" 指定は、ルビを横書きでは上側、縦書きでは右側に配置することを示す。一つの本文の両側にルビをふる場合は、2 つめのルビコンテナには "ruby-position: under left" を指定する。

また、台湾の一部の国語教科書では、書記方向に沿った漢字の脇にボボモフォを振って発音を示す、日本とは異なるルビの用法があり、そのために inter-character という値が用意されている。

ruby-merge は、本文とルビが複数のペアを構成する際、「モノルビ」か「熟語ルビ」かを指定する。また、ruby-align は、ルビの本文に対する配置方法を指定する。

◆ 今後の課題

CSS Ruby は国際標準であり、ルビが日本語だけではなく、世界中の様々な言語（英語やアラビア語、またはそれらの混在）において使用されることも考慮しなければならない。このような「ルビの世界化・多言語化」において、JLREQ は無力な点があり、今後はこの部分を、日本市場のニーズを満たしつつ、世界的に公平な形でどのように精密に定めていくかが課題になる。

また、CSS Ruby は（HTML5 Ruby とは異なり）W3C の規格制定プロセスにおける Public Draft の段階であり、テストの作成やブラウザへの実装などはまだ不十分な状況である。これらの実現には応分のコストがかかると見られ、その負担などが今後の課題になる。

回 最後に

現在、HTML5 におけるルビの標準化は、WebKit と Gecko の 2 つのブラウザエンジンで実装されたことにより、「十分な実装経験を蓄積している」状態にある。2014 年 12 月に予定される、W3C 諮問委員会による HTML5 の勧告において、ルビの規定が含まれる可能性は高い。

一方、CSS Ruby は、そのルビの配置方法の規定や多言語への配慮、それらに基づくテストの作成、ブラウザへの実装などはまだ不十分であるなど課題が山積している。現在の Public Draft が、Last Call の段階に進むのにもしばらく時間がかかる見込みである。

CSS Ruby が、Public Draft の状態であるとは、すなわち CSS ワークグループの公開メーリングリスト（www-styles@w3.org）で幅広く意見を募り、議論を行う状態であるということでもある。策定作業の発展と規格の完成に向けて、より多くの方に本メーリングリストでの議論に参加し、または規格のテストの作成やブラウザの実装などの貢献をしていただけると幸いである。

参考文献

- [1] “日本語組版処理の要件,” 2012. [Online]. Available: <http://www.w3.org/TR/2012/NOTE-jlreq-20120403/>
- [2] C. M. Nevile and others, World Wide Web Consortium Process Document. 2013 [Online]. Available: <http://www.w3.org/2014/Process-20140801>
- [3] M. Smith, HTML Working Group Charter. 2013 [Online]. Available: <http://www.w3.org/2007/03/HTML-WG-charter>
- [4] M. SUIGNARD and others, Ruby Annotation. 1999–2001 [Online]. Available: <http://www.w3.org/TR/ruby/>
- [5] R. Ishida, Use Cases and Exploratory Approaches for Ruby Markup. 2013 [Online]. Available: <http://www.w3.org/TR/ruby-use-cases/>
- [6] R. Berjon, Ruby Extension Specification. 2013 [Online]. Available: <http://darobin.github.io/html-ruby/>
- [7] R. Berjon and others, HTML5 Candidate Recommendation. 2014 [Online]. Available: <http://www.w3.org/TR/html5>
- [8] “Add some tests to match changes made in new ruby model #27.” 2014 [Online]. Available: <https://github.com/html5lib/html5lib-tests/pull/27/files>
- [9] “Catch up ruby and its tag omission rule changes in HTML5 CR Feb 2014,” 2014. [Online]. Available: https://bugs.webkit.org/show_bug.cgi?id=131175
- [10] “Implement spec changes to ruby parsing.” 2014. [Online]. Available: https://bugzilla.mozilla.org/show_bug.cgi?id=664104#c33
- [11] “CSS Ruby Layout Module Level 1,” 2013. [Online]. Available: <http://www.w3.org/TR/2014/WD-css-ruby-1-20140805/>

2014 年度夏期公開シンポジウム

情報化時代における
中国学研究基盤を考える

本特集は、2014 年 8 月 2 日に関西大学で開催された「2014 年度 夏期公開シンポジウム 情報化時代における中国学研究基盤を考える」(科学研究費補助金・基盤研究 (B)「情報化時代における中国学次世代デジタル研究基盤の確立」との共催)の報告である。

今日、ICT を利用した研究支援や、研究成果の発信の重要性は、以前よりも高まっていると言ってよい。しかし、我が国の人文学、特に東洋学・中国学における ICT 利用は、他国に比して遅れをとっている。中国や韓国では、様々な学術情報がデジタル化され、インターネットを通じて利用することが当たり前になってきている。ある韓国人留学生は、「日本では、図書館の本は図書館に行かないと読めないのが不便だ」と嘆いていた。

また日本では、せっかく公開された学術情報が、検索可能になっていないためにユーザの目に届かないことも多い。検索できない最大の原因は、論文や写本などの文字資料が画像データとして公開されていることである。中国では、大規模なテキストデータベースを開発・販売するビジネスが定着しており、すでに膨大な文字数の文献が検索可能となっている。検索可能な文献こそが研究者の関心を集め、検索できない文献は研究者の目に触れる機会がどんどん失われていく。

今後、論文や史資料を検索可能なデータとして公開することが国家レベルで求められるが、我が国はデータ開発が研究業績として評価されたり、ビジネスになったりするような風土ではなく、実現のためには課題も多い。このような状況では、個人レベルでのスキルを高めることが重要になってくる。どこにどのような資料があるのかを見つけ出す能力や、自分の論文を一人でも多くの研究者に読んでもらうための公開のテクニックは、インフラが整備されていない現在、研究者個人が身につけるべきスキルである。

本シンポジウムでは、これらの課題について情報を共有するとともに、今後のあるべき方向性について議論を行った。報告者による報告の後、一般参加者も交えた活発なディスカッションが行われた。その内容については各記事を参照されたい。

なお、関西大学アジア文化研究センター (CSAC) には、会場や懇親会など、多くのご支援をいただいた。ここに記して感謝申し上げる。

Contents

学術情報収集のスキルとインフラ整備	千田 大介	12
学術データベースの構築と公開	師 茂樹	21
学術情報共有プラットフォームとしての CiNii	大向 一輝	30
2014 年度夏期公開シンポジウム		
「情報化時代における中国学研究基盤を考える」レポート	佐藤 信弥	36

学術情報収集のスキルとインフラ整備

千田 大介（ちだ だいすけ）

はじめに

近年、研究・教育の分野においても、情報化が急速に進展したことにより、従来の紙媒体などのアナログ・メディアを主体とした研究・教育の方法は変革を余儀なくされている。こと中国学—中国を研究対象とした人文科学研究—については、主に中国本土や台湾で、学術データベースの構築が大規模に進展しているため、好むと好まざるとにかかわらず、今やそれらデジタル資料を活用することなくして研究の水準を確保できなくなっているといっても過言ではなかろう。

インターネットや情報技術によって提供される学術情報は、研究の素材となる資料と研究成果とに大別できよう。前者には文献データベース（コーパス）、統計データ、画像・映像などが含まれ、電子ジャーナルや論文データベースなどが後者にあたる。本稿ではこのうち前者、特に中国古典文献の全文データベースに絞り、情報化時代の中国学にデジタル研究インフラの整備、情報収集に必要なスキルといった問題について考えてみたい。

中国古典文献のデジタル化

中国学研究と文献調査

中国古典文献を研究資料として読解する際には、典故引用という修辞法が多用されるため、それら

の出典を確認してニュアンスを把握しなくてはならない。

研究ジャンルや手法によっては、出土資料や口述資料、あるいは今まで存在が知られていなかった文献資料の発掘など、新出資料が用いられることもある。しかしそれらも、読解に際して典故引用の把握が必要となるのは同様である。したがって、語句の出典・用例の調査と把握は、文学・史学・哲学を問わず、中国学の文献資料を解読する上で共通の基礎的作業であるといえる。

出典・用例の調査は、従来、記憶に頼る、あるいは辞書や類書・百科事典、逐字索引などを使うことによって行われていた。筆者も大学院生時代の古典詩歌の演習では、諸橋『大漢和』を引いて、それに無い言葉は『駢字類編』や『佩文韻府』を引いて、出典に当たって……という作業を何度となく繰り返し、ときに『佩文韻府』の非實在出典に泣かされたりしたものだ。

この出典・用例調査には、デジタル化のメリットが直ちに現れる。従来、辞書などの工具書や記憶に頼らざるを得なかったのが、瞬時に数億字の文献を網羅的に検索できるようになったのだから。とはいえ、成語化した言葉とその出典とで表現が異なることは多々あるし、また検索結果の中から最も古い用例を判断しなくてはならないなど、データベース検索のみで出典・用例調査は完結しないことには留意が必要であり、辞書とデータベースを往き来しながら読み解いていく必要がある。もっとも、辞書についても『漢語大詞典』

CD-ROM や CNKI 工具書館のようなサービスがあるので、いまや複数データベースを往き来する方がより効率的であるのだが。

当然のことながら、文献読解のための語句の調査のみならず、ある事件や人物、思想などの具体例の検索にも、大規模古典文献データベースは大いに威力を発揮する。データベースによって網羅的調査を行っていない場合、資料を見落とすかもしれない、データベースを使った研究と比べて、引用資料の数と精度とに大きな差が生まれることだろう。それが研究の評価に結びついていく可能性もある。その意味で、大規模古典文献データベースは、今後の中国学研究のクオリティを左右するインフラであると位置づけることができよう。

あるいは、中国学の基礎は資料を読み込むことと、より多くの文献に目を通すことにあり、データベースに頼るのは邪道である、と考える人がいるかもしれない。筆者もそうした修養は当然必要であると考えます。しかし、そうした基礎を身につけた上で、さらにデータベースを使いこなすことができれば、研究を大いに効率化させ深化させることができるだろう。大規模古典文献データベースの持つ研究上の意義は、決して軽視すべきではないし、軽視できるものでもない。

中国の大規模古典文献データベース

中国古典文献の大規模データベースについては、本誌バックナンバーや『電腦中国学』シリーズなどでこれまでに何度も紹介してきたが¹⁾、行論の都合上、改めてまとめておく。

大規模古典文献データベースの嚆矢となったのは台湾中央研究院の漢籍電子文献である。MS-DOS 時代の 1984 年に構築が始まり、1990 年代のインターネットの普及期にいち早く二十五



漢籍電子文獻

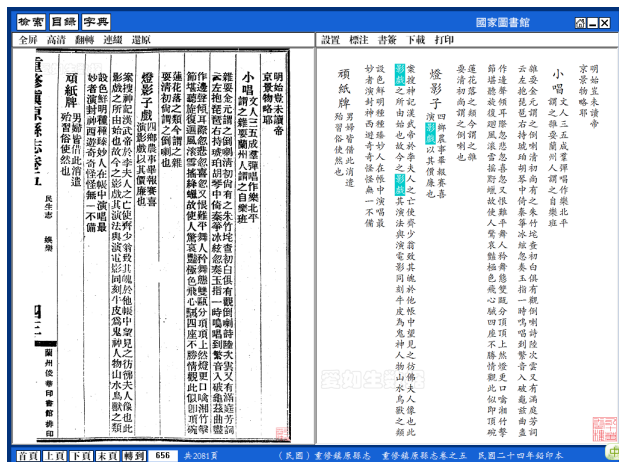
史をはじめとする古典文献検索を無償提供し、古典文献全文検索データベースの威力を大いに知らしめたものである。

現在、同データベースの規模は 5 億字弱で、主要な中国古典文献を網羅している。ただし、全てのデータを使うには、ユーザー登録あるいは機関登録が必要となる。中国本土の大規模古典文献データベースがいずれも有償であるのに対して、インターネットを通じて無償公開されていることは特筆に値し、我が国の中国学研究者の間でも広く使われている。

2000 年にはミレニアムプロジェクトとして、中国で文淵閣『四庫全書』の全文検索版、約七億



『四部叢刊』



愛如生『方志庫』

字が作られた。その作業を担った北京の書同文社は、引き続き『四部叢刊』『十通』などを開発した。

2006 年には北京の愛如生社が、合計文字数が 17 億字に達する『中国基本古籍庫』を発売、引き続き方志庫・道藏・叢書・別集など、いずれも十億字を超える規模のデータベースを開発し、古典籍のデジタル化を網羅的に進めている。

このほか、瀚堂典藏^[2]も重複が多いとはいえ、合計 20 億字にもものぼる古典文献や近代の新聞資料等をデータベース化している。

一方、2014 年は、信頼性の高い中華書局の標点本を底本とする約 2 億字規模のデータベース、中華經典古籍庫が発売されている^[3]。

中国本土のデータベースの特色

2000 年以降、主に中国本土で続々と開発・発売されている大規模古典文献 DB には、いくつかの共通する特色がある。

まず、公的助成を受けて作られているものが多いにもかかわらず、企業が制作・販売している点である。我が国では、科研費等の公的助成を受けて作成したデータベースについては、無償で社会に還元することが求められるため、企業と一体化した開発・販売などを行うのには困難が伴う。しかしそのため、人文学情報処理に詳しい研究者個人に、データベースの制作・運用の負担が集中し、往々にしてメンテナンスが充分に行われなくなる

し、研究者が移籍・退職した後のデータベース運用が保証されない。この点、中国のやり方であれば企業の商品としてメンテナンスが行われるため、研究者個人が運営するより継続的なメンテナンスが期待できる^[4]。

データベースの規模は、いずれも億字単位の巨大なものであるが、大規模な文献を効率的に電子テキスト化するため、底本を校訂せずにそのまま入力している。そして異体字の揺らぎについては、検索に異体字・同音字などの同一視検索機能を持たせることで対応している。また複雑な検索条件設定には対応していない

(後述)。

これは、『四庫全書』電子版の制作時に書同文社によって確立された大規模文献データベース構築手法である。愛如生社の文献デジタル化システムは書同文社から購入したものであり、中華經典古籍庫の開発元である北京創新力博社は書同文社の元エンジニアが起こした会社であるなど、中国の大規模古典文献データベースの大半が書同文社の何らかの影響下に制作されていることが反映されている。

大規模古典文献 DB 導入の問題点

決定版の不在

これらの大規模古典文献データベースが中国学の研究上、必要不可欠であることは論を俟たないが、しかし我が国の大学や研究所・図書館などの機関にとって、その導入は決して簡単なことではない。

大学・機関の図書・データベース購入予算には限りがあるので、決定版といえるデータベースを 1 つ導入するだけで済みたいところであるが、しかしその決定版が存在しない。

収録文献の量については、『中国基本古籍庫』が他者の追随を許さない。しかし『中国基本古籍庫』の全文検索用電子テキストは「通行本」に基

づいており、しかも底本とされるテキストと字句が異なる例が見られるなど検索結果の信頼性にやや欠ける。また、書籍によっては序跋や注釈などが電子テキスト化されておらず、影印が収録されているが全文検索できない例が多々見られる^[5]。

文献データベースの利点の1つに、用例を発見できるだけでなく、用例が無いことを証明できる点が挙げられるが、残念ながら『中国基本古籍庫』の電子テキストの品質では、検索できない用例が存在する可能性が排除できないため、用例を発見するためにしか使えない。その上、収録される書籍の影印の画像がきわめて粗く、読めないものさえある。これらは、全文データベースがいかにあるべきか、その基本方針が十分に練り上げられていなかったことを示唆している。高額な大規模データベースでありながら、いささか残念である。

文淵閣『四庫全書』は世界史の教科書でも言及される清朝の優秀な国家的プロジェクトである、という程度の認識で使っている学生も多いようだが、抄写ミスが多くテキストの信頼性が低いというのは中国古典文献を扱う者の常識である。それに加えて、我が国の多くの機関で導入されている2000年に発売された『四庫全書』v.2では、表が検索できるようになっていない^[6]。このため、やはり用例の不在を証明できないし、また決定版データベースと見なすこともできない。

本文のクオリティという点では、中華書局の標本本に基づいた『中華經典典籍庫』が最も優れている。しかし、合計2億字程度と収録される文献の量が限られており、地方志や戯曲・通俗小説なども含まれていないため、やはりこれ1つだけで事足りるものではない。

≡ データベースの価格

このため、大学・機関は複数のデータベースを買いそろえなくてはならないことになる。ここでネックになるのが、データベースの価格である。

数十万円で導入できるのは『四庫全書』くらいで、その他はいずれも数百万から数千万円もの価格である。愛如生製品に至っては、全て揃えるのに一億円を超える予算が必要になる。

	文字数	価格
四庫全書 v.3.1	約 10 億字	約 420,000+216,000 円 / 年～
基本古籍庫	約 20 億字	約 25,000,000 円～
愛如生大型数拠庫	各 約 10 ～ 20 億字	約 15,000,000 円 (スタンドアロン) 約 1,500,000 円 / 年 (Web)
中華經典典籍庫	約 2 億字	約 3,000,000 円～

中国の国家重点大学であれば、自国の歴史や文化を研究するために必要なツールであるので、予算を確保して購入するのはある意味当然といえる。しかし我が国の中国学は、人文学の中の大きいとはいいがたい一分野に過ぎず、一般に予算も潤沢ではない。その潤沢とはいいがたい予算の中で、紙版の書籍・雑誌のほか電子ジャーナルや新聞データベースなども購入・契約しなくてはならないのだから、百万円・一千万円にもなる古典文献データベースを簡単に購入できるところは、ほとんどないだろう。

≡ データベース導入のネットワーク化

前述のように、大規模古典文献データベースは、情報化時代における中国学の研究インフラであるといえる。そうであるならば、インフラの整っていない国や地域の経済発展が望めないのと同様、データベース導入の成否が大学の研究環境を決定し、研究水準を規定するようになると考えられる。

しかし、日本の現状では、中国の大規模古典文献データベースを大学・機関がそれぞれ網羅的に買いそろえることがほぼ不可能である。ともなれば、これらのデータベースを存分に使いこなせる中国に比して、我が国の中国学はそもそも研究インフラにおいて太刀打ちできず、落後してしまうことになりかねない^[7]。

この問題は、おそらく複数の大学・機関の連携によって解決するしかないと思われる。このとき、特定の大学・機関が中国学のデータベースを集中的に導入するよう取り決める方法もありうるが、

むしろ大学・機関の連繋・ネットワーク化を進めて、東京・関西などの地域単位で大規模データベースが網羅されるように導入するのが現実的だろう。

例えば、学部教育レベルで必要となるデータベースとして『四庫全書』・『中華經典典籍庫』などを各大学が買いそろえ、より高価な大学院生・専門研究向けデータベースを各大学の得意とする分野にあわせて分担購入する、といった方法が考えられる。その前提として、図書館相互利用協定の拡充なども必要になるだろう。従来も大規模叢書については、関東・関西などの地域ごとに、他の大学が買ったので我が大学は見送る、といった調整が事実上行われてきた。同様のことを、データベースについても行えばよい。

ところで紙の書籍についても、日本では各大学・機関の図書館がそれぞれに選書して購入しているため、いずれも似たような蔵書構成となる。また、中国書籍を網羅的に収集する図書館がないため、一般向け書籍やマイナーなジャンルの研究書などで、どこも購入していないものが出てきてしまう。中国において学術書を含め出版点数が増加している現状にあって、これについても、理想的には大学間で集中購入する分野を取り決めるなどして対応してゆくべきであろう。

紙の書籍にせよデータベースにせよ、中国学の資料の収集・導入は、もはや大学・機関が単独で対応できる時代ではなくなっており、早急に対策を講ずる必要があると考える。

情報収集のスキル

Ⅲ 大規模古典文献データベース検索の問題点

大規模古典文献データベースにも、しかし多くの問題点がある。

底本とテキストの質の問題については前述したが、そのほかに例えば検索機能がある。中国の大規模古典文献データベースの多くは、単純な検索機能しか備えていないものが多い。書同文の製品や『中華經典典籍庫』は and・or 検索程度の条件設定にしか対応していないし、愛如生の製品は

and・or・not 検索に対応しているものの、同時に検索できるのは 2 語に限られる。例えば書同文社のデータベースが「A %n B」という条件式で、2 つの語が出現する範囲の字数を設定できるように、個別の検索式が設定されている例はあるものの、総じて複雑な検索には対応できない。

また、検索結果はヒットした書籍名・篇名などのリンケ一覧として表示されるのが一般的である。データベースによっては語句がヒットした箇所の前後の文も表示してくれるが、1 ページに表示される検索結果数には限りがある。このため、データベースで検索した結果から、二次的に用例・事例などのカード型データベースを作成して研究を進める、というような使い方がしにくい。言い換えれば、ある語句の出現箇所を調べる以上の使い方が想定されていない。

前述のように、中国の大規模古典文献データベースは、いずれも底本を校訂せずにそのまま電子テキスト化し、異体字・同音字などの同一視テーブルによる検索によってテキストの字体の揺れを吸収する方式を採っているが、その漢字同一視テーブルが企業によってまちまちなのも問題である。『中国基本古籍』のテーブルは簡体字・繁体字の同一視検索にしか対応していないため、例えば「總」・「總」・「総」などで検索すると、それぞれまったく異なる結果になる。このため、ユーザーの側が Unicode に収録されている漢字の知識に基づいて異体字を逐一指定して検索する必要がある。これは愛如生社が書同文社から文字入力システムのみを購入しながら、その文献デジタル化思想への理解が不十分であった結果、齟齬を生じたものといえよう。

台湾中央研究院の漢籍電子文献は、not 検索その他、比較的柔軟な検索条件設定が可能で、しかも検索結果の一覧出力に対応しているなど、研究者の実際のニーズに沿ったデータベース設計が行われている。その一方で、文字コードが Unicode の URO のみの対応である、おそらく著作権の問題から底本となった評点本の校注・注釈などが入力されていない、といった問題もある。

総じて大規模古典文献データベースは、それぞ

れに独自の機能や制約、検索のクセなどがあり、それらを的確に把握しておかないと、調査効率が悪くなるばかりか、検索漏れといった事態をも起こしかねない。そうしたノウハウはデータベースのマニュアルには必ずしも記載されていないので、ユーザーサイドで情報を蓄積するとともに、研究者のみならず学生にも、それらを周知あるいは教育していく必要がある。

中国の人文情報教科書・書籍

それでは、大規模古典文献データベースの利用を前提とした中国学の基礎教育はいかにあるべきであろうか。この問題を考えるにあたり、それらデータベースの本場である中国本土の状況について見ておくのは決して無駄なことではあるまい。

この問題を考える上で参考になると思われるのが、中国で出版された人文系学生向けの情報教育教科書、あるいは初年次教育教科書である。その内容を検討することで、中国において大規模古典文献データベースの利用について、何がどの程度教えられているのか、また何が問題であると認識されているのかをうかがい知ることができると思われるからである。

筆者がこれまでに入手した同種の教科書は、以下の4種である。

- 《文献信息检索与论文写作》第2版, 花芳编著, 清华大学出版社, 2014
- 《文献信息检索与论文写作》第四版, 王细荣、韩玲、张勤编著, 上海交通大学出版社, 2013
- 《文献信息检索与利用》, 王丽萍主编, 华南理工大学出版社, 2013
- 《电子文献检索教程》, 胡光林、李雪萍主编, 北京理工大学出版社, 2010



『中国基本古籍庫』の「総」と「總」の検索結果



《电子文献检索教程》

このうち、前の3者は一般的なOPAC・論文データベースなどの使い方解説が主であって、古典文献データベースには言及していない。一方、『電子文献检索教程』は〈第八章 古籍資料的检索〉で、古典文献の調査方法について、アナログとデジタルを取り混ぜて解説している。その目次を以下に引く（括弧内は筆者訳）。

- 第一节 书目的利用（図書目録の利用）
- 第二节 类书的利用（類書の利用）
- 第三节 政书的利用（政書の利用）
- 第四节 索引的利用（索引の利用）
- 第五节 数字化资源的利用（デジタルリソースの利用）

以上のように第1～4節では伝統的な工具書による調査方法が解説されており、第5節でデジタルリソースの活用方法について述べている。

全体として、第1～4節では中国学のリファレンスに不可欠な目録・類書などの使い方をオーソドックスに解説しており、中国学の入門にふさわしい内容になっている。一方で、そうしたアナログ調査方法のうち、デジタルの手法に置き換え可能なものについての言及が必ずしも十分ではない。例えば叢書の子目検索について、同書では『叢書綜録』の使い方を解説しているが、台湾の全国図書書目资讯网（NBI Net）^[8]、あるいは我が国の全国漢籍データベース^[9]など、オンラインで叢書を検索できるサービスが多々ある現状を考えると、最善の方法とはいいいがたい。類書についても、台湾でオンライン公開されている『古今圖書集成』に言及していない。

第5節では、各種データベースやWebサイトについて概説している。主立ったリソースが網羅されているし、また後半部では、データベース検索の注意点についても言及しており、データベースに収録されている書籍の版本に注意しなくてはならないこと、また検索に際してキーワード指定の工夫が必要であり、「李白」・「李太白」・「太白」だけでなく「李杜」までも検索する必要があることなどを、具体的に解説している。

一方で、前述の個々のデータベースの検索のクセや異体字処理機能の違いなどには言及していない。学部生レベルではそれでも良いのかもしれないが、大学院生以上の研究レベルを考えた場合、データベースから遺漏なく情報を引き出すのにはいささか解説が不足していると言わざるをえない。

ここから類推すると、中国国内でも大規模古典文献データベースの利用方法に関するノウハウの集積、およびデータベースを前提とした教育は、まだ十分に確立されていないようである。

中国学に必要な検索スキル

前掲の中国の人文系情報教科書と我々の『電腦中国学』シリーズとの大きな違いの1つに、正規表現の解説がある。『電腦中国学』シリーズでは、ネットから取得するなどした電子テキストの分析・検索方法として、正規表現による検索を解説している。我が国では、例えば『文系のための情報処理入門』^[10]・『講座 IT と日本語研究』シリーズ^[11]などがいずれも正規表現を取り上げていることから、人文情報系の研究者の間で正規表現は必須のスキルであると認識されていると思われる。しかし、中国の教科書には正規表現への言及はおろか、テキストファイルの検索についての解説も見られない。

この原因として考えられるのは、中国の人文学における情報処理が事実上、Windows 95の登場を契機とするGUI時代に起源することである。DOS時代の1980年代、中国はまだ経済発展の途上であり、当時まだまだ高価であったパソコンの普及は進んでおらず、Grep・Perlなどのツールが知られることもなかった。人文系研究者は、南巡講話以降、1990年代後半になって初めてパソコンに触れた人が大多数であり、そのためWindowsとInternet Explorer・Officeこそがパソコン、サーチエンジンやWord・メモ帳などの検索機能こそが検索、という感覚が定着したものと推測される。

正規表現の柔軟な検索の便利さを知らなければ、and・or検索程度の機能しか持たないOPACやサーチエンジンこそが検索であると考え、その程度の機能しか要求しなくなるのは自明である。つまり、

正規表現の不在が各種データベースの検索機能の貧弱さに繋がっており、また上掲の教科書における検索条件設定やテキスト検索に関する解説の薄さに結びついているものと解釈できる。

翻って、正規表現のスキルは中国学研究者にとってどの程度必要なのかについても、一考の余地がある。正規表現を習得することで、検索やデータ加工の精度や効率を高めることができるのは事実であるが、実際問題として大多数の中国学研究者は正規表現を身につけていない。それ以前に、正規表現検索の前提となるテキストファイルを、独自に収集・作成・整理している研究者はほとんど居ないし、正規表現のスキルは大規模古典文献データベースの検索では役に立たない。つまり、情報処理スキルがごく平均的な多くの中国学研究者にとって、正規表現の習得はコストとベネフィットが見合っていないのであろう。逆に、古典文献の全文データベース化が進んでいないからこそ、我が国では青空文庫などから必要なテキストを収集して独自に検索を掛ける必要があり、エディタや検索ソフトなどを使った正規表現による検索が必須のスキルとなっているのかもしれない。

ともなれば、大規模古典文献データベースの普及が進む現状下においては、正規表現よりもむしろ、それらを使いこなすための検索スキルを優先的に身につけるべきかもしれない。このとき、異体字検索の問題をクリアするために、Unicodeの仕組みと収録漢字、さらにコードセパレートや異体字に関する十分な知識が必要となる。つまり、コンピュータで使いこなせる漢字の知識とその入力スキルという、ある意味パソコン多漢字・多言語処理の基礎中の基礎こそが、大規模古典文献データベースの使いこなしにおいても最も重要な基礎的スキルであることになる。その上で、台湾中央研究院漢籍電子文献などのフリーデータベースの検索方法や、各大学・機関などに導入された大規模古典文献データベースの検索条件設定を具体的に習得していくべきであろう。

一方、正規表現のスキルについては、そのベネフィットとともにプレゼンテーションする必要がある。前述のように、正規表現による検索スキ

ルは、テキストデータの構築・集積・整形などの作業を前提としている。そしてそうした作業は情報収集よりもむしろ、テキストデータ、ないしはデータベースの公開といった情報発信と、より密接にかかわっているといえるのではないか。ともなれば、中国学の領域において正規表現は、学術情報発信の必要性とともに、大学院の演習などで教授されるのがふさわしいではなかろうか。

■ おわりに

中国学の研究において、大規模古典文献データベースはもはや研究インフラというべき、必要不可欠の存在になっている。一方、複数の高額なデータベースを導入するのは、単独の大学・機関の手には余るので、ネットワーク化などの方策を講ずる必要があると思われる。しかしその一方で、我が国においても中国においても、それらを使いこなすためのノウハウは十分には知られておらず、今後、大学・大学院における教育等を通じて、教授されてゆくべきであると思われる。

ところで、それら大規模古典文献データベースを利用する上でのノウハウについては、導入された大学・機関でしか検証作業ができないこともあり、情報の蓄積も発信も充分であるとはいいがたい。『電脳中国学入門』も、データベースの紹介に主眼を置いたため、検索方法の詳細や使いこなす上でのノウハウなどに関する解説は詳しくない。

また、古典文献大規模データベースの評価方法の確立も課題であろう。紙媒体の文献の場合は、主に校訂の精度によって評価がなされるが、データベースについては宣伝的な記事が目につく一方で、客観的かつ学術的な批評はあまり見られない。実際、前に言及したように、大規模古典文献データベースは、大規模叢書・通行本・標点本など、底本の種類がさまざまであるほか、底本を電子テキストに起こす方法や電子翻刻ルールもまちまちである。

考えるに、データベースの電子テキストの質は、底本の選定とデジタル化の方法・正確性、2つの側面から評価されるべきである。また、動作速度・

検索機能・異体字テーブル・検索結果の出力方法などについても合わせて考慮しなくてはならない。そうした客観的かつ精密な評価方法を確立してこそ、各データベースをツールとして十分に活用することができよう。

いずれにせよ情報化時代にあって、中国学研究・教育の資料や方法は、大きな変化を迫られており、それに適応した新たな研究・教育基盤を早急に構築する必要があるだろう。

注

- [1] 『電腦中国学』（好文出版、1998）『電腦中国学Ⅱ』（好文出版、2001）『電腦中国学入門』（好文出版、2012）など。
- [2] <http://www.hytung.cn/>。詳細は本誌レビュー「瀚堂典藏」（p.114）参照。なお、本稿で言及した Web サイト・ページについては、いずれも 2014 年 9 月 15 日に最終確認している。
- [3] 詳細は本誌レビュー「中華経典典籍庫」（p.110）参照。
- [4] 企業であっても、経営が立ちゆかなくなるケースや、

研究投資がなされずに新たな OS や情報環境に対応できなくなるケースがあるので、あくまで相対的な特色に過ぎない。

- [5] 筆者のブログエントリ「結構微妙な基本古籍庫」（<http://www.wagang.jp/blog/logdata/eid71.html>、2006.09.13）参照。
- [6] 『四庫全書』は v.3.0 へのアップグレードに伴い、図表内の文字が検索できるようになり、また Unicode の Ext.B に対応した。
- [7] 例えば、中国基本古籍庫導入大学・機関は中国：88、台湾：22、香港：8、日本：8、韓国 7 である（愛如生社のウェブページ <http://www.er07.com/front/jiaodian/kehujsp?name=30> による。2014/09/03 確認）。
- [8] <http://nbinet.ncl.edu.tw/>
- [9] <http://kanji.zinbun.kyoto-u.ac.jp/kanseki>
- [10] 中村康夫・安道百合子著、泉書院、2008 年。
- [11] 荻野綱男・田野村忠温編、明治書院、平成 23～24 年。『2 アプリケーションソフトの基礎』・『4 Ruby によるテキストデータ処理』などで正規表現が取り上げられている。

学術データベースの構築と公開

師 茂樹（もろ しげき）

はじめに

本報告をはじめるとあって、筆者の個人的な体験を記すことをお許しいただきたい。いずれも、外国人研究者との対話の中での話である。

一つは、欧米の研究者の会話を小耳にはさんだことから始まる。その会話では、日本にある古文庫の話となり、それについて一方の研究者が「日本には文字資料が少ない。なぜなら日本の文化はビジュアルが得意だから」という発言をした。文脈から推測すると、「ビジュアルが得意」という発言の裏には、現代日本のマンガ文化と、それが『鳥獣人物戯画』などの前近代の作品に淵源する、という歴史観¹⁾が念頭にあるようであった。筆者は、その会話に割って入って、日本に多くの文字資料が残されていることを力説してみたが、語学力の問題か、相手を納得させることができなかった。

もう一つは、台湾の仏教研究者とのやりとりである。この研究者は、日本に調査に来た際、日本に多くの写本・刊本が残されており、その中には中国語圏で知られていない注釈書等があることをたまたま知り、「日本にこんなに重要そうな文献があるとは知らなかった」と興奮気味に語っていた。そして、どうして日本人はこれを使って研究しないのだ、どうしてこれらの文献を出版しないのだ、と述べていた。その後筆者は、その研究者の求めに応じて日本古典籍総合目録データベース

（国文学研究資料館）²⁾ や日本古写経データベース（国際仏教学大学院大学）³⁾ など、いくつかの関連データベースを紹介したが、外国の研究者だけでなく、日本の研究者にとっても、日本の古文献を調査するということがそれほど容易ではないことを痛感することとなった。

ここで紹介したごくわずかな例を、筆者のバイアスが多分に入っていることもあわせて、安易に一般化することは避けるべきであろう。しかし、筆者がこの二つの発言に接した時に感じたことは、日本の古文献がいかに海外で知られていないか、ということである。そして、一つ目の例は極端な例かもしれないが、いずれ世界中の研究者の共通認識になってしまうのではないか、ということである。

現在、多くの研究者が情報収集のためにインターネットを用いている。特に、国外の資料の調査においては、インターネット上のリソースを用いることが当たり前のプロセスとなっている。研究者であれば、学部生のようにインターネットで調べるだけで満足するのではなく、各地にある図書館に足を運び、索引のない紙の目録を1ページずつめくってでも調査をすることが当たり前、と言われれば、理念としてはまったくそのとおりである。しかし現実には、インターネットで、もっと言えば Google でアクセスできない資料の調査が後回しになるのは否めないし、場合によっては調査できなくてもしかたがないと思われたり、そもそも資料自体が存在しないと思われたりしても

しかたがない、という状況になりつつあるのではないだろうか。

インターネットを通じていかに情報を発信するか、というのが、これからの研究のあり方を変えてしまう可能性がある。本報告では、どのように研究のためのデジタルリソースを構築し、公開すればよいのかについて、中国学を中心に若干の考察を行いたい。

■ 日本のデジタルリソースの特徴とその背景

≡ 中台日の代表的デジタルリソースの比較

現在、日本から発信されている中国学関連の学術データベースが、どのような特徴を持っているのかについて考えるための材料として、下表を作成した。これは、中国・台湾・日本で公開されている中国学関連の代表的なデジタルリソースを、開発・公開元の種類によって大まかに分類したも

のである。中国学関連のデジタルリソースは膨大な数になるため、ここでは漢字文献情報処理研究会編『電脳中国学入門』（好文出版、2012年）で紹介されているもののなかで、前近代に関するものを中心にとりあげて、全体を代表させている。したがって、きわめて大雑把な表であることはご了承いただきたい。

表では、中国・台湾・日本を縦の列に、大学や図書館といったデジタルリソースを開発・公開している担い手を横の行にして、マトリックスにしている。担い手についても、実際には、大学の教員が中心になって行ったプロジェクトの成果が、営利企業を通じて市販されているような場合もあり、どの欄に入れるべきか迷うような例もある。ここでは、大雑把な傾向を見たいと思っているので、複数の欄にまたがるようなものであってもどれか一つに寄せてしまっている。

表中にゴシック体で書かれているのはテキストデータを持っていて全文検索ができるもの（一部しかできないものも含む）、明朝体は画像しかなくて

	中	台	日
大学・研究機関		<ul style="list-style-type: none"> 漢籍電子文献(中研院) 文物図像研究室資料庫(中研院) 	<ul style="list-style-type: none"> 東方学デジタル図書館(京大人文研) アジア古籍電子図書館(東大東文研) 東洋文庫オンラインライブラリ 拓本文字データベース(京大人文研)
学術団体・学会等		<ul style="list-style-type: none"> CBETA 中華電子仏典協会 	<ul style="list-style-type: none"> 大正新脩大藏經テキストデータベース
図書館	<ul style="list-style-type: none"> 在線数拠庫(国家図書館) 		<ul style="list-style-type: none"> 近代デジタルライブラリー(国会図書館) 京都大学電子図書館貴重資料画像 古典籍総合データベース(早大図書館)
公文書館			<ul style="list-style-type: none"> アジア歴史資料センター(国立公文書館)
企業	<ul style="list-style-type: none"> 超星網 文淵閣『四庫全書』電子版 『四部叢刊』など(書同文) 『中国基本古籍庫』(愛如生) 国学網 	<ul style="list-style-type: none"> 故宮東吳古今圖書集成 	
個人	<ul style="list-style-type: none"> 中国哲学書電子化計画 	<ul style="list-style-type: none"> 故宮【寒泉】古典文献全文検索資料庫 網路展書読 	<ul style="list-style-type: none"> 中国古典戯曲データベース

全文検索ができないものである。

この表から中国学関連の日本のデジタルリソースの特徴を分析するのは乱暴かもしれないが、次のようにまとめることができるのではないと思う。

- 日本は、画像データが主流
- 日本は、大学や図書館が主体

一つ目の特徴は、中国・台湾ではテキストデータが主流であるのに対照的であり、冒頭に紹介した「日本には文字資料が少ない」という海外研究者の印象と重なる部分であろうと思われる。

筆者は画像データベースの意義^[4]を否定するわけではないが、そもそも資料の存在を知ることがなければ、資料の画像データがあったとしてもそれを研究・教育において活用することはできないだろう。

言うまでもなく画像データベースは、本文を検索することができない。多くの画像データベースにはメタデータが付されているが、作者名やタイトルにあるキーワードによって検索することはそれほど多くないと思われる。

筆者は論文の一部を Academia.edu で公開している^[5]。Academia.edu には Analytics というサービスがあり（図1）、公開している論文に対してアクセスしてきたユーザについて、プロフィール閲覧件数・論文閲覧件数・論文ダウンロード数・国別・検索キーワード・外部リンクなどの統計情報を知ることができる。本報告の執筆時点（8月18日）から過去30日について、筆者の論文に対する検索キーワードを見てみると、次ページの表のとおりとなる（いずれも Google や Yahoo! からの検索。ちなみに日本からのアクセスは Yahoo! 経由が多い）。

表中、ゴシック体で表したものが論文タイトルにあるキーワードであるが、検索キーワードのほとんどが論文タイトルと関係がないということわかるだろう。中には「蔵以 読み」「却庵妄記 十眼房被申云」のように、文献中に出てくる表現をそのままキーワードにして検索していると思われる

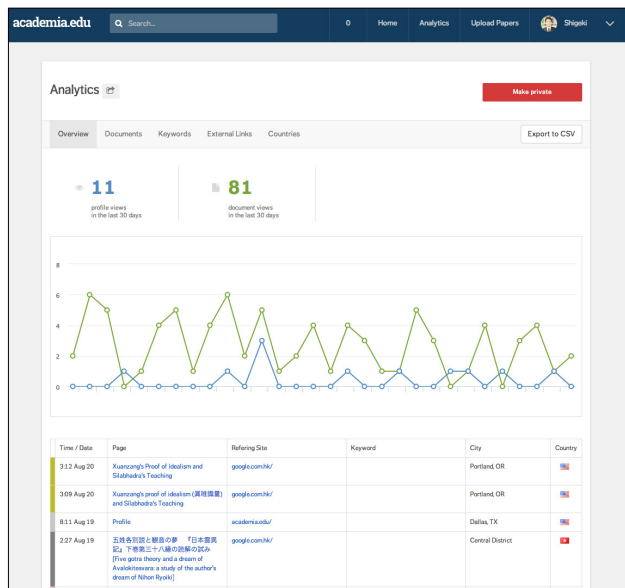


図1

例も見られる。以上は、あくまで筆者一人がインターネットで公開している論文に対するアクセスを例にとって述べたものである。性急に一般化することは避けるべきであろう。しかし、このような傾向は、ウェブ上の古典籍にアクセスするユーザについても言えるのではなからうか。

このような検索をするユーザが、メタデータしかない未知の画像データベースにアクセスするのは難しいであろう。古典籍、写本や刊本などのメタデータになると、論文のタイトルよりもより具体性を欠き、標準化されていない短いものになってしまうだろうから、検索される確率がより下がってしまうことになると思われる。

未翻刻資料が、たとえ多少の間違ひはあろうとも翻刻され、またそれがテキストデータとして検索可能な状態でインターネット上に公開されてこそ、画像データベースの存在が知られることになる。その意味で、画像データベースが主流の日本のデジタルリソースが、存在が知られないままであるというのは、テキストデータの不在が大きな要因となっていると思われる。

さて、中国学関連の日本のデジタルリソースの二つ目の特徴としては、大学や図書館が主体であるということがあげられる。一方で中国は企業が

検索キーワード	論文タイトル
元曉 pdf	元曉の唯識比量解釈 E. Franco 氏の説と比較しつつ
元曉 pdf	新羅元曉の三時教判批判 『大慧度經宗要』を中心に
五姓 差別	五姓各別説と観音の夢 - 『日本霊異記』下巻第三十八縁の読解の試み
五姓閣	
日本霊異記	
1 1 面神呪心経ぎりゅう	
蔵以 読み	楞嚴経惟懃疏の逸文をめぐる二、三の問題
却廢妄記 十眼房被申云	
夫靈蹤哉則 宗 尋容像不	
義寂	
唯識 凡天	義寂と新羅の唯識思想
99 人の最終電車 あらすじ	電子書籍の／とインタラクティビティ
電子書籍 インタラクティブ系	
僧 懷素 図	相部律宗定實の行状・思想とその日本への影響 『四分律疏飾宗義記』に見える仏身論を中心に
律宗綱要	
中国書籍 唐僧懷素伝	
律宗 日本 伝来	
八定の賓	
四禪の客八定の賓	
摂入	法相宗の「一乗方便」説再考 諸乗義林を中心に
法相華	
法花摂釈	撲揚智周伝についての二、三の問題
大西祝	大西祝の因明理解
善相法師？	占察經の成立と受容 なぜ占いが必要とされたのか
なぜ占い必要	
加我聞	徳一の「如是我聞」訓読をめぐる二、三の問題

中心であり、台湾では研究者・研究機関によるものの割合が比較的多い。

図書館司書・図書館員は、文献を利用者に提供するサービスの専門家・研究者ではあるが、実際に文献を読んで研究する研究者ではない。したがって、図書館が主体となるデータベース開発は、近現代日本の図書や雑誌などを除くと、自ずと画像データベースが中心となるのではないと思われる。図書館司書等が担っていることが問題なのではなく、台湾のように研究者がデジタルリソースの公開・開発に携わっていないことが、結果的に画像データベースを増やし、海外から日本のデジタルリソースが見えにくくなるような状況を作

り出しているのではないかと懸念されるのである。

中国の企業が作っているテキストデータは、これまで何度も指摘されているように学術的な信頼性の問題がある（中華書局がデータベース開発に参戦するというニュース¹⁶は、この評価を覆す可能性があり重要である）。このような問題はあるものの、彼らは自社でテキストデータベースを開発するだけでなく、そのために培った技術やシステムを用いて、刊本等を含む古典文献のデジタル化業務を受託する、というビジネスを行っているので¹⁷、今後テキストデータベースはどんどん増えていく可能性があるだろう。

Ⅲ 学術振興政策との関連

このような日本の状況の背景には、文部科学省などによる科学技術・学術振興政策もまた、大きな影響を与えていると思われる。

近年、科学研究費などでは、研究成果の社会還元ということが強く言われている。それを反映してか、人文系の分野でも、単なる研究ではなくデータベース開発・公開ということを掲げて申請すると通りやすい、というようなことがまことしやかに言われるようになった。つまり、ごく少数の専門家しか読まないような専門的な論文を書くよりも、データベースを公開するほうが比較的社会的に役に立つのだ、という発想である^[8]。

一方で、研究者の研究活動の評価という点では、相変わらず論文、特に国際的な学会やシンポジウムでの論文、あるいは査読付きの論文などといった、論文中心の評価が行われている。中国古典の研究においては、しばしば文献の翻刻や校訂、訳注などを大学の紀要で発表するが、そのような研究成果はどれほど有益であっても現在主流の評価基準では論文として評価されないような状況である。

グローバル COE のような最近の大型研究助成においては、大学院生やポスドク、若手研究者に論文を書かせ、博士号をたくさん出すことを要求するようになってきている。かつては、大学院生を含めた共同研究グループによる文献の翻刻・校訂・訳注・索引作り等の活動などが、結果的に大学院生らを鍛え、(ICT 以前の) インフラを構築していたように思われる。しかし、短期間で論文を書くことを要求するあまり、大学院生にそのような長期間にわたる作業を分担させることが難しくなってきた。

論文を増やせ、でも社会還元もしろ、という要求が、あまり負担のかからない、場合によっては業者に発注するだけで済む画像データベース公開の増加の背景の一つにあるのではないだろうか。

現在の学術振興政策は、安倍晋三首相の「今後10年で、世界大学ランキングトップ100に10校ランクインを目指します」^[9] という発言からまわ

かるように、グローバルな学術の世界における日本の順位をあげていこうというものだと思われる。人文学・東洋学の学術のあり方と、大学ランキング的な評価とがそもそも共存可能なのか、という問題は脇に置くにしても、これまで述べてきたような「日本には文字資料がない（から研究する価値がない）」というような理解を生み出すような状況においては、大学ランキングの上昇は望むべくもないのではなかろうか。

毎年大きなニュースになる「世界大学ランキング」の代表格、Times Higher Education (THE) 大学ランキング^[10] においては、次のような割合で点数化が行われている。

- International outlook（教員・学生・研究等の国際性）…7.5%
- Research（研究）…30%
- Citations（論文の被引用率）…30%
- Industry income（産業のイノベーションに対する貢献）…2.5%
- Teaching（教育・学習環境）…30%

安倍首相の方針では、「外国人教員の積極採用」「優秀な留学生の獲得」「海外大学との連携」「英語による授業のみで卒業が可能な学位課程の充実」「TOEFLの卒業要件化」など、「グローバル化」が前面に出ているが^[11]、THE 大学ランキングでこれに対応する International outlook は、実は全体の7.5%にすぎない。もちろん、「優秀な」外国人教員や留学生が首尾よく獲得できれば全体のランキングは上がるだろうし、英語の論文が増えれば論文の被引用率も上昇するであろう。

しかし、東アジア各国が積極的に行っているように、論文のインターネット公開を進め^[12]、論文の被引用率を上げたほうが費用対効果は高そうである。また、日本に残る古い文献もテキストデータベース化し、国内外の研究者に対する認知度をあげることで、これらの文献に対する研究が海外でも増えれば、結果的に海外の研究者が日本の研究者の論文を引用する割合も増えるのではないかと思われる。

■ 今後求められること

≡ 組織的なテキストデータ構築

では、日本から研究成果のインターネット発信を増やし、「日本には文字資料が少ない」というような誤解を払拭するためにはどのような対策をすればよいだろうか。ここでは次の3点について提案したい。

- 翻刻などへの評価
- ポータルの構築
- Google reachable であること

1 点目はもっとも本質的な問題である。翻刻とは、一般的には写本や版本などを活字化することであるが、ここでは画像データしかない古文書や論文のような文字情報をテキストデータにする意味にも用いている。

現在、Google ブックスや CiNii では、書籍・論文を画像データとしてスキャンするだけでなく、OCR をかけることによって検索可能にしている。OCR によって機械的に作られたテキストデータには不正確な部分も多々あるが、コーパス開発や統計的分析などに用いるわけではなく、検索の際に一つでもキーワードにひっかかればよい、という程度の精度が求められているわけであるから、現在の水準でもそれなりに実用的であると言える。

しかし、同じ活字文献であっても、近代デジタルライブラリーで公開されている画像データのよう、画質が不鮮明であったり、また現代とは異なるデザインの活字や、旧字旧仮名使いで書かれていたりすると、現代の一般的な OCR では歯が立たなくなる。ましてや、手書きの文書、筆で書かれたくずし字などについては、実用的な OCR 結果に至るまでに技術的にいくつものハードルを超えなければならず、短期・中期的には実現が難しいように思われる^[13]。

現在多くの学界で一般的な、国際学会での発表や査読付き論文に偏重しがちな評価ではなく、翻

刻や訳注などにもそれなりの評価を与えるよう、東洋学をはじめとする人文系の学界による独自の評価基準作り、制度作り（古典文献を扱う博士論文においては、翻刻・訳注などを行う資料編を設けることが望ましい、というようなもの）が求められるのではない。翻刻が論文として発表され、それがオンラインで公開されれば、DTP で作られたテキストデータ付きの PDF として公開され、そうでなくてもしかるべき機関において OCR かけられる可能性が高い^[14]。

もっとも、論文や翻刻された文献がウェブ上で公開されたとしても、それがユーザの検索にひっかからなければ意味がない。筆者は、10 年以上前から自分の論文をインターネットで公開してきた。その際、Google などでヒットしやすいようにと、論文を HTML で書き直し、自分が管理する Web ページとして公開していた。一方、最近になって論文の PDF を公開できる researchmap^[15] や Academia.edu などのサービスができたこともあり、論文の抜刷を自炊して公開したりもしていた（つまり、一本の論文をいくつものサイトから公開していたのである）。テキストデータとしての精度という点から言えば、単に OCR をかけただけの researchmap などの PDF と比べて、Word で書かれた提出原稿を元にし、ある程度の校正もしている HTML のほうが正確であることは言うまでもない。

試みに 1998 年に発表した論文「法相宗の「一乗方便」説再考」^[16] を Google で検索してみると図 2 のような結果となった（執筆時）。

つい最近公開したばかりの Academia.edu が 1 位、続いてメタデータのみを公開する CiNii、3 位が論文の掲載誌である『印度学仏教学研究』のバックナンバーを公開している J-Stage、そして 4 位に筆者が 10 年以上公開している HTML の Web ページ、となったのである。論文ごとの個別ページ（URI）を持たない reaserchmap にいたっては、19 位となっているため図 2 には載っていない。これは、手間をかけて HTML を作っていた筆者にとって、いささかショックな結果であった。長い時間をかけて公開し続けた個人サイトよりも、最近出たばかりの学術系 Web サービスのほうが上位

だったからである。

わずかに論文の例ではあるが、単にデータを公開するだけでは不十分であり、Google (あるいはその他の検索エンジン) からの検索にひっかかるための工夫が必要である、ということは言えるだろう^[17]。そのためには、メタデータの整備は言うまでもなく、固定した URI・オープンアクセス化、そして学術情報であることを公的に裏付けけるような専門のポータルサイトなど、個人レベルを超えた国家的・組織的な環境整備が必要となってくるのではないだろうか。

個人レベルでのデータ開発・公開

海外から日本の研究成果や古典文献が見えにくい現在の状況は、日本の学術界全体を考えれば問題があるが、個人レベル (特に若手研究者のレベル) で見ると逆にチャンスと言えるかもしれない。自身の論文や、自身が扱う文献をインターネットでうまく公開することによって、海外の研究者からの注目を集めることができるためである。

また、本誌のレビューで何度も指摘されてきたように、中国の企業が作ったテキストデータベースは、圧倒的な文字数を誇る一方で、検索機能としてはごく単純なものとししか備えていない。ウィルドカードや正規表現を使った柔軟な検索、あるいは確率統計的なテキスト分析のためには、自身の研究テーマに即してカスタマイズ (文字処理やマークアップなど) したテキストデータベースを用意する必要がある。

また、個人レベルでのテキストデータ開発に類するものとして、Transcribe Bentham^[18] や翻デジ^[19] など、ソーシャル翻刻も注目される。

まとめ

以上、日本発のデジタルリソースは海外から見えにくいのではないか、という問題意識から、画像データベースからテキストベースへの移行を中



図 2

心にいくつか提案してきた。

海外の研究者から日本の研究成果や文献が見えるようになるためには、検索可能なテキストデータの整備が不可欠である。テキストデータベースの構築・公開のためには、研究者が実質的な担い手になるべきであり、研究者が積極的にテキストデータの構築に参加するためには、翻刻やテキストデータ開発を評価するような学界の体制作り (研究活動の評価基準、データペーパーなど) が必要である。また、データの公開のためには、個人レベルの公開ではなく、学術的なデータであることを保証するようなポータルサイトなどの環境整備も求められるであろう。

一方で、現在の過渡的な状況においては、積極的に自分の論文や研究している文献をインターネットで公開することで、国内外の研究者に自身の研究をアピールすることができるかもしれない。また、自身の研究に即したテキストデータは、自分でしか作れない。そのための個人レベルでのデータ整備・公開と、そのためのスキルアップが

求められるだろう。

謝辞

本報告は、日本学術振興会・科学研究費補助金・基盤研究 (B)「情報化時代における中国学次世代研究基盤の確立」(研究課題番号 23320010、代表者・二階堂善弘)、および同・基盤研究 (C)「未翻刻仏教文献に基づく東アジア仏教の基礎的研究」(研究課題番号 25370054、代表者・師茂樹)による研究成果の一部である。

注

- [1] 一例として、高畑勲『十二世紀のアニメーション—国宝絵巻物に見る映画的・アニメ的なもの』(徳間書店、1993 年)など。マンガの起源としての『鳥獣人物戯画』をめぐっては、夏目房之介「仮説・コマの発達史 マンガはいつからマンガになったのか」(『別冊宝島 EX マンガの読み方』宝島社、1995 年)、山本陽子『絵巻物の図像学 「絵そらごと」の表現と発想』(勉誠出版、2012 年)等参照。
- [2] <http://base1.nijl.ac.jp/~tkoten/about.html>
- [3] <http://koshakyo-database.icabs.ac.jp/index.seam>
- [4] 古典籍等の画像データベースについて、中村未来は「特に、器物や文献をそのまま画像データとして公開している場合、それらの資料を研究対象とする者にとって、非常に有意義な情報を迅速に得られるという利点がある」として、資料の翻刻に誤りがないかを確認する作業の効率化をメリットとしてあげる。加えて、「また、インターネット上におけるデジタル図版の公開は、教育面での活用も期待できる。貴重な器物や文献について、単に無味乾燥な文字を羅列して説明するよりも、実物の画像を示し内容を確認することによって、学生の理解度や興味が深まると考えられるためである」と述べ、教育上のメリットもあるとする(『日本における漢籍デジタル図版の公開状況とその意義』『新しい漢字漢文教育』58、2014 年)。
- [5] <https://hanazono.academia.edu/ShigekiMoro>。Academia.edu については拙稿「研究コミュニケーションツールとしての SNS」(『漢字文献情報処理研究』14、2013 年)参照。
- [6] 本特集の千田大介氏の記事を参照。
- [7] 千田大介「中国の人文情報処理企業の最新動向 中易中標と創新力博」(『漢字文献情報処理研究』6、2005 年)
- [8] 菊池信彦氏は Digital Humanities (DH) による研究成果(データベース等)について、「DH の成果物には、パブリック・ヒューマニティーズ (Public Humanities) としての可能性、すなわちアカデミズムの枠を超えて人文学知を市民に提供するアウトリーチリソースとしての役割が期待されている」と述べる(『E1583 - 持続可能なデジタル人文学のために：大学における支援の現状』『カレントアウェアネス -E』No.262、2014 年 7 月 10 日)。
- [9] 「平成 25 年 5 月 17 日 安倍総理「成長戦略第 2 弾スピーチ」(日本アカデミア)|平成 25 年|総理の演説・記者会見など|記者会見|首相官邸ホームページ」(http://www.kantei.go.jp/jp/96_abe/statement/2013/0517speech.html、2014 年 8 月 20 日最終確認)
- [10] <http://www.timeshighereducation.co.uk/world-university-rankings/>
- [11] 前掲「平成 25 年 5 月 17 日 安倍総理「成長戦略第 2 弾スピーチ」(日本アカデミア)|平成 25 年|総理の演説・記者会見など|記者会見|首相官邸ホームページ」。安部首相の発言の根拠となっているのは、教育再生実行会議による「高等学校教育と大学教育との接続・大学入学者選抜の在り方について(第四次提言)」(http://www.kantei.go.jp/jp/singi/kyouikusaiei/pdf/dai4_1.pdf、2013 年 10 月 31 日)である。
- [12] 本誌・特集 2 参照。
- [13] 中国のように、古い文献のテキストデータ化するための専門技術を開発することで、テキストデータを増やすべきではないか、という方法もあり得るであろう。Google によって情報のあり方が変わったように、革新的な技術が先行することによってものの方見方や社会における価値観などが変化することは大いにあり得るので、そのような技術が発達することでテキストデータに対する認識が変わる可能性はあるだろう。実際、中国に発注する、という手段が出たことによって、筆者も含めた日本の一部の研究者はテキストデータの作成を積極的に行いつつある。しかし、一般的には、ニーズがあるからこそ技術開発が行われるものである。テキストデータについても同様で、研究成果をインターネットで公開したりテキストデータを作った

りすることが評価をされる風土を作ることが先決なのではなかろうか。

- [14] 本稿の元となったシンポジウムの質疑応答において、大向一輝氏が紹介したデータペーパーは、この問題についての大きなヒントになるだろう。真板英一氏によれば「データペーパーは、研究データ自体を論文として受け付けることで、研究データそのものに学術的価値を認め、優れた研究データを整備し公表することを学術的業績として評価しようという制度」であり（「データペーパー投稿者のためのメタデータ作成ガイド」『日本生態学会誌』63-2、2013年）、日本の人文情報学系の学会でも議論が始まりつつある。シンポジウムにおけるディスカッションでは、データペーパーを論文と同等のものとして評価すべきか、という問題提起や、中国の企業などに発注することによって研究費さえあれば業績をいくらでも水増しできるのではないか、というような批判的意見も出された。これらの問題は今後議論解決していくべきであるが、学界をあげて前向きに検討すべきことであると筆者は考えている。たとえば、現状でも、だれでも読めるような版本を単に活字化する

だけでは、アクセスが困難な希少本でもないかぎり翻刻論文としての価値は認められない。それと同様に、OCRでそれなりの精度でテキスト化できたり、中国の企業に発注するだけで完成したりするようなデータはデータペーパーとしては認められず、校訂作業を行ったり、読むのが困難なくずし字を解読したりしたようなものをデータペーパーとして認める、といったような基準作りが、翻刻・訳注論文の評価基準作りと同様に、求められるのではなかろうか。

- [15] <http://researchmap.jp>. researchmap についても前掲拙稿「研究コミュニケーションツールとしての SNS」参照。
- [16] 師茂樹「法相宗の「一乗方便」説再考—諸乗義林を中心に—」（『印度学仏教学研究』47-1、1998年）
- [17] シンポジウムにおけるフリーディスカッションにおいて大向一輝氏は、Googleの学術情報担当者との会談の様子を紹介しつつ、筆者のこの推測に賛意を示してくれた。
- [18] <http://blogs.ucl.ac.uk/transcribe-bentham/>
- [19] <http://lab.kn.ndl.go.jp/dhii/omk2/>

学術情報共有プラットフォーム としての CiNii

大向 一輝（おおむかい いっき）

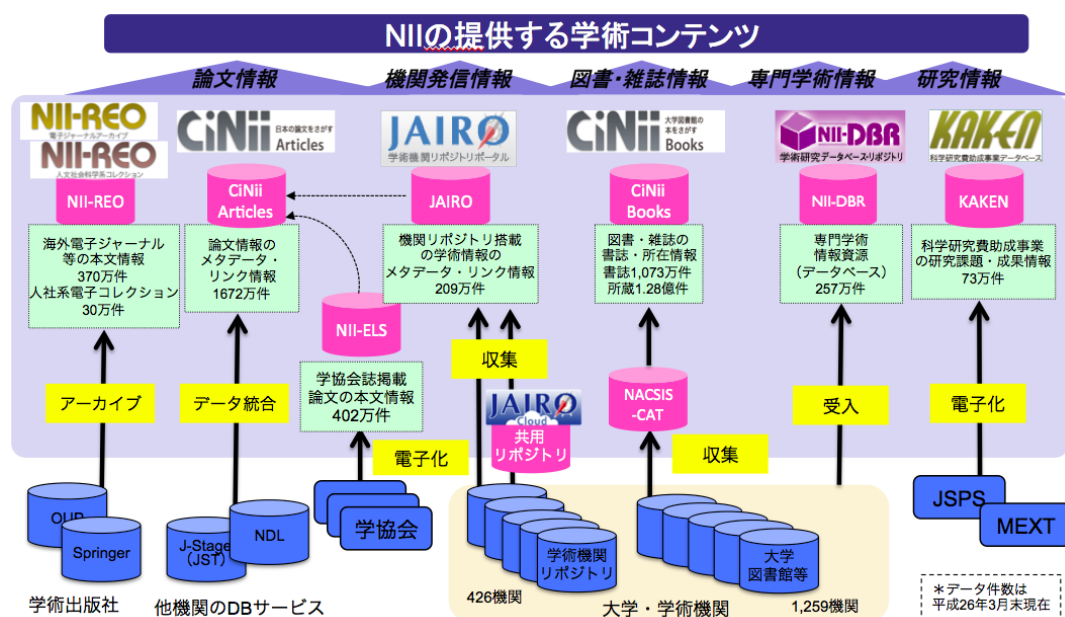
いて議論したい。

はじめに

国立情報学研究所が運営する CiNii は、国内最大規模の学術情報サービスとして人文科学、社会科学、自然科学の枠を超えて広く利用されてきた。本稿では CiNii が学術共有の枠組みの中でどのような機能を果たし、位置づけられてきたのかにつ

CiNii とは

CiNii の運営母体である国立情報学研究所（NII）は、2000 年（平成 12 年）に文部科学省学術情報センター（NACSIS）の改組によって立ち上げられた大学共同利用機関である。国立大学法人化と時



を同じくして 2004 年(平成 16 年)に情報・システム研究機構の一部となった。学術情報センターでは主な業務として国内の大学図書館・研究図書館を対象とした目録所在情報サービス NACSIS-CAT や情報検索サービス Webcat・NACSIS-IR などの事業を担ってきたが、改組の際に情報学の研究拠点としてのミッションが新たに与えられ、以降国立情報学研究所では車の両輪として研究と事業を位置づけている。事業面では情報サービスとともに大学間のインターネットサービスである SINET の企画運営が 2 本の柱である。

国立情報学研究所の情報サービスは多岐に亘っている(図 1)。前述の NACSIS-CAT ならびに本稿のテーマである CiNii に加えて、各大学・研究機関が運営する機関リポジトリの情報を収集し、検索機能を提供する JAIRO、機関リポジトリそのものをホスティングサービスとして提供する JAIRO Cloud、科学研究費補助金の報告書データベース KAKEN、海外電子ジャーナルの共同アーカイブ NII-REO、各種の学術情報データベースをホスティングする NII-DBR などがある。

CiNii は 2004 年に試験公開され、2005 年から本サービスを開始した論文情報サービスである。2011 年(平成 23 年)に NACSIS-CAT のデータを対象とした検索サービス CiNii Books が公開されたことを受け、CiNii Articles に改称された。本稿では CiNii Articles について詳述することとする。

CiNii Articles の基本機能は論文の書誌情報を対象とした検索・閲覧機能と、電子化された本文ファイルの提供機能に大別される。後者については、学術情報センター時代の 1997 年(平成 9 年)か

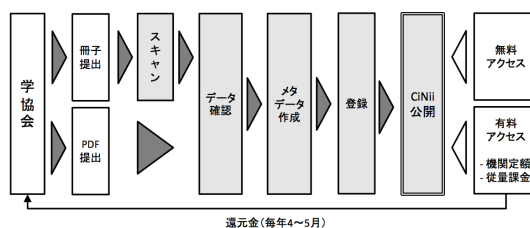


図2 電子図書館事業のプロセス

ら行われてきた電子図書館事業(NII-ELS)によって、学協会が発行した冊子体の論文誌をスキャンし、書誌情報を付与した約 400 万件の論文が対象になる(図2)。ユーザへの提供条件は誰もが無料でアクセス可能なオープンアクセス、国立情報学研究所との契約を結んだ機関の構成員であれば自由にアクセスできる機関定額制、有料で販売するペーパービューの中から学協会が選択する。

書誌情報の検索・閲覧機能は、電子図書館事業による約 400 万件のデータに加え、学術雑誌公開支援事業によって大学自らが電子化を行った紀要、国立国会図書館の雑誌記事索引、科学技術振興機構の電子ジャーナルプラットフォーム J-STAGE、機関リポジトリに格納された論文などを合わせて約 1,500 万件が対象となる。各情報源からは同一

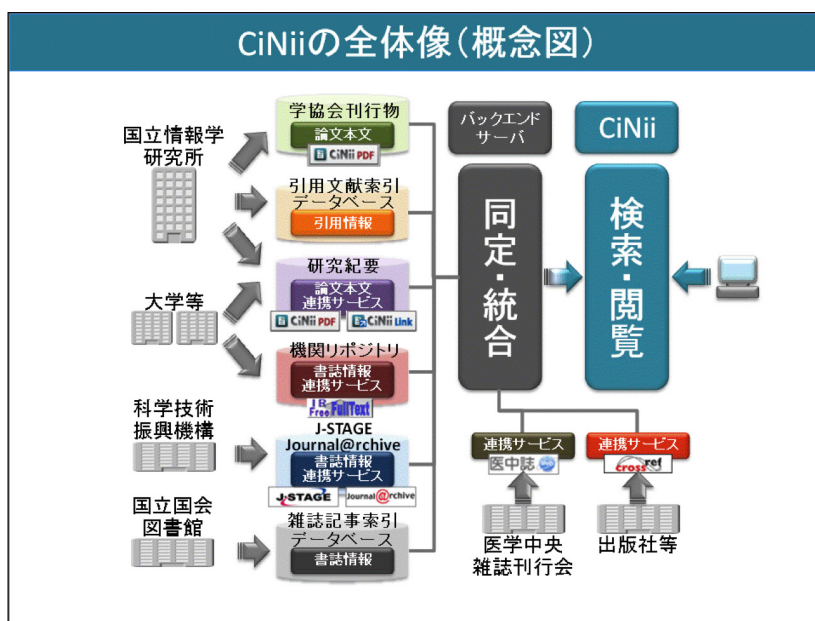


図3 CiNii Articlesの情報源とシステム構成

の論文に関するデータが複数含まれる可能性があるため、CiNii Articles では自然言語処理や機械学習を用いた名寄せのためのシステムを開発するとともに、一部については人手でのチェックを行うことで利便性の高いデータベース構築を行っている (図 3) [1]。

CiNii Article の利用回数は年々増加しており、近年では月間 1,500 万～2,000 万アクセス、年間では 2 億アクセスを超える。また、2009 年 (平成 21 年) からは開発者向けに構造化されたデータが利用可能なウェブ API を提供しており、その利用回数は年間 3 億回程度である。利用の増加やユーザ層の拡大に合わせて、CiNii Articles では年間 5～10 回程度のサービスのアップデート・修正を繰り返している。2014 年度 (平成 26 年度) の改善点としては電子図書館事業のスキャンデータ全件に OCR を適用し、透明テキストを付加した PDF を提供したこと、またシステムトラブル時にパブリッククラウドを利用することでサービス継続を可能にしたことが挙げられる。

今後は博士論文を対象とした検索サービス CiNii Dissertations (仮称) の開発を予定している。博士論文については 2013 年度 (平成 25 年度) の学位規則改正によりオンライン公開が義務化され、機関リポジトリを通じた公開が推奨されている。これらの情報と、国立国会図書館によってデジタル化された 1991 年 (平成 3 年) から 2000 年 (平成 12 年) までの発行分の情報などを組み合わせ、効率的な発見・入手を支援するサービスの提供を

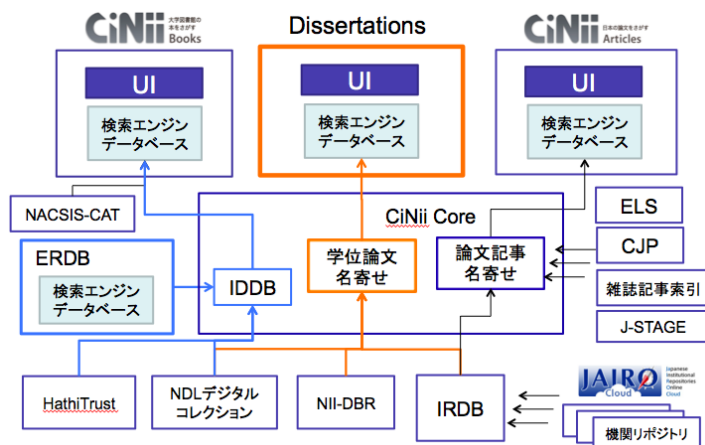


図 4 CiNii Dissertations (仮称) の情報源

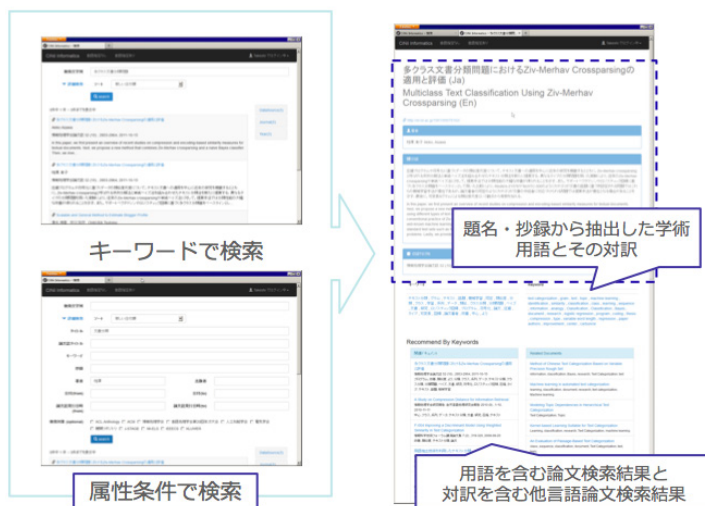


図 5 CiNii for Informatics (仮称) の検索・推薦



図 6 CiNii for Informatics (仮称) の情報提示

目指す(図4)。また、特定の研究分野に特化した学術情報サービスの可能性について検討する研究プロジェクトを開始している。このプロジェクトでは情報学を対象として海外論文をも収録対象とした CiNii for Informatics (仮称) を構築し、研究分野に適した情報提示やコンテンツ推薦の方式を検証している(図5・6)。このように、国立情報学研究所ではコンテンツの拡大とサービス品質の向上を目的とした取り組みを継続的に進めていく所存である。

■ CiNii Articles のコンテンツ

CiNii Articles 自体は研究分野を問わない学術情報サービスだが、電子図書館事業を含めてこの枠組みに参加している学協会あるいはコミュニティの数、あるいはオープンアクセスに対する姿勢については分野の違いが大きい。ここでは電子図書館事業のデータから傾向を把握する。なお、すべての数値は2014年(平成26年)3月現在のものである。

電子図書館事業では、参加している全428学協会を、人文科学・法学・経済学・理学・工学・農学・医学の7分野に分類している。実際には学際領域を対象とした学協会が多数存在するため、分類の正確さについてはここでは議論しない。分野ごとの学協会数を表1に示す^[2]。学協会数が最多の分野は人文科学(34.1%)である。ただし、理学・工学・農学・医学を自然科学として集約すれば計249となり、全学協会の58.2%を占める計算になる。以降は法学・経済学を合わせて社会科学と定義し、人文科学・社会科学・自然科学の3分野の比較を行うこととする。

表1 分野ごとの学協会数

分野	学協会数
人文科学	146
法学	4
経済学	29
理学	55
工学	61

農学	50
医学	83

各分野における雑誌数・論文数を表2に示す。雑誌数の割合は人文科学が249誌(18.3%)、社会科学が50誌(3.7%)、自然科学が1,062誌(78.0%)と、自然科学分野の割合が多い。このうち工学分野が605誌(44.4%)と突出している。一方で、論文数は人文科学が229,851件(6.4%)、社会科学が31,405件(0.9%)、自然科学が3,331,926件(92.7%)と、自然科学分野の論文数が他を圧倒する。うち工学分野は1,410,753件(39.3%)であり、最大のシェアであることは変わらないが、1誌あたりの論文数では152誌、1,004,849件(28.0%)の医学分野が最大となる。

表2 分野ごとの雑誌数・論文数

分野	雑誌数	論文数
人文科学	249	229,851
社会科学	50	31,405
自然科学	1,062	3,331,926

次に料金ごとの分類を見る。有料(ペイパービュー)で公開されている論文は232,056件(6.5%)、機関定額制は1,327,958件(37.0%)、無料公開は2,033,168件(56.6%)である。学術雑誌公開支援事業による紀要と合わせると CiNii Articles を通じて入手可能な文献の多くがオープンアクセス化されていることがわかる。また、国内の大学のほとんどが機関定額制の契約を行っており、それらの機関に所属する研究者や学生はペイパービューの文献以外は実質的に無料でアクセスが可能である。

■ 人文科学におけるユーザの意向と提供状況

CiNii Articles では定期的にユーザのアンケートを実施している^[3]。サンプリング調査ではなく回答者層の偏りがある可能性は否定できないものの、結果を通じて CiNii Articles のユーザ層を一定程度推定することができる。

ここでは2013年度(平成25年度)のアンケート

結果について述べる。アンケートの有効回答数は 1,215 件であった。職業に関する質問（総回答数 1,202）については大学院生を除く学生が 177、大学院生が 146、研究員が 39、教員が 267、図書館員が 267、官公庁等職員が 51、会社員が 147、その他・無回答が 108 である。表 3 に 2012 年度（平成 24 年度）と合わせた調査結果を示す。

表 3 職業に関する調査結果

職種	2013 年度	2012 年度
学生	177	151
大学院生	146	119
研究員	39	18
教員	267	164
図書館員	267	235
官公庁等職員	51	46
会社員	147	102
その他・無回答	108	76

分野に関しては 2 種類の質問を行った。CiNii Articles の利用目的として前述の 7 分野のうちどの分野を検索しているかという質問の回答（複数回答可）は、人文科学 672（34.2%）、社会科学 259（13.2%：法学 105・経済学 154）、自然科学 914（46.6%：理学 224・工学 329・農学 118・医学 243）、その他 118（6.0%）であった。表 4 に 2012 年度（平成 24 年度）と合わせた調査結果を示す。

表 4 利用分野に関する調査結果

分野	2013 年度	2012 年度
人文科学	672	481
法学	105	76
経済学	154	114
理学	224	154
工学	329	239
農学	118	98
医学	243	197
その他	118	120

また、回答者がどの分野に属しているかを科学研究費助成事業の系・分野・分科・細目表から選ぶ質問では、総合領域 148（14.5%）、複合新領域 26（2.5%）、人文学 196（19.2%）、社会科学 242（25.4%）、数物系科学 24（2.4%）、化学 11（1.1%）、工学 140（13.7%）、生物学 12（1.2%）、農学 43（4.2%）、医歯

薬学 98（9.6%）、無回答 81（7.9%）となった。表 5 に 2012 年度（平成 24 年度）と合わせた調査結果を示す。

表 5 所属分野に関する調査結果

分野	2013 年度	2012 年度
総合領域	148	120
複合新領域	26	16
人文学	196	129
社会科学	242	182
数物系科学	24	14
化学	11	8
工学	140	111
生物学	12	9
農学	43	38
医歯薬学	98	70
無回答	81	68

人文科学分野の学協会に限定して論文ごとの料金設定を見ると、従量課金制が 2,327 件（1.0%）、機関定額制が 63,492 件（27.6%）、無料公開が 164,032 件（71.4%）となる。全学協会での統計値と比較するとオープンアクセス文献の割合は約 15 ポイント程度高く、他分野と比較してオープンアクセスに対する意識が特段低いわけではないことがわかる。

しかしながら、アンケート結果から読み取ることができるユーザとしての電子情報の利用意向の強さと、実際に入手可能な論文の絶対数（229,851 件・電子図書館事業全体の 6.4%）に大きなギャップがあることは否定できない。また、自然科学分野では他に J-STAGE（約 200 万件）などの電子ジャーナルプラットフォームを利用した出版も多く、論文数の差はさらに開く。結果として、CiNii Articles を通じて入手できる人文科学分野の文献の多くが紀要由来であるという状態になっている。紀要の電子化は、前述の学術雑誌公開支援事業によるものであり、これまでに 2,868 誌が対象となっている。近年では機関リポジトリを通じた紀要の電子的出版も増加している。

紀要が相対的に目立つ状況の中、CiNii Articles に寄せられる批判として「論文の由来を理解しないまま引用する傾向を助長している」というもの

がある。主に査読を経ない論文の取り扱いに関する懸念であると思われるが、この批判にはいくつかの論点が含まれている。サービス提供者として、ある論文の掲載誌が専門雑誌であるのか紀要であるのか、あるいは査読の有無といった付加情報を積極的に収集し、利用者の判断に資する材料を増やすべきという意味では重要な指摘である。

一方、CiNii Articles の努力だけでこういった傾向を押しとどめることは困難であることも指摘しなければならない。抜本的な解決のためには電子的に入手可能な文献の大幅な増加が求められる。上記のような批判は、電子化・オープンアクセス化が相対的に進行している紀要が、その利便性のために積極的に利用されていることの証左である。広く読まれることを優先するならば、冊子体の学協会誌よりもむしろ紀要に掲載することが有効であるとの声も聞かれる。とくに若手研究者にとっては、キャリアパス形成のためには自身の専門分野だけでなく周囲の分野の研究者に活動が認められることが極めて重要であり、その手段として成果の可視化は必須である。また、評価方法についても論文の被引用数やその集計値としての h-Index といった指標、ソーシャルメディアでの言及をカウントする Altmetrics^[4] など、論文発表後になされる評価の重みが増している。このような潮流が正しいかどうかに関する議論は置いておくが、研究分野として多様な人材を引き寄せ、長期的な活性化を図るための戦略として電子化・オープンアクセス化の恩恵は非常に大きい。

翻って学協会にとっては、論文のオープンアクセス化は会費制による専門誌の発行とその頒布という運営モデルを根底から覆す可能性がある。購読者からの課金である会費制から、論文著者に対する出版料の徴収へと転換する学協会も現れており、今後学術分野全体に大きな変化が生じるものと予想される。一方で電子的発行・流通にかかるコストは着実に低下しており、運営体制の如何に関わらず上記のメリットを享受することは可能であろう。

ここまで電子図書館事業のデータを中心としてオープンアクセスの意義について述べてきたが、

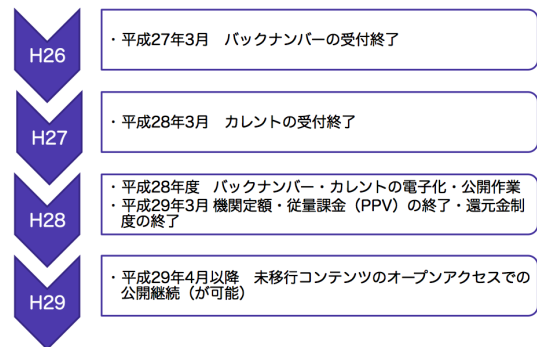


図7 電子図書館事業終了のスケジュール

国の方針によって本事業は 2016 年度 (平成 28 年度) 末を持って実質的に終了することとなった (図 7)^[5]。その後は簡易的な電子ジャーナルプラットフォーム J-STAGE Lite (仮称)^[6] 等を活用することで学協会自らが電子化・オープンアクセス化の主体となることが期待されている。CiNii Articles 自体は今後もデータ連携や名寄せの精度を高め、検索と文献入手のためのサービスを継続する。乗り越えるべき課題は多いが、人文科学分野の持続的発展のためにもよりよい学術情報共有プラットフォームに関する議論を続けていきたい。

脚注・参考文献

- [1] 大向 一輝: CiNii Articles のシステムデザインとデータモデル. 情報の科学と技術. Vol. 62, No. 11, pp. 473-477 (2012).
- [2] すべての学協会のリストは http://ci.nii.ac.jp/society/society_name/all_ja.html を参照のこと。
- [3] アンケートの概要は http://ci.nii.ac.jp/info/ja/a_result_2013.html (2013 年度) http://ci.nii.ac.jp/info/ja/a_result_2012.html (2012 年度) を参照のこと。
- [4] 林 和弘: 研究論文の影響度を測定する新しい動き—論文単位で即時かつ多面的な測定を可能とする Altmetrics 一. 科学技術動向. No. 134, pp. 20-29 (2013).
- [5] 電子図書館 (NII-ELS) の事業終了について. http://www.nii.ac.jp/nels_soc/about/
- [6] J-STAGE Lite (仮称) の開発について. https://www.jstage.jst.go.jp/pub/html/AY045560_ja.html

2014 年度夏期公開シンポジウム

「情報化時代における中国学研究基盤を考える」レポート

佐藤 信弥

はじめに

本稿は 2014 年 8 月 2 日に関西大学で開催された漢字文献情報処理研究会夏期公開シンポジウム「情報化時代における中国学研究基盤を考える」の報告である。今回のシンポジウムは主催が漢字文献情報処理研究会、共催が日本学術振興会科学研究費基盤研究（B）「情報化時代における中国学次世代研究基盤の確立」（研究代表者 二階堂善弘氏）及び関西大学アジア文化研究センターとなる。千田大介・師茂樹・大向一輝の三氏による講演の後にフリーディスカッションが行われた。

最初に司会の二階堂善弘氏（関西大学）より開催趣旨の説明が行われた。同氏によると、よく海外の研究者から「どうして日本人の研究はネットで閲覧できないのか？」という質問を受けるという。無論、研究者によっては論文を自分のホームページなどにアップロードしている場合もあるが、個人での対応には限界があり、これは外国のように国家単位で対応すべき課題である。そして最近では日本の大学院生なども、ネットで公開されていない、もしくは CiNii などの検索で引っかけられない

研究は無視してもよいという感覚になってきているという問題提起がなされた。

学術情報収集のスキルとインフラ整備

千田大介氏（慶應義塾大学）による講演。中国学の分野での情報化への対応や大規模古典文献データベースについて述べる。

近年中国及び台湾において国家的規模で電子ジャーナルや古典文献データベースなどの学術インフラの整備が進められている。研究の基本となる古典文献からの出典・用例調査をはじめとして、現在の中国学研究はこれらのデータベースの利用が欠かせないものとなっている。代表的な大規模古典文献データベースとしては、台湾中央研究院の漢籍電子文献のほか、中国の書同文の文淵閣『四庫全書』・『四部叢刊』、愛如生の中国基本古籍庫などがあるほか、最新のものとしては二十五史・史料筆記叢刊など中華書局刊行の点校本によるデータベース中華經典古籍庫が注目される¹⁾。

ただ、これらのデータベースは、比較的少額のものでも数十万円、中国基本古籍庫など高額のも

のとなると二千五百万円程度の導入費や使用料が必要となる¹²⁾。更にこの種のデータベースは、収録する文献の種類や数量、テキストの校訂などの面でそれぞれ問題があり、これひとつ導入すれば事足りるという決定版が現状では存在しない。よってどうしても一つの機関で複数のデータベースを揃える必要があるが、中国学が外国学の分野という扱いでしかない日本の大学の場合、大学間で協定を結んで分担導入を図ったり、特定の研究機関が集中的に導入するといった試みが必要となろう。

またデータベースのシステム上の問題として、検索機能が不十分である、検索時の異体字への対応方式が統一されていなかったり不完全であったりするといった事項が挙げられ、利用する側もそういった問題を認識し、大学院生などに検索などのスキルを教育していく必要がある。

このほかデータベースを利用する側の今後の課題としては、データベースの品質・信頼性や検索機能に関する評価基準を確立し、これらの情報を Web や書籍の形で集約・周知を図っていくことが挙げられる。

■ 学術データベースの構築と公開

師茂樹氏 (花岡大学) による講演。日本のデータベースの特徴と今後の開発・公開の課題について述べる。

近年は日本でも大学図書館所蔵の貴重書などを中心にネット上での文献の公開が進んでいるはずであるが、海外の研究者の反応を見ると、意外とそれらの情報が周知されていなかったり、日本にはあまり文字による資料がないと誤解されていたりする。日本に残されている文献の情報が完全にネットの中で埋没してしまっているのである。その原因はどこにあるのだろうか？

中国・台湾・日本のデータベースを比較すると、中国・台湾は台湾中央研究院の漢籍電子文献など本文検索機能を前提としたテキストデータの提供が中心であるのに対して、日本の場合は国立国会図書館の近代デジタルライブラリーや国立公文書

館のアジア歴史資料センターなど、画像での公開が中心である。これが海外の研究者から日本の文献が「見えない」主要な理由である。

画像での公開が中心となる理由としては、データベースの制作・公開が科研費などの助成金を獲得しやすい項目となっている一方で、データベース制作が研究者の業績としてあまり評価されず、その結果、研究者がほとんど制作に関与せずに下請けの企業に丸投げの状態になってしまうことなどが挙げられる。

日本のデータベースを海外から「見える」ようにするためには、画像によるものからテキストによるものへの移行と、研究者をデータベース制作の主体にすること、そして制作したデータを Google での検索にヒットしやすくすることが必要になる。そのための試みとして近代デジタルライブラリーの画像データを有志でテキストデータ化していく翻デジなどの「ソーシャル翻刻」が興味深いが、組織性・体系性・人員の確保に問題がある。海外へのアピールを強めるには、それらの活動を取りまとめるポータル構築など、戦略性が必要となろう。

■ 学術共用プラットフォームとしての CiNii

大向一輝氏 (NII (国立情報学研究所)) による講演。CiNii の事業の概要や今後の方向性・課題について述べる。

NII は昭和 51 (1976) 年に東京大学情報図書館学研究センターとして発足したが、昭和 61 (1986) 年に NACSIS (学術情報センター) として東大より分離、平成 12 (2000) 年に現在の NII となった。以来事業 (電子図書館事業などによる学術情報の収集・公開) と研究 (情報学の研究・教育) を二本柱としてきたが、近年は研究の方に軸足を移しつつある。

NII の事業としては、当初 CiNii Books 及び CiNii Articles を中心とする大学図書館の蔵書状況や機関リポジトリからの情報収集・公開のみであったが、現在は科研費の研究課題・成果情報のデータベース (KAKEN) 構築や博士論文のネット

による公開の義務化に伴う関連データの収集・公開 (CiNii Dissertations)、各大学の機関リポジトリ構築支援などに広がりつつある。

CiNii Articles についても、同じ論文の書誌や研究者の情報が複数にまたがるなどの問題が発生しており、その情報統合をどう進めていくかについて検討中である。CiNii Articles において論文の本文は PDF ファイル形式で現在約 402 万件提供されているが、現在これを画像データから透明テキスト付きデータへと変換する作業が進められている。またスマホやタブレットでの利用を想定した機能の追加、博士論文や海外の論文、電子資料、あるいは文献以外のデータ（動画など）検索への対応も課題に挙がっている。更にネイチャー・パブリッシング・グループが、論文作成のために使用されたデータセットを公表するための電子ジャーナル Scientific Data を創刊したことを受け、各論文がどのデータセットを使用したかという情報を検索できるようにしたり、CiNii for Informations という形で、情報学の分野から、論文本文の PDF に用語の解説・関連する文献や動画などの情報を付加して提供していく試みが進められている。

ただし、周知のように CiNii が担ってきた電子図書館事業は平成 27(2015)年 3 月から平成 29(2017)年 3 月にかけて順次終了する予定となっており、学会・協会誌の電子化などは JST (科学技術振興機構) の J-STAGE Lite に一本化されることになっている¹³⁾。

人文科学の分野による CiNii Articles の利用状況については、登録学会数では人文科学系の学会が全体の 34.1% を占め、登録雑誌数では 18.3%、更に利用者の割合も人文科学の分野に属する利用者が 32.5% を占めるものの、登録論文数は全体の 6.4% にすぎず、CiNii Articles に対して利用者の期待しているものと、実際に提供できているものが乖離しており、その溝をどう埋めていくかが課題である。

■ フリーディスカッション

講演の後は聴衆からの質問にパネラー（講演を担当した三氏）が答える、あるいは逆にパネラーから聴衆に向けて質問がなされたり、意見が出されるという形で、活発な議論が行われた。

特に問題とされたのは、人文科学の分野においてデータベースの制作者にどの程度、あるいはどのようにしてリスペクトを示したら良いのかである。データベースを利用して論文を作成する場合、その旨明記する必要があるのではないかという意見に対しては、例えばデータベースの利用を謝辞の形で示すなど、一定のルールを決める必要があるだろうという回答が寄せられた。これに関連して、データベースを利用しなければ成り立たないような研究では当然その旨明記する必要があるだろうが、単純に索引がわりにデータベースを利用する場合は必ずしもその必要はないのではないかという意見も出た。

データベースの制作をどの程度業績として認めるかという問題については、履歴書・業績書に書ける程度には評価すべきという意見が出る一方で、データベース自体は中国の専門企業に制作を委託するという形で、企画発案者が直接作業せずとも制作は可能なので、データベース制作を業績として認めた場合、制作資金さえあればいくらでも業績が水増しできるのではないかという懸念が寄せられた。またかつて研究室単位での索引づくりが大学院生の教育の一環となっていたように、データベース制作を大学院生の教育課程として取り入れたら良いのではないかと、大向氏の講演の中で話題に出た Scientific Data の事例を参考に、データベースあるいはデジタル翻刻などのデータそのものを査読・評価する仕組みをつくる必要があるという意見も出た。

データベースの評価とともに、新しい OS への対応など、その維持やメンテナンスも問題となる。制作時にあらかじめ検索・閲覧に関するソフトウェアの部分と中身のデータの部分を切り分けられるようにしておき、新しいソフトウェア上で

データを再利用できるような工夫が求められるという意見が出た。

■ おわりに

千田氏・師氏の講演に関して、中国で活発にデータベースの制作が進められているのは、そもそも紙媒体での工具書などの作成が業績として認められているのが強く影響しているのであろう。例えば筆者（佐藤）の専攻である中国古文字学の分野では、文字編（見出し字に該当する古文字の字形を羅列する形式の字書）などの工具書や博物館の解説付き図録（例えば欧米の特定の博物館に所蔵されている甲骨の目録など）、あるいは既出の工具書や図録の校訂が学位論文として認められ、その審査の手法も確立されている。また工具書や図録の書評も多く発表されている。このような学術的環境が有用なデータベース制作の原動力のひとつとなっていると思われる。

日本の学術的環境が中国とは相当に異なり、か

つ中国と同様の環境に改めていくのが困難であるという現状をふまえたうえで、日本からもネットを通じてより有用な学術情報を発信していくには、日本の環境に合わせた形を模索していくほかなのではないか。

注

- [1] 中華經典古籍庫の詳細については、本誌掲載レビュー（p.110）及び東方書店による紹介（<http://www.toho-shoten.co.jp/er07/zhonghua.html>）を参照。以下、本稿で引用する URL はいずれも 2014 年 8 月に確認したものである。
- [2] これら各種データベースの具体的な導入費等については、漢字情報システム HP（<http://www.naf.co.jp/kis/shikozensho.stm>）や東方書店電子商品カタログ（<http://www.toho-shoten.co.jp/export/sites/default/er07/er07-price.pdf>）を参照。
- [3] NII による告知（http://www.nii.ac.jp/nels_soc/about/）などを参照。

漢情研記念論集(仮)

原稿募集のお知らせ

かねてお知らせしておりますように、漢情研は来年度、2015 年度をもって、従来の会費制度による会の運営、いうなれば第一期の活動を終了し、その後はこれまでに培った人的ネットワークをベースに、科研費研究などのプロジェクトと連動してイベント開催・出版などの活動を行っていく予定です。

現在、漢情研第一期の終了と連動する形で、二階堂善弘元代表を研究者とする科研費研究「情報化時代における中国学次世代研究基盤の確立」(基盤研究(B))が実施されており、漢情研のこれまでの活動を総括する書籍『電腦中国学入門』・『大学で学ぼう』・『人文学と著

作権問題』(いずれも好文出版)を刊行してきました。

漢情研第一期および科研費研究の締めくくりの年である来年度には、記念論文集の刊行を計画しております。科研費研究メンバーにとどまらず、広く漢情研の内外から原稿を募り、漢情研第一期の締めくくりにふさわしい内容にしようと考えております。

つきましては、下記要領にて原稿を募集いたしますので、奮ってご投稿くださいませ。

また、本誌次号も第一期最終号として、漢情研の回顧特集を計画しております。こちらへのご協力も、あわせてお願い申し上げます。

記

漢情研記念論集(仮) 原稿募集要項

- 内容：漢情研の活動趣旨と合致する、人文学、とりわけ東洋学の研究・教育の情報化に関する論文・研究ノート・翻訳など。
- 言語：原則として日本語。
- 分量：400 字詰め原稿用紙 20 ～ 60 枚
- 執筆申込：2015 年 6 月末日までに漢情研執行部メイリングリスト (kanji@jaet.gr.jp) 宛に、タイトル・概要を添えてお申し込みください。
- 原稿締切：2015 年 9 月末日。ご投稿頂いた原稿は査読させていただきます。その結果によって、修正をお願いすることがありますので、ご承知置きください。

※ 2015 年春頃に、改めて BBS・メールマガジン等で詳細をお知らせする予定です。

以上

東洋学

論文検索

指南

近年、学術論文のデジタル化が急速に進んでいる。エルゼビアやシュブリンガーといった大手の電子ジャーナル企業に加え、新しい電子ジャーナル企業が各国でサービスを開始している。PDF だけの論文誌も珍しくなくなってきたし、既存の論文誌がオープンアクセス化される動きも進んでいる。

日本においては、最近「学位規則」が改正され、2013 年 4 月 1 日以降に博士号を授与された博士論文は、原則としてインターネットを通じて公表するという事になった。これは、日本の学術論文の流通、オープンアクセス化にとって大きな一歩となるだろうが、東アジア各国ですで行われていることでもある。ともあれ、有償・無償を問わず、学術論文のデジタル化は、世界中でますます加速していくことであろう。

しかし、このような流れに、日本の東洋学の研究者はついていけるのだろうか。中国では、CNKI をはじめとする学術データベースの利用方法について解説した書籍がいくつも出版されているという。一方日本では、国立国会図書館が公開している「リサーチ・ナビ」(<http://mavi.ndl.go.jp/mavi/>) にアジア諸国の情報を調べる方法がまとまっており、そのなかには学術論文・博士論文の調べ方も載っている。しかしここは、情報量が多い反面、具体的な検索方法などについては簡略な記事にとどまっており、初学者にとってはやや敷居の高いものとなっている。本会編『電脳中国学入門』(好文出版、2012 年)にも II-5「論文・雑誌記事の調査」という一節があり、日本・中国・台湾の主だったサイトが紹介されているが、具体的な使用方法については CiNii Articles と CNKI にとどまり、それ以外については簡単な紹介記事しかない。

東洋学のような外国文化の研究において、海外の研究情報へのアクセス方法が知られていない状況は、言うまでもなく危機的である。各研究者が情報をつかむことができているのであれば問題はないが、実際には入門書レベルで普及は不可欠であろう。

本特集は、『電脳中国学入門』を補足する形で、東アジア各国の学術論文・学位論文の検索方法などをまとめたものである。本特集が、学術情報のデジタル化という大きな時流に乗るための一助となれば幸いである。

CONTENTS

日本の学術論文データベース	小島 浩之	42
中国の論文を探す	千田 大介	51
台湾・香港の論文データベース	二階堂善弘	65
韓国の学術論文の検索方法	師 茂樹	70
ベトナムの文献を探索する	矢野 正隆	73

日本の学術論文データベース

小島 浩之（こじま ひろゆき）

＊ はじめに

日本の学術論文データベースについては、『電脳中国学入門』Ⅱ-5において「論文・雑誌記事の調査」と題して、様々なデータベースを紹介してきた。内容的に重複する部分があることも否めないが、『電脳中国学入門』の刊行から2年を経過していること、同書では簡単な紹介に留まらざるを得なかったものもあることなどから、こういった部分に特化して改めて論じてみたい。

＊ 博士論文のオープンアクセス化

『電脳中国学入門』刊行後の最大の変化としては、「学位規則」の改正により、2013年4月1日以降に国内で授与された博士論文の公開は、原則としてインターネットによるものとされたことだろう。ここで原則とあるのは、特許や諸権利との兼ね合いにより即時のインターネット公開ができればない場合があるのを考慮したことによる。

公開は原則として各大学等のリポジトリ¹⁾を通じてなされ、2013年度以降の日本の博士論文

は全てがオープンアクセスとなっている。

一方、法改正以前に学位が授与された論文についても、国立国会図書館（NDL）が著作権処理を行い、許諾の得られたものから順次インターネット公開されることとなっている。これまで、過去に提出された博士論文については、各大学が独自に著作権許諾を得て、リポジトリで公開してきていたが、今後は、平成22年度以前はNDL、平成23年度以降は各大学リポジトリ、という棲み分けがなされることとなった。

このため、博士論文の原文を閲覧する場合、平成22年度以前のは「国立国会図書館デジタルコレクション」から、平成23年度以降のものは各大学のリポジトリからアクセスすることとなる。



図1 国立国会図書館デジタルコレクショントップページ

● 国立国会図書館デジタルコレクション

<http://dl.ndl.go.jp/>

NDL は、平成 21 年の著作権法改正により、「図書館資料の原本を公衆の利用に供することによるその滅失、損傷又は汚損を避けるため」、権利者の許諾を得ずとも図書をデジタル化の上で、利用に供せることとなった。さらに平成 24 年の改正では、全国の図書館等は、NDL から入手困難な資料のデジタル送信を受けて、利用者に閲覧・複写させることが可能となった。

これらの経緯から、当該サイトで公開されるデジタルデータは、

- ①インターネット公開資料
- ②図書館送信資料
- ③国立国会図書館内限定 [資料]

の三種に分かれている。①はオープンアクセスのもの、②は NDL 館内のほか、NDL と配信協定を結んだ全国の図書館内で利用できるもの、③は NDL 館内限定で利用できるものである。2014 年 8 月現在で、1991～2000 年度に NDL に送付された 14 万件超の博士論文が登録され、このうち①に該当するものは、約 1 割の 14,661 件、②は約 11 万 6 千件、③は約 9,200 件となっている。

なお、NDL とデジタル配信の協定を結んでいる図書館は「図書館向けデジタル化資料送信サービス参加館一覧」(http://dl.ndl.go.jp/ja/soshin_librarylist.html) から確認できる。

閲覧はトップページで「博士論文」をクリックすると、専用の検索画面に進む。ここで必要に応じて先述の①から③に該当する



図2 博士論文検索ページ

チェックボックスにチェックを入れ、検索ボックスに検索語を入力した上で検索ボタンをクリックする。

なお、画面左メニューからは、「博士論文タイトル一覧」および「インターネット公開分のみの一覧」へアクセスでき、便利である。

● JAIRO

<http://ju.nii.ac.jp/>

ジャイロ

JAIRO (=Japanese Institutional Repositories Online) は日本における機関リポジトリの統合検索サイトである。各大学のリポジトリに登録された博士論文は JAIRO から検索・閲覧が可能となっている。

ただし、JAIRO には簡易検索、詳細検索のいずれのインターフェースにも博士論文に限定して検索する機能は搭載されていないようである。

試みに「博士論文」というキーワードで検索す



図3 JAIRO トップページ



図4 not 検索を利用したの博論一覧検索

ると、論文そのもののほかに、要旨や概要、一覧が数多くヒットする。このため登録されている博士論文を一覧するには、以下の要領で検索するとよい。

- ①簡易検索画面で検索ボックスに「博士論文」のキーワードを入れて検索
- ②検索結果一覧、右上の「検索エリアの追加」をクリックし、プルダウンメニューから NOT を選択し、追加の検索ボックスに「要旨」を入力
- ③同様に「検索エリアの追加」を繰り返し、「一覧」と「概要」を not 検索で入力
- ④改めて検索ボタンをクリックする

これにより、博士論文 not 要旨 not 一覧 not 概要の not 検索、つまりキーワード「博士論文」を

含み、「要旨」・「一覧」・「概要」の語は含まないものを検索できる。

なお、各大学のリポジトリに登録されてから、JAIRO で検索可能となるには、一定のタイムラグがあるように感じられるが、この期間がどの程度なのかは明らかにされていない。このため目当ての博士論文が JAIRO でヒットしない場合は、学位を授与した機関のリポジトリに直接アクセスすることをお奨めする。日本の機関リポジ

トリの一覧は、NII が以下のページで公開しているので参照されたい。

<http://www.nii.ac.jp/irp/list/>

✳ 一般論文の検索と閲覧

● CiNii Articles

<http://ci.nii.ac.jp/>

サイニイ

CiNii Articles (以下、本稿では CiNii と略) は国立情報学研究所 (NII) の提供によるわが国最大の論文情報データベースである。2014 年 8 月現在で約 1,700 万件の論文書誌データが検索できる。

CiNii に搭載されるデータは、NII がその前身の学術総合センター (NACSIS) 時代から、学・協会とともに作り上げてきた電子図書館サービス (① NACSIS-ELS)、② NDL の雑誌記事索引、③各大学の機関リポジトリに登録された学术论文、④科学技術振興機構の J-STAGE、⑤その他各種電子ジャーナルからなる。

このうち①については、2017 年 3 月末をもって終了することが発表されており、これらの事業は形式的には④に統合・継承されることとなる。このため 2017 年 3 月末以降は、NII として作成するデータは皆無となり、CiNii はポータルサイトとしての役割のみを有



図5 CiNii Articles トップページ

することになる^[2]。

要は、各研究機関の紀要類はリポジトリを、学・協会の学術雑誌類は J-STAGE を、それぞれ通じて CiNii で検索・閲覧できるようになるということである。ただし、学・協会誌の中には、事務局のある研究機関のリポジトリで公開を進めているところもある。なお、雑誌記事索引や、その他各種電子ジャーナルの情報は、両者に搭載されていない論文データを補うものという位置付けになるといえよいだろう。

CiNii から閲覧できる論文は、全て PDF 形式となっており、フリーアクセスのものと有料のものがある。有料の場合、所属機関への課金と個人への課金という二つの精算方法が用意されている。大学などでは定額制や従量課金制に機関として参加していることも多く、この場合、所属機関に接続するパソコンからは、自宅からより多くの論文本文を閲覧することができる。ただし、CiNii のあり方が変わる 2017 年 4 月以降は、これらの課金方式も様変わりすることになるだろう。

CiNii の利用で注意すべきは、搭載される論文情報は必ずしも網羅的ではないということである。これは、良くも悪くも CiNii のデータは、各データ提供元の採録スタンスに依存せざるを得ないからである。機関リポジトリに全ての紀要が入っているとは限らないし、NDL の雑誌記事索引も採録誌が限定されている^[3]。

このため、CiNii は日本の論文情報を検索するにあたって、まず参照すべきサイトではあるが、CiNii の検索結果が万能だとは限らないことを肝に銘じておくべきだろう。元となるデータベースがそれぞれどういった性格のもので、どのような雑誌が採録対象となっているのかについて、自身の専攻と照らし合わせてみておく必要もあろう。

CiNii で不足する情報については、各分野の学術雑誌の近刊総覧などの記事や、後述するような各分野に特化した目録やデータベー

スなどにあたってみなければならない。

なお CiNii のインターフェースは、『電脳中国学入門』執筆時とほぼ変わっていないので、検索方法や検索状上の留意点については、同書 152-153 頁を参照されたい。

● 雑誌記事索引

雑誌記事索引は「^{ざっさく}雑索」と通称され、NDL が 1949 年以来、作成している国内刊行雑誌記事の書誌データ検索ツールである。雑索は NDLOPAC (<https://ndlopac.ndl.go.jp/>)、もしくは国立国会図書館サーチ (<http://iss.ndl.go.jp/>) から利用できるが、ここでは NDLOPAC からの利用法を解説する。

NDL の利用者登録が済んでいれば、ID、パスワードを入力してログインする①。そうでない場合は、「登録利用者 ID が無い方」というメニューから、「ゲストログイン」をクリックする②。トップページ下部には簡易検索機能③もあるが、ここでの検索結果は雑誌記事索引に限定されないの、ここでは②のゲストログインからの流れをみてみよう。

ゲストログインすると、簡易検索画面に遷移するので、画面上部のタブメニューから「雑誌記事」を選択すると、雑索専用の検索画面となる。

雑索の検索画面では、論題名・著者・雑誌名・出版者・編者・ISSN・出版年などから検索できるので、必要な項目を選びキーワードや数字を入力して検索ボタンをクリックする。

雑索には、本文を閲覧する電子ジャーナル的な



図6 NDLOPAC トップページ



図7 雑索専用検索画面

機能は無いが、利用者登録されていれば複写を申し込むことができる（有料）。

先述のように雑索のデータは、基本的に CiNii でも検索可能である。ただし、機関リポジトリと JAIRO の関係と同様に、雑索に登録されたデータが、CiNii で検索できるようになるまでにはある程度のタイムラグがある。雑索の書誌データの更新は意外と早いので、最新の情報を得たい場合は、CiNii と合わせて雑索の検索も心掛けるようにしたい。

国立国会図書館サーチからの雑索の利用方法は、『電脳中国学入門』144 頁を参照されたい。ただし、国立国会図書館サーチから検索しても、最終的には NDLOPAC の画面に遷移することになる。

国会図書館サーチは、Google に慣れた人には使いやすく感じるかもしれないが、実際はヒット数が膨大で、検索結果をうまく絞り込まないと満

足する結果が得られないことも多い。検索対象や検索目的が決まっている場合は、個別のデータベースから検索を試みる方が、往々にして時間にロスが無く、スムーズに結果を出せる。このため、雑誌論文の検索に目的が限定されるのであれば、NDLOPAC からログインの方がよいだろう。

● J-STAGE

科学技術情報発信・流通総合システム (J-STAGE) は、独立行政法人科学技術振興機構 (JST) が構築・管理する「日本国内の科学技術情報関係の電子ジャーナル発行を支援するシステム」である。

NII は、東京大学情報図書館学研究センターから始まる旧文部省系の組織であるのに対し、JST は日本科学技術情報センター (JICST) と新技術開発事業団 (JRDC) を前身とする旧科学技術庁系の組織である。JST は自然科学分野の学・協会誌の電子ジャーナル化を進めているという印象が強いが、人文科学分野についても 317 誌 (2014 年 8 月末現在) が登録されている。このため、端から見るとは NII と JST の電子ジャーナル事業の境界は曖昧模糊としている。

ただし先述のように、2017 年 3 月までで、NII は学・協会の電子ジャーナル事業から撤退するので、以後は J-STAGE が日本の学・協会電子ジャーナルのポータルサイトとして機能することになる。

最近の大学図書館の OPAC には、My OPAC などの名称で、OPAC の機能を自分専用にカスタマイズできるしくみがある。J-STAGE にも My J-STAGE といって、無料の ID を取得することで、資料、記事、検索条件等を登録・保存しておけるサービスがある (Yahoo!JAPAN ID、Google ID、livedoor ID を取得していれば、これらの ID を使用



図8 J-STAGE トップページ

することも可能)。これら ID でのログインはトップページ上部中央から行える。

2014 年 9 月 2 日現在で、全収録誌数(ジャーナル+予稿集)は 1,831 誌で、全収録記事数は 2,591,085 件、このうち無料で論文本文(PDF)が閲覧できるのは 1,724 件となっている。

メニュータブの「資料を探す」には、「資料一覧」(収録雑誌からの一覧)、「機関別一覧」(発行機関からの一覧)、「分野別一覧」(学術分野の分類からの一覧)の三種の一覧メニューが用意されている。検索さえできればよしとする Google ばりの割り切ったインターフェースが流行する中、内容をブラウジングできる機能を用意しているのは、利用主体である研究者の目線にたったものと言えるだろう。

大学図書館は研究者へのサービスが主たる業務の一つなのであるから、こういった部分は是非学んで欲しいものである。

資料の検索はトップページ右上部に、検索ボックスがあり、ここにキーワードを入力して検索ボタンをクリックすればよい。

✳ 東洋学文献類目検索

『東洋学文献類目』は「^{るいもく}類目」と通称され、京都大学人文科学研究所(人文研)で、1935 年以来発行されている東洋研究文献の書誌目録である(初号発行は東方文化学院京都研究所)。ここに収められている書誌データは、人文研の所蔵する図書・雑誌から採録されたもので、雑誌論文のみならず、単行書のほか、記念論集など書籍形態の出版物に掲載された論文も収録している。このため雑誌論文に特化した CiNii・雑索・J-STAGE などより検索できる範囲が広い。

ただし、あくまで人文研所蔵資料からの採録であるため、どんなに優れた研究業績であっても人

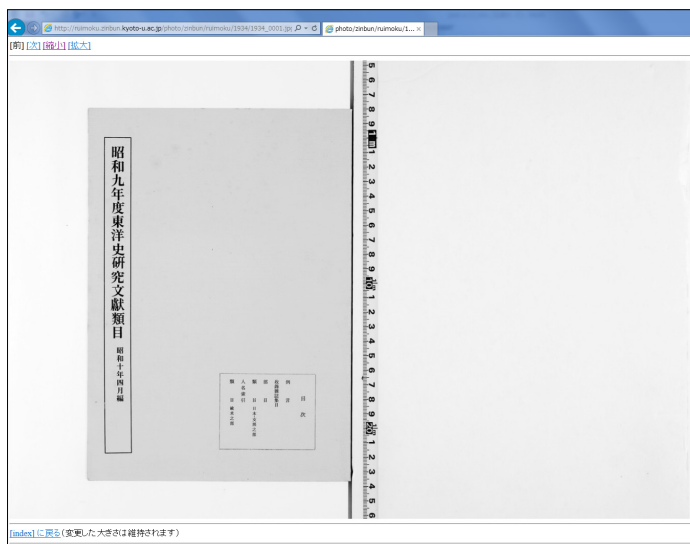


図9 類目初号(1935年発行)のアーカイブ

文研に所蔵がなければ掲載されていない点は、注意が必要である。

日本語文献に限らず、世界の諸言語で書かれたアジア研究の成果を収録している。インターネットで容易に論文の検索や閲覧が可能となる以前、類目の冊子体は大変重宝された。筆者なども学生時代に、過去の研究成果はまず類目で探すよう指導を受けたし、実際に利用してきた。冊子体は現在でも版を重ねているが、本誌で紹介しているような論文検索サイトの影に隠れてしまった感は否めない。なお、1980 年版までの冊子体は人文研のサイトでアーカイブされており、電子的に閲覧が可能となっている(<http://ruimoku.zinbun.kyoto-u.ac.jp/annual/index.html>)。

実は、1987 年に CHINA3 という名称で汎用機の検索システムが構築され、京都大学大型計算機センター(当時)から利用に供されていた。つまり、類目は雑索や CiNii などより数段早く電子化が進んだ論文目録なのであった。ただし、有料(機関による支払のみで私費利用は不可)かつ telnet 経由での利用のため、一部の大学を除き自由に利用することはできなかった。

現在では、Web 版として、「東洋学文献類目検索」第 6.10 版と「東洋学文献類目検索」第 7.3 α 版が公開されている。複数の検索サイトが存在す

図 10 類目検索 6.10 版トップページ

8 件見付かりました。	
ID	A016026-00004 (CHU002)
著者	顔世安 (yan shi an)
題名	齊桓公蒞政基礎の探討
雑誌	江海學刊 2001 年 1 期
発行年	2001
頁	113-117
分類	歴史 / 中國史 (13000)
ID	A86082300001 (ronbun)
著者	陳芳妹
題名	戦国田齊桓公の「金」碗
雑誌	故宮文物月刊 3
発行年	1985
頁	4-12
分類	藝術 / 工藝 (116000)
ID	A93203400002 (ronbun)
著者	野間 文史
題名	齊桓公の最期と『左傳』の「立」
雑誌	東方學 87 輯
発行年	1994
頁	28-41

図 11 検索結果一覧 (「齊桓公」で題名検索)

る理由については、守岡知彦氏の「レガシーとの付き合い方：東洋学文献類目の場合」^[4]に詳しいが、要約すると早い段階からの電子化への努力により、複数の形式のデータが作成され、それらの統合が一朝一夕には難しく、バージョンを分けて公開せざるを得ないということらしい。

● 東洋学文献類目検索 第 6.10 版

<http://ruimoku.zinbun.kyoto-u.ac.jp/ruimoku6/index.htmlja>

前掲守岡論文によれば、第 6 版は CHINA3 のデータ (1981 年度版 - 2000 年度版) に、人文研の現行システムにより入力された 2001 年度版以降のデータを統合したものだという。ただし、現行 6.10 版の検索トップページでは、2001 年度版以降のデータについては、第 7 版を利用するよう推奨されている。

6.10 版では検索ボックスが「題名」用と「著者」用が用意されているので、適宜キーワードを入力の上で検索開始ボタンをクリックする。ただし、入力字体によって検索結果が異なるので注意が必要である。たとえば題名検索の場合、「齊桓公」は 7 件、「斉桓公」は 5 件、「齐桓公」は 8 件のヒットとなる。また検索は部分文字列検索で、論理和 (and) 検索や論理積 (or) 検索はできない。

検索ボックス下部には、「歴史・地理・社会・経済・政治・法制・宗教・学術思想 (附教育)・科学・文学・芸術・考古学・金石古文書学・民族学・言語文字学・書誌学・雑纂 (附革命文物)・学会消息」といった 18 の主題分類がチェックボックスの形で示してある。必要に応じてチェックを入れれば、検索結果を絞り込むことができる。また、キーワードを入力せずに分類にチェックだけを入れて、検索開始

ボタンをクリックすると、当該分類に該当する書誌が一覧できる。

検索結果は、目録カードをイメージしてあるのか、ID・著者・題名・雑誌名・発行年・掲載頁・分類が、一論題ごとに野線で区切られて一覧表示される (図 11)。

● 東洋学文献類目検索 第 7.3 α 版

<http://ruimoku.zinbun.kyoto-u.ac.jp/ruimoku7/index.htmlja>

類目検索の 7.3 α 版については、検索サイトトップページに「1934 年度版から 1980 年度版までと 2001 年度版以降の東洋学文献類目のデータ (含未発行分)」を収録すると記されている。電算化以前の出版分を遡及入力したデータに加えて、最新のデータが逐次的にアップされているもののようである。

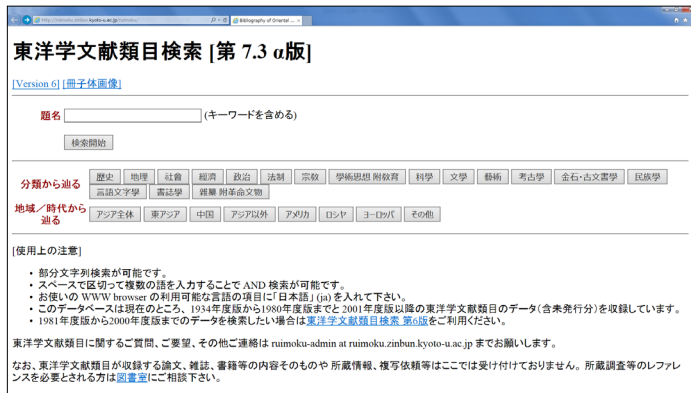


図 12 類目検索 7.3 α版トップページ



図 13 類目検索 7.3 α版で「歴史」分類をクリックした画面

検索ボックスは題名用しか用意されておらず、6.10 版のように著者名からは検索できない。ただし論文タイトルのほか、論文に付与してあるキーワード（類目検索の項目名として

は subject note）も検索対象の文字列に含まれている。たとえば、図 14 に示した「唐慕容暉皓墓志考釋」という論文は、subject note 欄に「唐代」・「吐谷浑」とあるので、これらのキーワードで検索してもヒットする。検索は、部分文字列検索のほか、複数のキーワードをスペースで区切ることで and 検索が可能となっている。入力字体によって検索結果が異なるのは 6 版と同じなので、この点は注意が必要である。

6.10 版の主題分類チェックボックスは、7.3 α 版では選択ボタンに変更されており^[5]、クリックすると当該主題に属する論文を出版年ごとに包括したものや、下位分類^[6]へのリンクが張られている。このように 7.3 α 版では、出版年や下位分類を順次たどることで、論文をブラウジングできるしくみが構築されている。

さらに、7.3 α 版では、地域や時代からも論文をブラウジングすることができる。検索画面のトップページには、分類に続けて「アジア全体・東アジア・中国・アジア以外・アメリカ・ロシ

ヤ・ヨーロッパ・その他」という地域分類ボタンがあり、これらをクリックすると、当該地域を時代で細分した項目の一覧画面に遷移する。一覧画面に表示される下位分類をさらにたどることで、地域や時代に特化して論文をブラウジングできるよう設計されている。

個別の書誌については、6.10 版とは異なり、1 書誌 1Web ページとなっている。検索結果一覧から詳細を見たい論文名をクリックすると、当該論文の詳細書誌画面に遷移する。

一覧画面、詳細画面ともに、著者、掲載誌、分類等にリンクが張られおり、これらをクリックすると、当該カテゴリで名寄せされた一覧データが表示される。たとえば、著者名をクリックすると、当該著者の著作が一覧できるのである（図 14）。

ただし、図の例からもわかるように、7.3 α 版の結果表示は、CHINA3 を彷彿とさせるコマンドラインで得た検索結果のようで、Web 用のユーザーインターフェースとしてはまだ改良の余地がある。最初に指摘したように、著者名検索ができないというのも難点だろう。こういった点を考慮

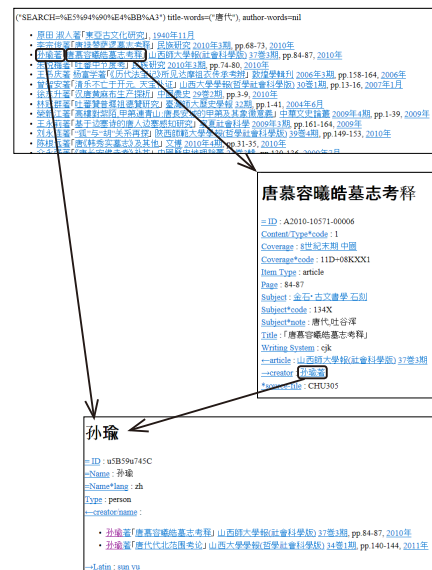


図 14 各画面の関係

してバージョンにテスト版を示す α が付けられているものと考えられる。今後の改良に期待したい。

✳ おわりに

以上、『電脳中国学入門』出版以後の変化と、同書では簡略な説明で済まされている論文検索サイトについて紹介と寸評を試みた。検索できるデータや閲覧できる論文本文が年々増加し、研究環境は便利になってきている。

しかし、改めて内容を検討して感じることは、いずれのデータベースも完全無欠ではないということである。どのような主旨で収集されているデータなのかを理解して使用する姿勢が必要なのはもちろんのこと、場合によっては旧来の冊子体などアナログデータにもあたってみなければならぬだろう。この意味で、各種研究入門書や学術雑誌の近刊総覧などの意義も決して失われたわけではないことを付け加えておく。

注

- [1] リポジトリ (repository) ないしは機関リポジトリとは、各機関で生産された研究・教育等に関する業績・情報を公開するためのオープンアクセスなプラットフォームのこと。
- [2] CiNii については本誌本号掲載の犬向一輝「学術情報共有プラットフォームとしての CiNii」も参照のこと。
- [3] このほか、雑誌記事索引の採録誌は学術雑誌とは限らないので、CiNii には一部の大衆誌の記事データも含まれてしまっている。
- [4] 『漢字文献情報処理研究』11, 2010, p82-95
- [5] 7.3 α 版の分類には 6.11 版にあった「学会消息」がなく、合計 17 分類に変更されている。
- [6] データ上は「包摂」とされているので、複数の分類を統合したものかもしれないが、最上位の「歴史」の下に「歴史 通論」「歴史 史籍」といった形で分類が並んでいるので、下位分類と表現しても差し支えないと判断した。

中国の論文を探す

千田 大介 (ちだ だいすけ)

＊ はじめに

つい十数年前まで、中国で刊行された論文を探すのには何かと苦勞がつきまとった。目録は我が国で作られた『東洋学文献類目』が頼りだった上に、所属大学が論文掲載雑誌を購入していなかったら、他大学の収蔵状況を知人に聞いたり、直接調査に赴いたりしなければならなかった。

しかし1990年代後半以降、経済発展と情報化が平行して進展する中、中国が後発優位を生かして論文のデジタル化・ネットワーク化を一気に推進したお陰で、いまや居ながらにしてインターネット経由で論文を調査し閲覧することが可能になっている。

中国の電子ジャーナルを牽引してきたのはCNKIである。我が国でも東方書店が代理店となりライセンスやプリペイドカードを販売しているので、CNKIをメインで使う人が多いと思われる。そのため、本稿もCNKIを中心に解説する。

しかし中国では現在、いくつかの商用電子ジャーナルが有償でサービスを提供している。CNKIを含め、いずれも執筆者や出版社と契約を交わして論文を収集・収録しているので、何らかの理由で契約が締結できずに収録されない雑誌や論文が発生することは避けられず、1つのサービスを検索して事足りるとせず、複数の電子

ジャーナルを検索して万全を期することも必要となる。このため、万方数拠・維普網についても、あわせて取り上げることとしたい。

＊ CNKI で論文を入手する

<http://www.cnki.net/>

● 使用の準備

CNKI (China National Knowledge Infrastructure、中国知网) は清華大学およびその関連企業である清華同方社によって開発された、学術情報サービスである。1999年に構築が始まり、今や学術論文のほか、学位論文・新聞・会議資料・統計データ・工具書など、さまざまな情報を提供している。我が国では東方書店が代理店を務めている。

CNKIでは、大学・機関などを対象としたライセンス契約と、個人がアカウントを取得し、従量課金で論文をダウンロードする個人契約の、2つの契約形態が用意されている。

ライセンス契約は、大学や機関が論文のカテゴリごとの契約することができ、LANを通じて自由に論文の検索・ダウンロードができるようになる。その場合、アクセス方法については図書館などの説明に従うことになる。

所属機関がライセンス契約していない場合は、個人アカウントを取得して利用することになる。

ライセンス契約していても、契約していないカテゴリの文書はダウンロードできないし、CNKIが提供している電子ジャーナル以外のサービスやリソース——学位論文・新聞・統計データ・特許・規格・工具書・国学宝典などについても、我が国での大学・機関で契約しているところはほとんど無いだろう。それらCNKIの全てのリソースを使いこなすため、やはり個人アカウントを取得しておいた方が良いと考える。

● 個人アカウントの取得

個人アカウントは、以下の手順で取得する（日本版のCNKIカードを購入した場合は、購入時の指示に従っていただきたい）。



- ① CNKI トップページ右上の「注册」（登録）をクリックする。



- ②「用户名」（ユーザー名。アルファベット・数字・アンダーバーの組み合わせ6～64字）、「密码」（パスワード。6字以上）、「确认密码」（パスワードの確認入力）、「您的邮箱」（メールアドレス）、「验证码」（キャプチャ。右側に画像で表示されるコード番号）を入力する。
- ③「我已阅读CNKI用户注册协议」（CNKIのユーザー登録規定を読みました）をチェックする。
- ④「立即注册」（直ちに登録）をクリックする。

後は、指定したアドレスに届いたメール（中国語）

の指示に従って、アカウントを有効にする。

● ログイン

アカウントの登録が完了したら、CNKIにログインしてみよう。



- ① トップページ右上の「登录」（ログイン）をクリックする。



- ②「用户名」（ID）、「密码」（パスワード）を入力する。
- ③「登录」をクリックする。

これで、CNKIの各種サービスが利用できるようになる。なお、ログインしていないと、論文の検索やサマリーの閲覧はできるが、ダウンロードはできない。

個人利用の場合、論文のダウンロードに対して従量課金される。金額は、一般の論文が1ページあたり0.5元、新作が1ページ1元などとなっており、それがアカウントにチャージされた金額ポイントから差し引かれる。

ポイントのチャージには、中国国内の各種決済サービス、中国国内の各種決済サービス、中国移動神州行携帯電話料金プリペイドカード（神州行充值卡）、携帯電話SMS送金などの方法があるが、日本国内では東方書店で購入できるプリペイドカード「CNKIカード」を使うのが一般的だろう。

なお、中国国内で販売されているCNKIカードは、以前、国外での使用が禁止されていたが、現在ではその縛りがなくなっているため、中国に渡航した際に購入するのも1つの方法である^[1]。ま

た、中国移動のプリペイドカードは、中国の街角のマガジンスタンドで手軽に手に入り、後で採りあげる万方数拠や維普網などの電子ジャーナルのチャージにも使えるので、中国国内でCNKIカードを買いに行く時間がない、あるいは人に頼んで買ってきてもらうというような場合には、こちらの方が便利だろう^[2]。

ポイントをチャージするには、トップページ右上の「充值中心」をクリックして、以下の画面を開く。



ここで支払い方法を選択する。CNKIカードの場合は、右上の「知网卡充值」になる（日本版CNKIカードを購入した場合は、購入時の説明に従っていただきたい）。移動した先の画面で、プリペイドカードのID・パスワードといった必要な情報を入力する。

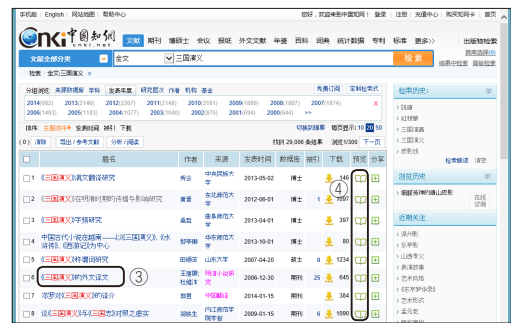
● 論文の検索とダウンロード

まず、基本的な論文検索とダウンロードの手順を解説しよう。


CNKIではトップページの検索ボックスにキーワードを入力して、学術論文を検索することができる。なお、ライセンス版の場合は、トップページとして別の画面が表示される（次節参照）。その場合も、検索からダウンロードの流れ自体は大差ない。



- ① 検索語を入力する。
- ② 「检索」（検索）をクリックする。



③ 検索結果一覧が表示されるので、閲覧したい論文の題名をクリックする。

④  をクリックすると、論文の1/10程度をプレビューできる。



⑤ 論文の作者・サマリー・掲載誌などの情報が表示される。「PDF 下载」（PDF 形式でダウンロード）をクリックする。



⑥ ダウンロードの確認画面が開く。個人アカウントの場合は金額を確認した上で、「下载」（ダウンロード）をクリックする。ファイル名が文字化けすることが多いので、そのときは名前をつけて保存する。

CNKIでは、論文をPDFと独自形式のCAJ形式で提供しているが、CAJファイルの閲覧には独自ソフトのインストールが必要になるので、PDF形式の方が使いやすいだろう。

PDFファイルは、DTPソフトから直接PDF出

力されたものもあれば、書影をスキャンした画像 PDF もあるが、後者であっても透明テキストが貼り込まれているものが多く、論文の必要箇所をコピー＆ペーストすることができる。

また、CNKI では簡体字・繁体字を同一視検索してくれるが、常用漢字体などの異体字には対応していないので、検索語は中国語 IME で入力した方がよい。

④の画面を下の方に移動すると、当該論文について、引用論文、引用された論文、テーマが近い論文などの一覧が表示されており、芋づる式に関連論文を探すことができる。

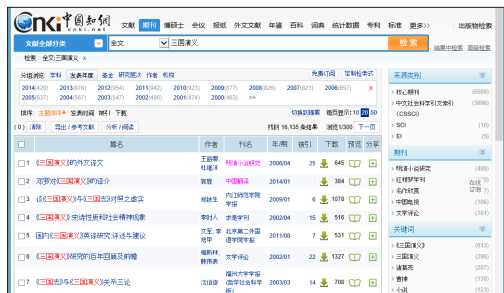
● 文献の種類を指定する

CNKI トップページから検索すると、ときに大量の文献にヒットして途方に暮れることがある。前の検索結果でも、雑誌論文のみならず学位論文（「数据库」が「博士 / 硕士」になっている）にまでヒットしている。自分と同じ題材を扱った論文ならば全て読むのが当然だが、隣接分野や特定の事項について調べたい場合には、大量の検索結果から本当に役に立つ論文や、必読の論文を見つけなくてはならない。その作業を効率化するテクニックを、いくつか紹介していく。

まず、文献の種類を絞り込む。



①検索ボックス上の「期刊」をクリックする。



②検索結果に雑誌論文だけが表示される。

このように、検索ボックスの上の文献の種類をクリックすることで、対象を絞り込むことができる。「文献」で全種類を検索できる。

はじめから雑誌論文だけを検索したいのならば、トップページの「期刊」をクリックして雑誌論文検索ページを使ってもよい。



①クリックする。



②雑誌論文検索ページが開く。

CNKI をライセンス契約で使っていると、はじめにこの画面が表示され、契約している範囲に限って検索するように設定されていると思う。値段を気にせずに使えるのは嬉しいが、そのままでは契約していない学術領域や論文の種類（学位論文・会議論文など）の論文が検索できないというデメリットもある。より広範に論文を調査するためには、ライセンス版のトップページだけではなく、CNKI トップページからも検索してみた方がよからう。

● 論文の学術領域を指定する

文学の論文だけ調査したいのに史学の論文までヒットしてしまう、というような場合がある。そのときには、以下の手順で学術領域を指定する。





- ①「学科」(学術領域)をクリックする。
- ②任意の学術領域名をクリックする。



- ③当該学術領域の論文だけが表示される。

雑誌論文の場合は、前述の雑誌論文検索ページ左側のツリーでジャンルを指定する方法もある。

● 学術論文だけを検索する

CNKIには紀要や学会誌などにとどまらず、一般誌の記事までもが収録されており、検索語によってはそうした雑誌記事が大量にヒットしてしまう。その場合には、研究レベルを指定する。



- ①「研究层次」(研究レベル)をクリックする。



- ②「基础研究(社科)」(基礎研究・社会科学)をクリックする。
- ③人文系の研究論文に絞り込まれる。

「大众文化」(大衆文化)などを選択すると、逆に一般誌の記事だけを拾うこともできる。さまざまな検索条件を指定して試してみたい。

● 検索対象を調整する

近年、中国の紀要や学術雑誌の掲載論文は、キーワードやサマリーを付した所定のフォーマットで

書かれるのが一般的になっている。そうした要素を検索対象とすることもできる。



- ①クリックする。



- ②「关键词」(キーワード)・「摘要」(サマリー)などをクリックして選択する。
- ③「检索」をクリックする。



- ④検索結果が絞り込まれる。

キーワードやサマリーは執筆者が設定するので、検索語について取り扱った論文を効率的に検索できると思われる(もっとも、それらがいい加減に書かれている論文も少なくないようだが)。

● 複数のキーワードを指定する

CNKIでは複数のキーワードを指定した複雑な検索条件を指定することもできる。



- ①「高级检索」(詳細検索)をクリックする。



- ②詳細検索画面が開く。

詳細検索では、複数の検索語を指定できる
①。検索条件は「并含」(and)・「或含」(or)・「不
 含」(not) を指定でき **②**、また「精确」(完全一致)・
 「模糊」(あいまい) を選択できる **③**。+ ボタン **④**
 をクリッして検索欄を追加できる。

また、「**专业检索**」(プロフェッショナル検索)タブ(5)をクリックすると、検索条件式による検索ができるし、「**作者发文检索**」(作者発表文書検索)・「**句子检索**」(フレーズ検索)なども用意されている。

● 評価の高い論文を探す

論文の評価をはかる重要な指標が被引用率である。CNKIでは以下の手順で引用頻度の高い論文を探することができる。



①「被引」(被引用数)をクリックする。

[illegible]

②検索結果が被引用数順に並べ替えられる。

同様に「ダウンロード」(ダウンロード)で並べ替えると、ダウンロード数の多いよく読まれている論文を探すことができる。

以上にとどまらず、雑誌や研究者の所属機関名で検索したり、論文の発表年代を指定したりといった、さまざまな機能があり、それらを組み合わせることもできる。

また、CNKI では論文だけでなく、統計データ・特許・法規などさまざまなデータベースを提供している。なかでも、数多くの辞書や工具書を一括検索できる「工具書」、古典籍を全文検索できる「国学宝典」（開発元是北京国学時代文化伝播社）なども非常に便利なサービスなので、是非試してみたい。

✿ 万方数抛

<http://www.wanfangdata.com.cn/>

● 万^{データ}方数拠とは

万方数拠は中国科学技術省の学術雑誌のデジタル化プロジェクトによって構築されたデータベースで、万方数拠社が運営している¹³⁾。電子ジャーナルのほか、学位論文・特許・法規・医薬などのデータベースを提供している。

万方数拠は CNKI に比べて知名度は劣るものの、CNKI でダウンロードできない論文・雑誌が登録されているケースもあるので、やはり中国学研究に欠かせない電子ジャーナルであるといえる。

● 使用の準備

万方数拠にもライセンス契約・個人契約の2種類の契約形態があるが、まだ国内で正規販売されていないので、個人アカウントを取得しポイントをチャージして使うことになる。

アカウントは以下の手順で取得する^[4]。



①トップページ右上の「**注册**」をクリックする。



- ②左側の「必填項」(記入必須項目)の「用户名」(ユーザー名)・「用户邮箱」(メールアドレス)・「用户密码」(パスワード)・「确认密码」(パスワードの確認)を入力する。

- ③「确认提交」(確認して提出)をクリックする。

登録成功のメッセージが出たら完了である。確認メールは送られてこない。トップページ右上の「登录」をクリックしてログインし、アカウントが登録されたことを確認しよう。

支払いだが、論文のダウンロードごとに代金を支付宝・銀聯カード・携帯電話SMSなどで支払う方法と、アカウントに金額ポイントをチャージしておきダウンロードごとにポイントが差し引かれる方法がある。後者では、中国の各種決済サービス・独自のプリペイドカード・中国移动神州行プリペイドカードなどが使えるが、いずれも日本からの直接支払いが少々難しい。やはり中国で中国移动神州行プリペイドカードを購入しておき、個人アカウントにチャージするのが現実的だろう。

金額ポイントをチャージするには、トップページ右上の「充值」(チャージ)をクリックする。チャージ方法の一覧メニューが表示されるので選択し、指示に従って手続きする。

● 論文の検索とダウンロード

万方数拠も、必要な論文を検索してダウンロードすることになる。



- ①検索キーワードを入力する。

- ②「检索」をクリックする。



- ③検索結果が表示される。必要な論文のタイトルをクリックする。




- ④論文の情報画面が開く。「下载全文」(全文をダウンロード)をクリックする。



- ⑤支払い画面が開く。金額を確認し「万方数据“我的钱包”」(万方数拠「私の財布」)の「购买」(購入)をクリックする。

- ⑥論文のファイルがダウンロードされる。

価格は、論文1篇3元で、ページ数によってはCNKIよりも安価になる^[5]。

検索結果画面で論文タイトルの前のアイコンがになっているものは、サマリーの閲覧しか提供されていない。またダウンロードしたPDFファイルは、書影のスキャンであっても透明のテキストが貼りつけられているものが大半であり、テキストデータをコピー＆ペーストすることができる。

● 検索・絞り込み機能

万方数拠は一般誌を収録していないので、検索結果はCNKIほど膨大にはならないが、キーワードによっては、やはり大量の論文にヒットしてしまう。

万方数拠も、さまざまな検索・絞り込みのため



の機能を備えている。

- ①論文タイプの選択。
- ②学術領域の選択。
- ③発表時期の選択。
- ④全ての結果表示・全文掲載のみ表示の切り替え。
- ⑤検索結果の表示順序の選択。「新論文優先」(新作論文優先)・「经典论文优先」(定番論文優先)などのほか、右の「其他▼」をクリックすると、「仅被引次数」(被引用数のみ)なども選択できる。
- ⑥検索結果から絞り込み検索。
- ⑦アドバンス検索。複数検索語による検索が可能。

以上のように、万方数拠も CNKI とほぼ同等の検索・絞り込み機能を持っている。ただ被引用数については、万方数拠と CNKI とでは統計の範囲が異なるためか、相違が見られる。

● 論文以外のサービス

万方数拠では、検索ボックスの上のメニューに見えるように、電子ジャーナルのほかにもいくつかのサービスを提供している。「专利」(特許)・「标准」(規格)・「法规」(法規)など、CNKI と同じものが多いが、万方数拠独自のものも見られる。「学者」は、研究者の所属と研究キーワードを掲載しており、また執筆論文の抽出もできる。ただ

し、同一人物を旧所属機関・現所属機関で2人に数えている例もあるので、いささか注意が必要である。

特筆されるのは「新方志」(新地方志)で、中華人民共和国成立後に編纂された地方志の類を全文検索・有償ダウンロードできる。我が国では見ることが困難な資料も多数収録されており、現代中国の社会・文化を研究する人にとって必見である。

また、トップページ上の「相似性検測」(相似度チェック)をクリックすると、ある論文・文章と既発表論文との類似度の検査、いわゆる論文のコピペチェックサービスに移動する。



チェック対象となった文章が、万方数拠に登録された学術論文とどれだけ似ているかを自動で付き合わせるもので、ライセンス契約による雑誌掲載論文や学位論文の一括チェックを主なターゲットとしているようだが、個人でも利用できる。その場合、価格は1万字あたり10元になる。

漏れ聞く噂によれば、中国におけるコピペ論文の猖獗は我が国の比ではないようである。このため中国語の論文、特に学位論文などを引用する際には、それがオリジナルであるのかを事前にチェックするくらいの慎重さが必要となろう。また、我が国でも中国語論文の投稿を受け付ける学術誌が多いし、学位論文が中国語で提出されるケースもあるので、昨今世間を騒がせている研究不正の防止、ひいては研究・教育や学術誌のクオリティ維持のために、こうしたサービスを積極的に活用すべきであろう。

以上のように、万方数拠はCNKIの補完にとどまらないさまざまなコンテンツや機能をもっている。我が国における使いにくさを解消するために、国内中国書籍店には、是非とも万方数拠のライセンスやプリペイドカードの正規輸入販売をお願いしたい。

✳ 維普網

<http://www.cqvip.com/>

● 維普網とは

維普網は重慶維普資訊社が運営する電子ジャーナルである。重慶維普資訊社の前身は中国科学技术情報研究所重慶分所データベース研究センターである。万方数拠社も前身は中国科学技术情報研究所万方データネットワークセンターであるので、両者はともに中国科学技術省系であることになる。

維普網の電子ジャーナルとしての機能はCNKIや万方数拠と大差ないが、学位論文や特許・法規などのサービスは提供されていない。一方で、Google Scholarに協力しており、また海外への展開を視野に入れているのか、英語や日本語の解説ページが設けられており、日本語で機関ライセンス契約の試用を申し込むことができる。また、個人アカウントのポイントチャージに、中国国内の決済サービスや中国移动のプリペイドカードのみならず、VISA・MasterカードやPaypalが使える。

● 使用の準備

維普網の個人アカウント取得手順は、以下のとおり。



①トップページ左上の「注册」をクリックする。



②必要事項を記入する。ユーザー名はメールアドレスになる。

③「确认提交」をクリックする。

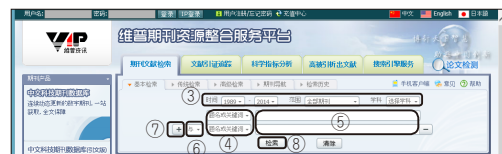
維普網でも登録完了通知メールは送られて来ない。登録完了画面が表示されたら成功なので、トップページ左上の「登录」をクリックして、入力したユーザー名・パスワードでログインできることを確認しておこう。

● 論文の検索とダウンロード

維普網で論文を検索する際には、通常のトップページではなく、「专业版」(プロ版)を使った方がよい。基本的な使い方は以下になる。



①トップページ左上の「专业版」をクリックする。



②プロ版が開く。

③「时间」(論文の発表時期)・「范围」(雑誌のタイプ)・「学科」(学術領域)などを設定できる。

④検索項目を選ぶ。デフォルトは「题名或关键词」(タイトルもしくはキーワード)。

⑤キーワードを入力する。デフォルトでは検索語ボックスは2つ。

⑥複数キーワードの検索条件「与」(and)・「或」

(or)・「非」(not)を選択できる。

⑦クリックして検索語ボックスを追加できる。

⑧クリックして検索を実行する。



⑨検索結果が表示される。読みたい論文のタイトルをクリックする。



⑩論文の情報が表示される。作者・掲載誌などの関連情報や引用文献一覧なども表示される。ダウンロードする場合は、「下载全文」(全文をダウンロード)をクリックする。

⑪「在线阅读」(オンラインで閲覧)をクリックすると、冒頭1ページを無償で閲覧できる。



⑫ログインしていない場合はログインする。ダウンロード画面が表示されたら、金額を確認して「账户余额支付」(アカウントの残額から支払い)をクリックする。



⑬ダウンロードサーバの選択画面が表示されるので、任意のサーバを選択する(「高速下载」など)。ファイル名が文字化けすることが多いので、その場合は名前を付けて保存する。

論文のダウンロードは、4ページ未満が1ページ0.5元、4ページ以上が一律1篇2元で、ページ数にもよるが、他の電子ジャーナルに比べて若干安い。

● その他の機能とサービス

維普網も論文・雑誌などの被引用データを提供しているが、しかしそれらは機関ライセンス契約でないと使えない。このため個人契約では、専ら論文の検索・閲覧用として使うことになる。

また、維普網も論文の類似度チェック機能を提供している。こちらは千字あたり1元で、万方数拠よりも文字数の単位が小さい分、個人使用に向いているといえよう。

✳ 国家図書館で学位论文を閲覧する

● 有償サービスと無償サービス

中国では、複数のサービスを通じて学位论文が提供されている。1つは有償の電子ジャーナルで、前述のCNKI・万方数拠などがそれである。日本の大学・研究機関で学位论文までライセンス契約しているところはまれであり、個人アカウントで購入すると30～50元ほどのコストがかかる。

もう1つが中国国家図書館のWebサイトで、全国の博士論文が無償公開されている。ただし、初めに書いておくと、国家図書館の博士論文に

はFlashが使われており、閲覧は問題ないのだが、プリントアウト機能がない。このため、印刷してじっくり読みたい場合には、有償の電子ジャーナルを使うことになる。また、国家図書館の博士論文検索は、タイトル・著者名など複数項目についてキーワードを指定できるものの、サマリーや全文の検索機能・引用論文機能などを持つ電子ジャーナルに比べて、いささか見劣りがする。

目的の博士論文がはっきりしている場合は直接国家図書館で検索すればよいが、そうでない場合はまずCNKIや万方数拠で検索してから閲覧した方がよからう。

● 国家図書館のアカウントを取得する

国家図書館では2・3年前から、デジタルコンテンツの利用に個人のアカウントが必要になっている。このアカウントの登録過程で中国の身分証明書番号の入力を求められるので、外国人にはアカウントが取得できないかのように思える。

筆者も当初、そう考えていたが、国家図書館で司書から身分証明書番号は入力しなくても大丈夫だと聞き、試したところ、確かに問題なくアカウントを取得できた。以下、その手順を解説する。

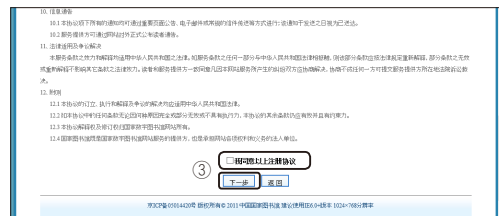


- ① 中国国家図書館のWebサイト、<http://www.nlc.gov.cn/>を開き、画面右側の「读者门户登录注册」（読者ポータルログイン・登録）をクリックする。



- ② 読者ポータルのログイン画面が開く。右側

の「注册新用户」（新規ユーザー登録）をクリックする。



- ③ 規約画面が開くので、画面下方の「我同意以上注册协议」（以上の規約に同意して登録します）をチェックし、「下一步」（進む）をクリックする。



- ④ 「邮箱激活」（電子メールによるアクティベート）タブをクリックする。



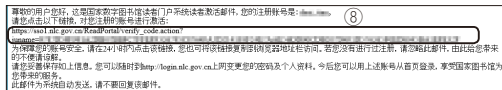
- ⑤ 「*」の付いている必須項目、「登录账号」（登録アカウント。6～30字、アルファベット・数字・アンダーバーが使用可能）・「密码」（パスワード。6～32字、アルファベット・数字のみ）・「确认密码」（パスワードの確認）・「电子邮件」（電子メール）・「验证码」（キャプチャ。隣に画像で表示される文字列）を入力する。

- ⑥ 「提交注册信息」（登録情報を提出する）をクリックする。



- ⑦ 個人情報登録画面が表示されるが、何も入力せずに下の「跳过」（スキップ）をクリック

クする。



⑧確認メール発送通知画面が表示され、登録確認メールが届くので、メールに書かれたURLにアクセスし、アカウントをアクティベートする。

これで登録完了のメッセージが表示され、アカウントが使えるようになる。

● 博士論文を検索・閲覧する

②のログイン画面で、登録したアカウントとパス・キャプチャ（入力ボックス内をクリックしてカーソルを移動させると画像が表示される）を入力し、国家図書館読者ポータルにログインする。すると、以下の画面が開く。



ここを起点として、国家図書館が提供する各種コンテンツを利用することができる。
博士論文は以下の手順で閲覧する。



①左メニューの「电子论文」（電子論文）にマウスカーソルをあわせ、展開されたメニューの「博士论文」（博士論文）をクリックする。



②博士論文の検索ページが開く。任意の項目にキーワードを入力し、「检索」（検索）をクリックする。



③検索結果が表示される。必要な論文のタイトルをクリックする。




④論文の情報画面が開く。画面下方の「在线阅读」（オンライン閲覧）をクリックする。



⑤論文が表示される。

表示の拡大・縮小、ページの移動などは、上の

メニューバーで行う。また、をクリックして全画面表示した方が読みやすいだろう。

● 国家図書館のデジタルコンテンツ

前述のように国家図書館のアカウントを取得すると、博士論文に限らず、以下のような膨大なデジタルコンテンツが利用できるようになる。

电子图书	民国法律
	民国图书
电子期刊	民国期刊
电子报纸	方正 apabi 报纸资源
古籍	数字方志
	西夏文献
	宋人文集
	西夏论著
	碑帖菁华
	年画撷英
	前尘旧影
	中国学网络
	中国学汉学家
	中华古籍善本联合书目
	敦煌遗珍
	哈佛大学善本特藏
	古代典籍
	甲骨世界
	数字善本
地方馆资源	地方馆藏特色资源
	地方志
	民国文献
	家谱
	老照片
	地方视频资料
	少数民族资源
	少年儿童资源
	地方非物质文化遗产
	年画
	科普科教
	历史文化

少儿资源	少儿连环画
	少儿文津奖图书
	少儿新闻总署推荐图书
	少儿图书

民国時期の図書・雑誌、国家図書館や地方の図書館が所蔵する地方志、さらには善本まで、画像データを中心とする大量のコンテンツが公開されており、さまざまな分野の研究に教育に、大いに活用できることだろう。

このほか、現代の図書や視聴覚資料もあるが、館外からは一部分しか閲覧・視聴できないものが多い。

✳ 電子ジャーナルと言論統制問題

全体主義国家である中国では、周知のように大学・学術であっても、共産党政府の統制下に置かれており、往々にして研究に対して政治的な正しさが求められる。本稿で紹介してきた電子ジャーナルも、その例外ではない。

例えば、ノーベル平和賞受賞者である人権活動家「劉曉波」氏を検索すると、CNKI・維普網ではエラーとなってページが表示されない。一方、万方数拠では多くの同姓同名研究者の論文なども含め、執筆時で709篇ヒットする。特定の話題を取り上げた論文を排除するのも問題であるし、同姓同名であるために論文が電子ジャーナルから排除されてしまうのも酷い話である。

また、2014年に拘束され国家分裂罪で起訴されたウイグル族の経済学者、イリハム・トフティ（伊力哈木・土赫提）氏を検索してみても、CNKI・万方数拠・維普網いずれも、数多くあるはずのその著作にまったくヒットしない。

こうした電子ジャーナルのあり方は、中国の国情を反映したものであるとはいえ、学問の独立という原則に大いに抵触する露骨な言論統制であることは論を俟たない。

現代中国研究においては、政治的に敏感な記事の事後的な書き換えがしばしば見られる中国の新聞記事データベースに全面的に依存してはならず、

紙版の新聞の保存・参照が欠かせないとされるが、中国学の学術論文についても、同様のことが行われる危険性を考慮すべきだろう。従って、政治的に敏感な問題に触れる蓋然性の高い近現代の社会・歴史・文化・思想などに関する研究はいうまでもなく、古代・古典研究であっても、政治的立ち位置による研究者への圧力といった問題が想定されるので、重要な学術誌については紙媒体をも購読・保存しておく必要がある。

一方、中国の研究者やライターの中には、雑誌掲載時に書き換えや削除を余儀なくされた箇所をブログなどに公表する人もいる。このため、電子ジャーナルで参考となる論文を見つけたら、Googleなどで執筆者名や論文のタイトルを検索し、関連情報がないかも確認しておいた方がよからう。

ところで、側聞するところでは、我が国の一部の学会が会誌論文のCNKIへの登録を目指しているという。研究情報の発信という観点からは、中国において我が国の中国学論文が読まれ、引用される機会が増える効果が期待され、望ましいものであるといえるが、他方、それは中国の言論統制体制下に進んで飛び込むことをも意味している。学問の独立性、および我が国中国学の独自性が損なわれないように、細心の注意を払う必要があろう^[6]。

注

- [1] 中国国内の販売店情報は、以下のページで確認できる。
<http://ec.cnki.net/skwd/skwd1.htm>。なお本稿で引用したWebサイト・ページは、いずれも2014年9月15日に確認している。
- [2] 中国移动のプリペイドカードには、全国と地方限定の2種類があり、確認していないが地方版のカードではチャージできない危険性がある。北京・上海などの大都市の場合は問題ないが、地方都市で購入する場合は注意した方がよい。
- [3] CNKIが教育省系、万方が科学技術省系ということであり、似たような電子ジャーナルが複数作られた原因が透けてみえる。
- [4] 筆者のWindows 8.1+Internet Explorer 11環境ではエラーが出てアカウント登録画面が表示されなかった。その場合はGoogle Chromeなど他のブラウザで試していただきたい。
- [5] 万方数拠のダウンロード画面には、価格が中国国内向けであると注記されている。現在は日本国内からもその価格でダウンロードできるが、今後、変動があるかも知れない。
- [6] 万方数拠・維普網にはユーザーが論文をアップロードする機能がある。日本の研究者がそうした機能を使って研究情報を発信することもできる。

※ 本稿で言及したWebサイト・ページはいずれも2014年9月15日に最終確認した。

台湾・香港の論文データベース

二階堂 善弘（にかいどう よしひろ）

＊ 台湾の修士・博士論文データベース

台湾の学術論文においては、修士論文や博士論文が引用されていることが多い。台湾においては、修士論文や博士論文は早くから図書館などで公開されており、利用しやすい面があった。

1990年代、筆者が台北に留学していた時、国家図書館（当時は国立中央図書館）の開架書庫に行けば、修士論文と博士論文がずらっと並べられており、自由に閲覧・コピーが可能であった。また幾つかの大学の附属図書館においても、博士論文が無造作に置かれているのを見たことがある。台湾ではとにかく修士・博士論文が利用しやすい環境なのである。

また台湾では、時に非常に優れた修士論文が存在する。まれに雑誌掲載論文を凌駕するといっても過言ではないほどの修士論文に行き当たることもある。そのため、一般の学術論文や学術書でもよく修士論文が引用されたりする。

筆者に関わる分野では、かつて沈淑芳氏の『封神演義研究』という論文があった。これは東呉大学の修士論文であったが、当時においては恐らく最も優れた『封神演義』に関する総合的な研究であった。しかしその後その内容は雑誌論文にもならず、かつ書籍として刊行されることもなく、非常に入手しづらいものであった。これも筆者は国

家図書館で懸命にコピーしたりしていたことがある。ほかにも、多くの優れた修士論文を国家図書館で見えており、大いに活用させてもらった。

そのような経緯をふまえてか、現在台湾の国家図書館で公開されている「台湾博碩士論文知識加値系統 National Digital Library of Theses and Dissertation in Taiwan」（台湾博士・修士論文利用データベース）も、非常に利用しやすいデータベースである。これを使えば、インターネット経由で修士論文・博士論文が簡単にダウンロードできる。台湾の論文を参照することの多い筆者にとっては非常にありがたいデータベースである。

ただ、このデータベースでは、近年の論文はPDFで全文がダウンロード可能であるが、古いものについては対応していないので、注意が必要である。

＊ 「台湾博碩士論文系統」の利用

● 台湾博碩士論文知識加値系統

<http://ndltd.ncl.edu.tw/cgi-bin/gs32/gswweb.cgi>

上記URLは少々長いので、台湾の国家図書館のサイト <http://www.ncl.edu.tw/> からアクセスすることをお勧めする。国家図書館トップページで「学術研究」の欄を開く。すると幾つかのデータ



台湾国家図書館のサイト

ベースの中に「台湾博碩士論文知識加値系統」が存在する。画面中央の「熱門資源」のところからもリンクしている。

初めて利用する場合は、画面右上の「註冊」をクリックしてアカウントを取得する必要がある。メールアドレスとパスワード、それに氏名（ハンドルネーム）などを入力して申し込む。登録確認メールに記載された URL をクリックすれば、すぐにデータベースが利用可能になる。その後は登録したアカウントで何度でも利用できる。

検索画面は簡易検索と詳細検索があるが、だいたい簡易検索において探すことが可能であると思う。



「台湾博碩士論文知識加値系統」の検索画面



「哪吒」で検索した結果

「論文の名称」「執筆した院生の氏名」「指導教員の氏名」「口頭試問委員の氏名」「キーワード」「参考文献」などで検索可能である。幾つか試してみたほうが、より正確に検索できると思う。

とりあえず例として、「哪吒太子」の「哪吒」で検索してみた。

タイトルに「哪吒」を含む論文は9本程度あることがわかる。このうち5本の論文については、PDFの電子データがダウンロード可能である。そのデータには「電子文本」のマークがあるので、それをクリックすればダウンロードページに飛ぶ。ここでは画像の数字をパスワードとは別に入力する必要がある。ダウンロードは1本であれば数秒で終わる。

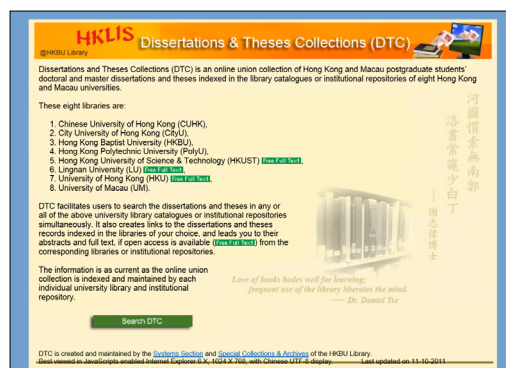
収録件数は現在約57万件とされている。このように台湾の修士論文・博士論文の利用は非常に容易になった。

✿ 香港の修士論文・修士論文データベース

香港にも修士論文・博士論文のデータベースは存在する。ただ、台湾のデータベースほど使い勝手はよくない。

● Dissertations and Theses Collections (DTC)

<http://library.hkbu.edu.hk/electronic/libdbs/dol.html>



Dissertations and Theses Collections (DTC)のサイト

主催しているのは香港バプテスト大学 (Hong Kong Baptist University)。

このサイトからは、香港の主要な大学の修士論文・博士論文を横断検索することが可能である。含まれる大学は、香港中文大学 (Chinese University of Hong Kong)、香港城市大学 (City University of Hong Kong)、香港理工大学 (Hong Kong Polytechnic University)、香港科技大学 (Hong Kong University of Science & Technology)、嶺南大学 (Lingnan University)、香港大学 (University of Hong Kong)、マカオ大学 (University of Macau) である。現在すべてで3万9千件とされている。ただ、タイトルの検索のみの大学が多く、電子テキストがダウンロード可能であるところはまだ多くない。

DTCは英語のサイトであるが、漢字入力で検索することも可能である。

台湾のサイトほど便利ではないが、現在どのような修士論文・博士論文が出ているかは確認する

ことが可能である。恐らく、今後中身はもっと充実していくと考えられる。

✿ 台湾の学術雑誌論文データベース

さて、それでは修士論文・博士論文以外の学術雑誌などに掲載された論文についてはどうだろうか。

こちらでも、検索については幾つかのサイトがあり、横断的に行えるようになっている。ただ、電子データのダウンロードについては、雑誌発行元の機関により対応がまちまちである。台湾の場合、大学や大学院、研究所などの紀要が一般に書店などで発売されている書籍であることもある。その場合はたぶん無料でのダウンロードが難しいのではないかと推察する。

雑誌論文を探すなら、まずは国家図書館である。

● 台湾期刊論文索引系統

<http://readopac.ncl.edu.tw/nclJournal/>

先に示した国家図書館のサイト、<http://www.ncl.edu.tw/> からたどる場合は、画面中央の「熱門資源」のところに「台湾期刊論文索引系統」があるのでクリックする。



台湾期刊論文索引系統

ここで雑誌論文を、「タイトル」「キーワード」「作者」「概要」「全文」などを指定して検索することが可能である。ここで検索した論文のタイトルをクリックし、もし論文の情報画面から全文データにリンクしていれば、それをダウンロードできる。リンクしていない場合、上の雑誌名のリンク先にURLが書いてあって、大学などのサイトで公開されていることもある。

公開されていない場合はどうするか。まずは、その雑誌が所蔵されている日本の図書館を探すべきであろう。

台湾の学術論文をアーカイブしているサイトで、再度検索してみる方法もある。これらのサイトでは、台湾のみならず、香港・マカオさらに中国本土で発行された雑誌もカバーしているところが多い。基本的に有償サービスだが、日本からのクレジットカード払いで論文のダウンロードが可能なサイトもある。それを以下で紹介する。

なお、近年台湾では公式領収書の扱いが厳格化し、クレジットカード払いでも領収書を送付する台湾の住所を求められることが多い。領収書を創世基金会などの慈善団体に寄付することを選択し、住所が自動で入力されなければ、Googleで検索してコピーする。

● 華芸線上図書館

<http://www.airitilibrary.com/>



華芸線上図書館サイト

台湾と大陸、香港・マカオを網羅した学術プラットフォームで、学術雑誌論文、修士論文・博士論文、会議論文などの検索・ダウンロードが可能。現在の登録件数は、雑誌論文が約3000万件、会議論文が約250万件、修士・博士論文が約320万件にもなる。

まず上の「加入會員」からユーザー登録を行う必要がある。論文のダウンロードは有償で、あら

かじめアカウントにチャージしたポイントで支払うほか、論文をダウンロードするごとにクレジットカードで決済することも可能である。雑誌論文・会議論文で1頁4台湾ドル、修士論文・博士論文は1本200台湾ドルだが、プロフィールで国籍が海外になっていると、雑誌論文が1本1,200台湾ドル、修士論文・博士論文が1本2,100台湾ドルになる。安い価格ではないので、無償公開されているものや、国内の図書館をまず探すべきだと考える。

検索結果は、台湾の雑誌論文、会議論文、修士・博士論文、大陸の雑誌論文、修士・博士論文、電子書籍などに分かれて表示される。

ここでも「哪吒」で検索してみよう。



華芸線上図書館の「哪吒」検索結果

台湾の雑誌論文が18件ヒットした。修士・博士論文は7件のヒットで、国家図書館でダウンロードできるのは5件だったので2件多い。論文題名の下に「加入購者車」(ショッピングカートに入れる)をクリックし、画面の上の方の「儲値&購物車」(チャージ&ショッピングカート)から購入手続きする。ショッピングカートの右の方にダウンロードに必要なポイント数(価格)が表示される。

● 台湾学智慧蔵 TAO

<http://tao.wordpedia.com/>

TAOは台湾の論文を収集したアーカイブで、原則として2000年以降の論文の全文が収録されている。

TAOもはじめて使うときに「會員服務」の「加入會員」をクリックし必要事項を記入してユーザー登録する。そして同じく「會員服務」の「個



台湾学智慧蔵 TAO サイト

人會員中心」をクリックし「付費方案」(支払い方法)でポイント通販画面になる。途中で台湾の郵送先住所の入力を求められるが、ポイントはバーチャル商品で実物が郵送されることはない、慈善団体などの住所を入れておいても問題ないと思われる。

検索はトップページ上のボックスで行うが、詳細検索もトップページ中程に用意されている。国家図書館の修士論文・博士論文データベースと一括検索することも出来る。

やはり「哪吒」で検索した結果を掲げておく。

雑誌論文が7件、修士論文・博士論文が4件ヒッ



台湾学智慧蔵の「哪吒」検索結果

トした。ヒット件数は華芸の方が多いが、おそらく2000年以降に刊行されたものを主に収録しているためであると推察される。

論文題名の後ろの「↓」アイコンから全文ダウンロードできる。クリックすると即ポイントが引かれてダウンロードが始まる。必要ポイント数(価格)はその右側に表示される。雑誌論文の場合は、1頁2台湾ドルのようだ。必要ポイント数が書かれていない修士論文・博士論文などは、論文を公開している各大学院の公開ページにリンクしているものが多い。

このように台湾の雑誌論文の入手も以前に比べると遥かに容易になっている。こうしたアーカイブは今後さらに充実し使いやすくなっていくことであろう。

韓国の学術論文の検索方法

師 茂樹 (もろ しげき)

✳ RISS とは

韓国に限らず、アジア各国の学術情報を調査するためには、国立国会図書館のリサーチ・ナビにある「アジア諸国の情報をさがす」^[1]に情報がまとまっている。論文情報の収集についてもまずはここからはじめることがおすすめである。韓国の論文であれば、「アジア関係論文の検索(朝鮮語)」^[2]に、各種サイトがリストアップされている。

しかし、リサーチ・ナビには、具体的に検索する方法についての解説はない。本稿では、代表的なサービスである 학술연구정보서비스(学術研究情報サービス、RISS)^[3]の利用方法について簡単に解説したい。



図 1

RISS は、韓国教育研究情報院が提供している CiNii と NDL-OPAC を足したような韓国の学術情報検索サービスである。RISS には、韓国の博士論文・修士論文、韓国国内外の学術雑誌に収録された学術論文について、書誌情報を検索し、本文データがあればダウンロード(無料・有料)することもできる。

インターフェースはハングルのみで、英語等他言語版は存在しないが、ハングルが得意でない人でも検索はそれほど難しくない^[4]。

✳ RISS の使い方

RISS のインターフェースは、CiNii などと同様、シンプルなものである(図2)。基本的にはテキストボックスにキーワードを入力して、右端の 검색(検索) ボタンを押せば検索ができる。



図 2

図2の①は、기본검색(基本検索)と상세검색(詳細検索)を切り替える部分である。詳細検索にすれば、テキストボックスの左にあるドロップダウンリストで전체(全体)・제목(タイトル)・저자명(著者名)・주제어(キーワード)・발행처(発行元)・초

록 (抄録)・목차 (目次) を指定して、AND・OR 検索をすることが可能である (図3)。

図3

図2の②の部分は、左から다국어입력 (多言語入力)・검색도움말 (検索ヘルプ)・검색환경설정 (検索環境設定)・최근검색목록 (最近の検索リスト) となっている。多言語入力をクリックすると、図4のようなウィンドウが表示され、ひらがな・カタカナ・ハングル・ギリシア文字・特殊なラテン文字・キリル文字などが入力できるようになる。

図4

図2の③で、検索対象を指定することができる。ここに並んだチェックボックスは、左から학위논문 (学位論文)・국내학술지논문 (国内の学術誌の論文)・해외학술지논문 (海外の学術誌の論文)・학술지 (学術誌)・단행본 (単行本)・공개강의 (公開講座) となっている。

✳ 検索上の留意点

検索はハングルで行うのが原則であるが、漢字で検索してもハングルに変換して検索してくれる。たとえば、新羅時代の僧侶の名である「義寂」で検索してみると、義寂のハングル表記의적に相当する「義賊」や「祭儀的」=제의적にもヒットする。

なお、データ中の漢字は、韓国の字体 (いわゆる旧字体) に統一されているわけではないので、注意が必要である。たとえば「對馬」で検索すると 4,303 件 (대마も検索されるので、「大麻」なども含んでしまう) となるが、新字体の「対馬」で検索しても 82 件、簡体字の「对马」でも 1 件ヒットする (以上、執筆時点の結果)。さらに言えば、「대마」のローマ字表記である daema、taema や、日本語発音のローマ字表記である tsushima などでもヒットする。ヒット数の差が出るのは、「對馬」の「對」や「对马」が韓国で使われない字体であるため、「대마」に変換されないためである。しかし、「對馬」「对马」というキーワードをタイトルや抄録などに含む論文があるため、少ないながらもヒットするのである。「對馬」と「対馬」の表記が同一論文で混在することはまずないので、検索結果はほとんど重複していない。

RISS には韓国外の論文もデータとして登録されているので、「對馬」「对马」「tsushima」などでヒットする論文は韓国語以外の言語で書かれた論文である場合が多い。しかし、これらのキーワードを含む韓国語の論文も存在する。日中韓で字体の違いがある漢字、表記にゆれがあるキーワードで検索する場合には、複数の表記を試してみることで検索漏れを減らすことができるだろう。

✳ 本文のダウンロード

RISS では、論文本文を閲覧して読むことができる。学位論文の場合であれば、検索結果一覧で各論文情報の下に「원문보기」(本文表示) というリンクがある。また、学術雑誌の検索結果の一覧には、「원문보기」(本文表示) の隣に㉞や㉟といったマークがついている (図5)。このマークは、前者が有料、後者が無料で、本文データをダウンロードできることを示している。

本文データの形式は、PDF だけでなく、TeX の DVI 形式、TIFF 形式などでも提供され、ダウンロードの際には随時 Adobe Reader・ezPDF Reader・TeXplus Viewer・Cox Guard などのソフトウェアのインストールが求められる [5]。

图 5

· 로그인 · **회원가입** · 장바구니 · 고객센터 · about RISS · 사이트맵

図6の左から2番目の회원가입(会員登録)^[6]をクリックすると会員登録のページとなる(図7)。

图 7

会員登録を済ましログインすれば、本文を閲覧
することができる。学位論文や無料の学術誌の論

图 8

有料の論文を個人で購入する場合は、韓国のオンライン決済会社 KG INICIS^[7] 経由での決済となる（ここでは説明を省略する）。

- [1] <https://rnavi.ndl.go.jp/asia/>
- [2] https://rnavi.ndl.go.jp/research_guide/entry/theme-asia-39.php
- [3] <http://www.riss.kr>
- [4] Google Chrome には閲覧中のページを翻訳して表示してくれる機能がある。韓国語は日本語と文法的・語彙的に近いので、Google の翻訳であってもかなり実用的な精度で翻訳してくれるが、RISS のサイトではダウンロードリストやチェックボックスに添えられたハンゲルをすべて翻訳してくれるわけではないので、検索用にはあまり役に立たない。
- [5] http://www.riss.kr/customer/svchelp/HelpView.do?help_id=393
- [6] 図 6 のメニューは左から ログイン (ログイン)・회원가입 (会員登録)・장바구니 (ショッピングカート)・고객센터 (お問い合わせ)・about RISS (RISS について)・사이트맵 (サイトマップ) となる。
- [7] <http://www.inicis.com/>

ベトナムの文献を探索する

矢野 正隆 (やの まさたか)

＊ はしがき

日本国内で文献を探索しようという時には、まず複数機関の所蔵資料を横断的に検索できるサイト、例えば国立情報学研究所 (NII) の CiNii Books、CiNii Articles や、国立国会図書館 (NDL) の NDLサーチ等を手がかりにすることであろう。本稿が対象とするベトナムでは、現在のところ、このような全国的な「総合目録」は存在しない^[1]。しかし、ベトナム国内における文献の目録・画像のデジタル化・オンライン化の動きは、必ずしも低調というわけではなく、以前本誌でも、漢喃文献のデジタル化について紹介したことがある (11号)。本稿では、それ以外の、図書・雑誌を中心とした情報基盤整備の現状について、簡単な解説を加えることを目的とする。

＊ ベトナム国家図書館^[2]

主な情報探索ツールとして以下の3つがある。

- ①蔵書検索 ilib opac
- ②各種データベースの全文統合検索
- ③国家書目 Thư mục quốc gia

①は国家図書館の蔵書検索システムである^[3]。

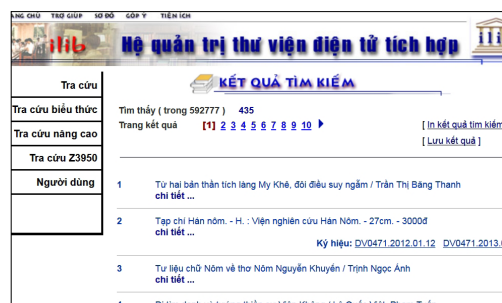


図1 国家図書館蔵書検索結果

検索対象には、図書だけでなく雑誌記事も含まれる。図1は「Tập chí Hán Nôm」を検索した結果であるが、435件ヒットしたうち、2番目に雑誌名があがっている以外は、この雑誌に掲載されている論文タイトルである。ただし、雑誌記事の採録は必ずしも最近には及ばないようで、この「Tập chí Hán Nôm」についていえば、2014年8月現在、この目録で確認できたのは2009年分までであった。また、検索語の入力は、補助記号込みにする必要がある。「Tập chí Hán Nôm」は「Tập chí Han Nom」と入力したのではヒットしない。こうした入力方法による検索結果の相違については、検索サイトによって様相が異なるので注意が必要である。

②はデジタル化された文献のテキスト全文検索である^[4]。図2は「cổ tiền」を検索した結果で、この図では途中までしか見えないが、以下の5つ

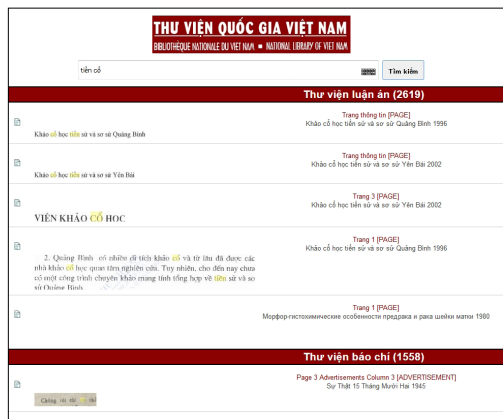


図2 各種データベースの全文統合検索^[5]

のデータベースの検索結果が順に提示されるかたちになっている。

A. 博士論文^[6]: ベトナム人が国内外で授与された博士論文のコレクションである。現在のところ2万タイトル以上がデジタル化されており、これは所蔵分のおよそ8割に相当する。このようにデジタル化が進んでいるのは、国の規定で、論文著者は印刷版と併せて電子版を納めなければならないことによる。

B. 新聞・雑誌^[7]: Gia Định báo ほか 1890 年から 1955 年までに刊行された 20 タイトルが含まれる。ただし、欠号は補充されておらず網羅的なデータベースとはなっていない。収録誌名および号数がリスト化されており、全容を把握することができる。

C. 植民地期図書^[8]: フランス大使館、フランス外務省、ベトナムのいくつかの図書館による「ベトナムに所在のある古いフランス語資料のデジタル化」プロジェクトを通じて作製されたもの。1837 年から 1953 年までの刊行物を含む。

D. 漢喃文献^[9]: 本誌 11 号で紹介したデータベースであるが、インターフェースは、この統合検索に組み込まれることによりやや変化している。採録されている 1907 タイトルについてはタイトルのブラウジングができ、また刊行年による抽出も可能である。検索は漢字やチューノムではできず、現代ベトナム語表記 (クオックグー) を用いる必要がある。

E. 地図: 26 部。ハノイ遷都 1000 年 (2010 年) を機に、ハノイに関連した関連図書・地図を、国家図書館がハノイ出版社と共同してデジタル化したものの一部。

この統合検索には含まれないが、国家図書館では、この他に、英文のベトナム関連図書やマイクロフィルム、テープ・CD・DVD 等についてもデジタル化を進めている。

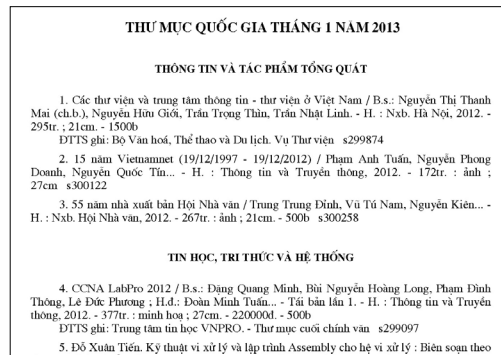


図3 国家書目 (2013 年 1 月)

③は月刊の全国書誌^[10]であり、主題ごとに著者名順にリスト化されている (図3)。これをまとめたものも年1度刊行されており、現在、月刊版は 2010 年 11 月分以降、年刊版は 2009 年版以降が PDF で閲覧できる^[11]。全体は「図書」「音楽・絵画・地図」「定期刊行物」の3部構成になっている。ただし、定期刊行物については、誌名のためのリストであり、そこに含まれる記事の情報は含まれない。現在のところベトナムではオンラインの総合目録がないため、こうした昔ながらの書目も、研究情報を得るための重要なソースとなっているようである^[12]。

✳ 社会科学図書館^[13]

ベトナムの人文科学系研究機関としては、まずベトナム社会科学翰林院 (以下 VASS^[14]) を挙げておかなければならない。この下部組織となる専門研究機関は、現在のところ、社会科学部門7機関、人文科学部門9機関、国際研究部門8機関、地方研究部門4機関、その他7機関の、計35機関を数えるが^[15]、これら研究機関にはそれぞれ図書

館が設置されており、その蔵書は、各研究領域において重要な情報基盤となっている。このうち、社会科学通信院^[16]の社会科学図書館は、フランス極東学院の蔵書の一部を引き継いだ、ベトナムにおける情報基盤の一つの拠点となっている^[17]。この社会科学図書館が構築している検索システムは、自館の蔵書だけでなく、上記の研究機関の蔵書も含むもので、書誌ごとに所在が表示される。

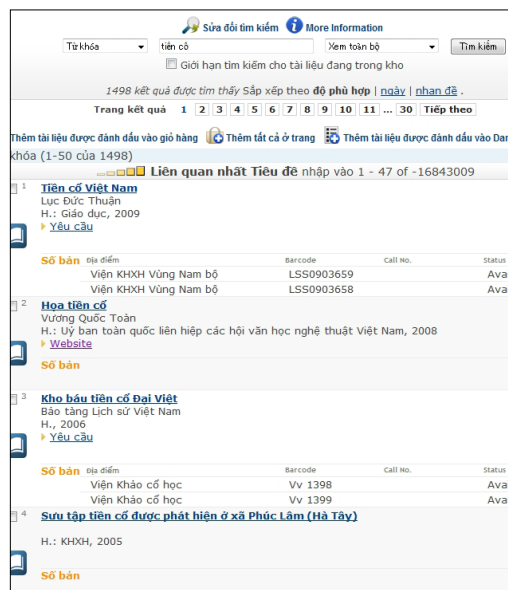


図4 社会科学図書館蔵書検索結果

この検索システムでも、国家図書館と同様に図書と雑誌記事をまとめて検索することができる。図4をみると、検索結果の1と3は図書であり、各コピー（Số bản）の所在場所（Địa điểm）が示されている。1は南部地方社会科学院に、3は考古学院に所蔵があることがわかる。これに対し、2と4は雑誌記事であるためこうした情報は空白になっている。この場合、所在を知るためには掲載誌を改めて検索する必要がある。

✳ ベトナム学術雑誌オンライン^[18] ベトナム科学技術翰林院雑誌オンライン^[19]

VASSのもとにある研究機関はそれぞれ学術雑誌を刊行しており、研究者は研究を進めるにあた

り、まずはこれらにアクセスする必要がある^[20]。こうした雑誌のデジタル化・オンライン化については、漢喃研究院のように「Tập chí Hán Nôm」の掲載論文全文が掲げられているところもあるが^[21]、これは例外的であって、ほとんどの研究院は、ウェブサイト上に近年刊行分の目次を掲載する程度である。従って、雑誌論文にアクセスするためには、掲載誌の現物の所在を確認して、その在処へ直接赴くという昔ながらの方法がまず基本となる。

雑誌刊行の主体である研究院レベルではこのような状況であるが、これとは別に、学術論文のデジタル化・オンライン化を進めようという動きがいくつか存在する。

そうした動きの中でも特に注目されるのが、ベトナム学術雑誌オンライン（以下VJOL）で、これは、International Network for the Availability of Scientific Publications (INASP) の数あるプログラムの一つとして、科学技術省ベトナム科学技術情報局^[22]が中心になって進めている、文字通り、ベトナムの学術雑誌記事のデジタル化・オンライン化事業である（図5）。現時点で、ここに含まれる

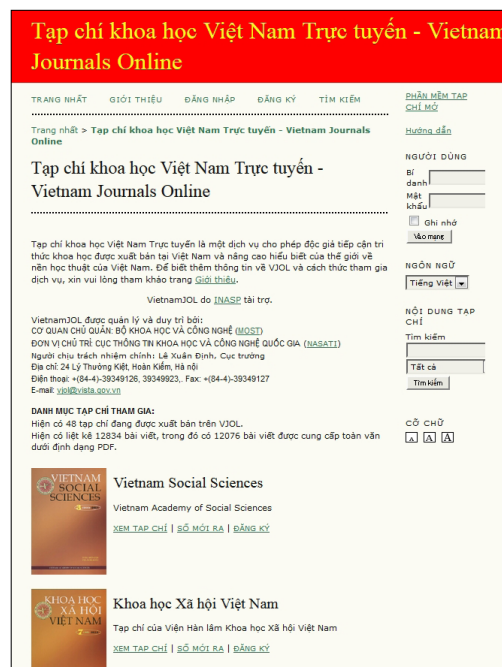


図5 ベトナム学術雑誌オンライン

のは、人文・社会科学系だけでなく理系分野も含む 54 誌で、トップページからタイトルやキーワードによる検索が可能である。

この 54 誌のうち、VASS の学術雑誌は 14 誌を数え、各研究院においてデジタル化・オンライン化が進んでない部分を、ある程度は補うことができる。ただし、このなかには、誌名は掲げられているものの、まったくコンテンツの入っていないもの、記事のタイトルのみ掲げられているもの、アブストラクトのみのもの等、その作業の進捗は各雑誌によってまちまちであり、また、カバーする時期については概ね 2000 年以降であるが、中途に欠号のあるものも少なくない。このように、このデータベースは網羅的なものではないが、総合目録が存在しないという現状からすると、この検索サイトは文献を探索するにあたって、最初の手がかりになると思われる。

このデータベースに含まれる学術雑誌のうち、VASS 関連以外では、大学関係が 16 タイトル、政府や省庁関係が 12 タイトル含まれているが、情報基盤という意味で特に注目されるのは、ベトナム科学技術翰林院（以下 VAST）^[23] が刊行する学術雑誌である。VAST は、VASS と並ぶ政府直属の研究機関で、理系の領域をカバーする。VASS と同じく、下部組織として、各領域の研究機関が存在するが、学術雑誌の多くは、各機関ではなく VAST が直接刊行しており、ウェブサイトでは 11 誌が紹介されている^[24]。このうち、VJOL には 6 誌が採録されており、他と同様に網羅的ではないものの、検索や全文の閲覧が可能になっている。更に、VAST の学術雑誌の場合には、この VJOL 以外にも、全文データベースとしてベトナム科学技術翰林院雑誌オンラインが存在する（図 6）。

ここには VAST 刊行の 11 誌すべてが含まれているが、VJOL と同様、目次だけのものや、欠号のあるタイトルもあり、網羅的なデータベースとは言いがたい。また、この二つのデータベースを見比べると、重複している部分も少なからず存在し^[25]、情報基盤の整備という意味では、まだまだ改良の余地がありそうである。



図 6 ベトナム科学技術翰林院雑誌オンライン

以上、紹介したデータベースについては、完成したものというよりは、現在構築途中中と考えたほうがよさそうである。ただし、そうした現状を理解した上であれば、それなりに利用価値はあるものと思われる。

本稿では紹介できなかったが、これ以外にも、科学技術省ベトナム科学技術情報局^[26]のデータベースなど、情報基盤整備の上で見逃せない事業は少なくない。これらについても、機会を改めて紹介出来ればと考えている。

（本稿は平成 26 年度科学研究費補助金基盤研究 C「インドシナ 3 国逐次刊行物データベースによる機関横断型ネットワーク構築の研究」（課題番号 26330382 代表・大野美紀子）による研究成果の一部である。）

注

- [1] 2014 年 8 月 21 日に大野美紀子が実施したベトナム社会科学通信図書館 Thư viện Khoa học xã hội Thông tin のヒアリングによると、現在のところ総合目録を構築する計画も特にないという。これは、ベトナムにおいて図書館を管轄する省庁が複数あるという事情も、ひとつの要因になっているようである。以上は大野氏からの私信による。

- [2] Thư viện Quốc gia Việt Nam (National Library of Vietnam) <http://nlv.gov.vn/> (参照 2014-9-8 以下同)
- [3] <http://103.23.144.229/opac/>
- [4] <http://dl.nlv.gov.vn/>
- [5] この図は、本誌面掲載用に実際のウェブページをやや加工してある。
- [6] Thư viện Luận án
- [7] Thư viện Báo chí
- [8] Thư viện Sách [Đông Dương]
- [9] Thư viện sách Hán-Nôm
- [10] 「ある一国で刊行されたすべての出版物を網羅的、包括的に収録した書誌」(『図書館情報学用語辞典 第4版』丸善、2013年、132頁)
- [11] <http://nlv.gov.vn/thu-muc-quoc-gia-thang/>
- [12] 注1 ヒアリングによる。
- [13] Thư viện Khoa học xã hội http://opac.issl.vass.gov.vn/*vie
- [14] Viện Hàn lâm Khoa học xã hội Việt Nam (Vietnam Academy of Social Sciences = VASS) <http://www.vass.gov.vn/Pages/Index.aspx> この機関は頻繁に改称しており、現在の名称は2012年12月以降のもの。その前は Viện Khoa học xã hội Việt Nam であり、日本語に訳すならば「ベトナム社会科学院」となる。現名称はこれに「翰林」を加えたもので、意味的には「ベトナム社会科学院」のままでよいが、本誌読者の多くが中国研究者であることに鑑み、ここでは敢えて「翰林」を加えて示してみた。
- [15] このほかに博物館や出版社といった非研究部門5機関が含まれる。
- [16] Viện Thông tin Khoa học xã hội
- [17] 注13 所引|ウェブサイトの紹介によると、図書 347,828 冊、雑誌 2,322 タイトルを数える。
- [18] Tạp chí khoa học Việt Nam Trực tuyến (Vietnam Journals Online = VJOL) <http://www.vjol.info/>
- [19] Vietnam Academy of Science and Technology Journals Online <http://vjs.ac.vn/>
- [20] 具体的な誌名は VASS のウェブサイトを確認できる。 http://www.vass.gov.vn/noidung/anpham/Lists/ThamKhaoTapChi/View_Detail.aspx?Page=1, http://www.vass.gov.vn/noidung/anpham/Lists/TapChi/View_Detail.aspx リストはいくつかあり、それぞれ数え方が異なるが、2014年8月現在で筆者が確認した限りでは、英語誌も含めて37タイトルである。
- [21] Viện Nghiên cứu Hán Nôm <http://www.hannom.org.vn/default.asp?CatID=7>
- [22] Bộ Khoa học và Công nghệ. Cục Thông tin Khoa học và Công nghệ Việt Nam (National Agency for Science and Technology Information = NASATI)
- [23] Viện Hàn lâm Khoa học và Công nghệ Việt Nam (Vietnam Academy of Science and Technology = VAST)
- [24] <http://www.vast.ac.vn/xuat-ban/tap-chi>
- [25] 例えば Tạp chí Hóa học の47巻6号を見ると、明らかに同一の雑誌記事が別々にデジタル化されている。
- [26] <http://www.vista.gov.vn/Default.aspx>

『人文学と著作権問題』余話

漢字文献情報処理研究会

はじめに

漢字文献情報処理研究会では、今年の2月に『人文学と著作権問題』を刊行しました。この本は、これまで当会が10年余りにわたって法律の専門家とともに考えてきた、人文学研究と知的財産権



漢字文献情報処理研究会編 好文出版
2014年2月28日刊 ¥2,500+税
ISBN 978-4-87220-177-2

の間に介在する様々な問題についての成果をまとめたものです。

タイトルには「著作権問題」と銘打っていますが、実際の内容は、アカデミズムにおけるコンプライアンスというものを正面に据えた議論となるように工夫しました。こう書くと、なぜ本のタイトルが「人文学とコンプライアンス」ではなく「人文学と著作権問題」なのかという疑問が出てくるかと思います。これは、人文学と法律の接点を考えた場合、やはり最も関わってくるのは著作権法をおいてほかに無いという考えからです。

このような観点から書かれた本書は、一問一答型の著作権マニュアルではなく、あくまで著作権法をはじめとする法というものの根底にある考え方をベースに学んでいこうとする作りになっています。いわば議論の書として世に問うてみたということになるでしょうか。

刊行から半年を過ぎ、人文学や法学の研究者はもちろんのこと、弁護士などの実務家や、図書館司書、博物館学芸員の方々などからも肯定的、批判的両面から大きな反響がありました。特に桃山学院大学の山本順一氏には、『専門図書館』No.266(2014.7)に書評の労をお執り頂き、いくつかの批判・提言をいただきました。

そこで、本稿では、刊行後に各方面からいただいた批判に応えるとともに、本書で書き尽くせなかったことについて、法條先生と菅原先生に自由闊達に議論してもらうような場を設定しました。

山本氏の書評に応える

菅原 法條さんとの議論が本になったことで、様々な反響があり、少し驚いているところです。

法條 基本的にいろいろな視点からこの本を読んでももらえること、そしてコメントを頂戴できるのは嬉しいですね。
もちろん批判的なものもありますが、こうした視点の違いによって、自らの立場が鮮明になるということがしばしばあります。ありがたいことです。

菅原 山本順一さんに書評を書いていただきました。これまでの本にない斬新な部分を評価してもらいつつも、手厳しい論評もいただきました。ここまで、読み込んでいただけたことは本当にありがたいことです。

法條 私の書評のイメージというのは、執筆者や出版社がお願いして、評者が持ち上げるというものだったのですが、割と辛口で、書評がきちんと役割を果たしていると感じました。

評者である山本さんも自らの視点を明確に意義付けてから、議論を進めていますよね。

菅原 山本さんの視点は、大学における研究活動に積極的に寄与する大学図書館の機能や役割を考える立場ということになるでしょうか。

この点、さすが図書館情報学の研究者というべきでしょう。

菅原 この本で提唱した「アカデミズムにおけるコンプライアンス」を山本さんは「世間から指弾されないよう安全に行きましょうね」という穏当な方向を説くものと評されましたが、これについてはどうでしょうか。

法條 なるほど、そういう読み方もできるのかなという感想を持ちました。

ただ、現代においてこれほどまでにコンプライアンスが重視されるのはなぜでしょうか。企業だけではなく、大学をはじめとする研究機関・研究者においても自らの問題としてコンプライアンスというものを考えてもらいたいというのが本書のメッセージなのです。

社会一般から隔絶された「ムラの論理」は通用しません。自己完結的な研究ならいざ知らず、何らかのかたちで社会と接点を有する研究であれば、社会の常識とうまく「折り合う」必要があるのです。もちろん、社会とうまく「折り合う」ためには、山本さんのいうように「人文学研究者の常識を世間に分かってもらう」ことも大切です。だから、コンプライアンスには、「説明責任」という要素が当然に伴うわけです。

菅原 この本は、「人文学の固有の価値に対する信念がいささか希薄」であるとの指摘も受けました。

この本は対談方式をとっていますから、その発言の半分を法学者の発言が占めています。本の中で人文学者はどちらかという受け身になっていますから、そういった指摘は当然といえば当然かもしれません。

法條 ただ、それはそれとして、この点は本書の価値がどの点にあるのかという問題と関わってきますよね。

この本は、人文学者に法的思考というものの勘どころを伝え、そのエッセンスを理解してもらうことに眼目があったはずです。人文学者に自らの学問分野の価値を高らかに謳ってもらうことが目的ではありません。人文学者が直面している問題に、人文学者が培ってきた学問研究の知恵と、法学者がこれまで紛争解決を通じて積み上げてきた法的思考という知恵とを持ち寄って、ともに考えていこうというのがこの本の主題であり価値だと思っています。

『人文学と著作権問題』余話

菅原 だからこそ、この本の英文タイトルにわざわざ“Humanities meets Jurisprudence”とつけたのですよね。

法條 その通りです。どちらかをどちらかに近づけ、適合させるというではありません。出会いであり、知の持ち寄りだと考えたわけです。人文学も法学も、自然科学とは異なる次元における、人間の経験知・実践知の統合による課題解決の成果です。山本さんは人文学固有の価値に対する信念が希薄と述べていましたが、本書で取り組んでいる作業それ自体が人文学的な価値の体現なのではないでしょうか。

菅原 山本さんはこの本について、「ステークホルダーを説得し、人文学研究者の著作権処理に割く時間と労力をできるだけ軽減するような論理を練り上げるべきだった……（相対契約への過剰な依存は不可解）」と述べています。

法條 人文学研究者の著作権処理の軽減は、山本さんの立場、つまり、大学における研究活動に積極的に寄与する大学図書館の機能や役割を考える立場から導きだされるものなのでしょう。

これに対してこの本の立場は、人文学者に法的思考の作法を伝え、その作法にある程度精通することが、ひいては著作権問題への適確な対応につながり、その方が、結果的に時間や労力の軽減をもたらすというものです。

山本さんは「相対契約への過剰な依存は不可解」と言われますが、ある問題に対応するルールがない中で、ある専門家あるいは専門家集団による部分社会的なかたちで行われるロジックの構築は、この本で最初の問題として指摘した新たな「ムラの論理」をつくりかねません。問題解決は数々のステークホルダー間の合意を通じての秩序形成が基本であるというのが、この本の考え方の基底になっています。

菅原 知人の法学者や弁護士数人に「法学を学ぶ

にあたって最も基礎となる法律は何か？」と尋ねたら、みな民法と答えました。

この意味で、民法の考え方をまず基本におくというのは、法学者として普通なわけですよ。ですから、民法から考え方をスタートさせて、合意の無い部分は民法に立ち戻るというスタンスをとる法條さんは、なんて真っ当な（失礼）法学者だと思ったりしました。

法條 それはどうも（笑）。

菅原 ところが、山本さんは「法條先生は、立法論に踏み込まず、実定法解釈と（保守的な）判例の解釈にとどまるべしという古いタイプの法学研究者のよう」だと認識しています。

法條 確かに、この本ではいわゆる立法論にはあまり踏み込んでいません。しかし、従来の規範を新たな事実当てはめ、新しい法的現実に向き合うのも、新たな法への構想であって、その意味で立法論だと言えるのではないのでしょうか。

菅原 この本では、これまで法的に考えられたことのなかった事例も多く採りあげていますよね。法條さんが答えに窮するような場面もありますから。

法條 もちろん、そうです。できる限り従来の法規範の解釈・適用に努めてはいますが、素材となっているのはいずれもこれまで法が向き合ってこなかった新たな事実なわけですよ。

また、法的思考についての説明を一つの大きな目的としている本書の性質上、法律学のオーソドックスなやり方を踏襲していることは事実なので、山本さんのような読み方をする人がいてもそれは仕方がないですね。

でも、もしかすると山本さんの方こそ、法学者を法解釈だけの人というステレオタイプなイメージに捕われているのかもしれない。すでに述べたように、法解釈だけに完結しない法律学は、新たな事実直面す

ることで、新しい法を生み出すのです。もし、私が真に古いタイプの法学研究者であつたならば、そもそも裁判所の判決にも現れていない人文学という学術研究活動それ自体に、きっと関心などもたなかったでしょう。

様々な読者の声から

菅原 法律家以外は、法律問題というのは、きちんとした割り切れる回答が出るように錯覚しがちです。

この点、読者の方々がこの本に違和感を抱くようで、そういった声がまま聞こえてきます。

法條 この本は性格的には法学入門といえると思います。日常の研究活動と関連付けて、自分の問題として法律を捉えてもらおうということがねらいの一つでした。そういう意味では法律に最初触れてもらう人に対して法律自体が、ある一つの価値観に縛られた一つの読み方しかできないというのは、勧められないと考えました。

ではどうするか、やはりそれを素材に皆さんに考えてもらいましょう、という部分にこだわったつもりです。

多分それぞれの業界には、固まったルールになっているところがあって、そこから見ると、逸脱が甚だしいというふうに読まれてしまうところもあるのだと思います。

菅原 これだけ話題になってくれると、それはそれでわれわれの目的は達成できているのかもしれないですね。

自分の頭で考え出そうとしなければ違和感は抱きませんから。

法條 法学の教科書というのは、一般論を書くと当たり障りがなくてよいのかもしれませんが、こういうふうに特定の分野や業界に特化した形で議論をすると、多くの反応が見られるのが面白いですね。

菅原 反応は望ましいことなのですが、結構ヒス

テリックな論調の方もおられて、驚きました。

法條 法律学も学問ですから、ある程度思考の自由が許されていると思うのですが、残念です。

菅原 司法判断と乖離することが書かれているのならまだしも、学術としての思考の自由の範囲で論じたにすぎず、一つのたたき台として次に生かせればという思いの方が強かったのですが。

法條 逆に単なる読み物として接した方からは、一気に読めて楽しめましたという感想ももらっています。

立法論にまで踏み込んではいないかもしれませんが、一定の法的な作法、法律の作法にのっとって見ると、このような事実として行われている活動をどういう風に理解するかが重要なのではないのでしょうか。ここでは答えが一つに決まるわけではなく、既にそういう事件をやってらっしゃる方からすると、違和感を抱く部分もあると思います。でも法律家としては当然そういったことを想定しているのです。

同じ解釈しか出てこないのであれば、そもそも事件や紛争は起こらず、裁判も必要なくなりますね。だから、そこは自由に議論をしながら、一緒に考えていかねばならないのではないのでしょうか。

菅原 歴史学者は、一部の歴史小説家から結論が曖昧だとよく批判されます。複数の可能性があるというような結論になったり、根拠の無いことには明確な結論を出さなかったりしますから。法学も学問として同じ雰囲気醸し出していますよね。

法條 逆に答えが一つにしか決まらないような学問というのは、学問としては二流だと思いませんか。資格試験で解答が一義的に決まっているようなものというのは、ある意味一流の学問ではないと思います。それは法律学が低い位置にあるということではなく、それを分かった上で謙虚に仕事をして

『人文学と著作権問題』 余話

いきましようということです。そうでなければ、法律家はともすれば権威主義的になってしまいますから。

菅原 どの分野でも一つの答えを追い求める人というのは常にいますからね。しかし健全な学問であれば、唯一無二の解答などそうそうないものなのですか。

教育と著作権余話

法條 この本では教育に関してもかなり多く紙幅を割いていますが、そちらの方面から反応はありましたか。

菅原 教育関係の研究会でもいろいろ議論がありました。図書館の所蔵資料は、教育する立場から考えると利用できなければ意味はないのですが、著作権が気になり積極的に利用しにくいと感じている人が多いようです。個々の教員が法律的な考え方を身につけて判断できるようになれば、図書館の資料に限らずさまざまなコンテンツを積極的に利用できるでしょうから、それが一番よいわけです。今はともすると著作権を固定的に考えてしまい、「使える」コンテンツがあるのに萎縮して手が出せず、なかなか有効に活用されにくいという問題もあります。法律の条文に書いてあることはとても大切なことですし、いろいろな議論が積み重なって形になったものだと思うのですが、社会的なニーズに合わせて合意形成をしてゆくほうがむしろ大切だと感じます。

ただ、一般的な本で勉強すると、条文にどう書かれているかからスタートして、「できる」・「できない」の二者択一のように考えてしまいます。そういう人にはこの本はやはり衝撃的だったようです。

法條 なるほど。

菅原 ある教育系の事務の方からは、法律の条文自体は知らねばならないことですが、そのもっとも根源にある条文の書かれる前の段階の話がこの本に載っていたので、これはみ

なで回し読みしなくてはならないと思った、という感想を頂きました。

法條 回し読みするなら、買ってくださいよ（笑）。

菅原 確かにそうですね（笑）。

要するに、細かい条文については、必要な時に専門の本を読めばよいので、そういう本には載っていない、根源的なスタート地点の話をするというのがこの本の目的なのですね。

そこを理解してもらえない人には、なんとなく逸脱しているように見えたり、都合よく見えたり、あるいは研究者の立場だけを主張しているように見えたりするのかもしれない。

使う側がいるから法律があるという根本的な問題が見えていない人が多いような気がします。

法條 一つの法律だけで一つの秩序が成り立っているわけではないとか、少し考えれば分かることではあるのですが、なかなか理解するのは難しいのです。菅原さんも、著作権問題は著作権法だけに留まるものではないということが分かるまで、それなりに私と時間をもったではありませんか。

だから、この本を初めて読む人にとっては、衝撃というほどのものかどうか分かりませんが、びっくりはしたでしょうね。

研究倫理との関係から

菅原 この本には、著作権の問題だと思い込んでいたことが研究倫理の問題だったという用例を載せています。たまたま STAP 細胞事件があり、偶然にトレンドを踏まえた内容になってしまいました。この点、何か反応はありましたか。

法條 とりたててなかったように思います。

菅原 STAP 細胞の問題では、小保方さんの発言は重要なことから形式はどうでもよいというような議論を目にしました。つまり、法律の条文を掲げて違反を指摘した人に、

STAP 細胞の重要さを考えれば、法律の方が間違っているのだと言わんばかりでした。研究ということになると、このあたりのバランスは難しいと思いました。

私は、法律は過去の議論の積み重ねからできていて、その議論は時に現実に即して上書きはされるけれども、過去の議論を完全に無視することはできないと理解しています。法律の個々の条文はそういう過去との接点として存在するわけなのでしょうが、小保方擁護の議論の中には、それらが全て捨象されているものもあると思いました。やはり、人文学研究者としてそういったことは、この本でももう少し述べておくべきだったのかもしれない。

法條 自然科学という学問分野は、論文の面に関しては非常に制度化されており、論証の方法や形式がカチッと決まっている分野であると思います。一方で、人文学はそこまでタイトな感じはしません。だから、「人文学と研究倫理」といった議論をするならば、他分野の制度化の状況を見るべきなのではないでしょうか。経済学などは、割とカチッとしているように見えます。

菅原 自然科学系では論文の文体がかなり定形化されていて、一定のフォーマットに名詞・数値を入れ替えただけで論文になる分野もあり、それが「コピペ」と呼ばれているようです。そのため、人文系と議論がかみ合わないこともあったようです。

法條 それにしても、自然科学系のようにディシプリンがきちんとしていても、真実にまで至るとは限らないことが STAP 細胞問題で明らかになったわけですから、この問題は簡単に言い尽くせません。

菅原 STAP 細胞問題では博士論文序論のコピペも騒がれましたが、コピー元がアメリカの公的機関の Web ページで、著作権保護の対象外だといいます。この事件は、法律問題として論じるのがふさわしいのでしょうか。

法條 今回の博士論文のコピペ問題に関して言う

と、法律問題ではないのではないかと思います。一方で、博士号の授与というのは法律問題と言えなくはありませんが、その授与の判断に関しては、やはり法律問題というよりは大学の自治の問題ということになるでしょう。

任意団体が行う契約の承継

菅原 この本では、法律に書いてないことや判例のないことは、原則、民法の契約の考え方にのっとるべしという方針を貫いています。その中で、契約の対象となるのは自然人か法人でなければならないと教えられました。では任意団体、具体的には法人格を持たない研究会などが著作権許諾を得たい場合、契約をどのように行い、承継はどうなるのでしょうか。

法條 任意団体の場合の契約は、研究会管理ではなく、個人の形で行うべきだと思います。契約者は、その契約の提案者になって構わないのですが、代表者にまとめるのが普通でしょう。その場合、承継も次の代表者になされると考えてよいと思います。

菅原 では、「代表者が交代した場合には、契約が次の代表者に承継される」というような一文を契約書に入れておいた方がよいのでしょうか。

法條 それはあたり前のことなので、そこまでする必要はありません。

菅原 書くまでもないことですか。

法條 契約というのは、当たり前のことを確認するものではなく、当たり前ではない扱いについて合意をとるものなのです。

菅原 よく、わかりました。

まとめると、法人格を持たない任意団体は、団体名による契約が法律上できないので、代表者という個人の名前で契約をする。そうであれば、代表者が交代した場合、まず常識的に次の代表者に契約が承継されることは疑いない。ということですね。

『人文学と著作権問題』余話

法條 そうなります。これは契約という法律問題でもあるけれど、著作物に関する契約であるので、著作権問題でもあります。そういった法律の重なり合う部分が重要だということを感じてもらえればありがたいです。

平成 26 年著作権法改正との関係

菅原 先般、著作権法が改正されましたが、改正案をみて驚きました。今回の改正では電子書籍に対する出版権について、紙の出版とは別の権利が設定されました。この本の中で、法條さんが出版契約を結んでいても、それを電子化するならば別の契約であろうとおっしゃっています。いみじくも今、著作権法の改正があって、法條さんの考え方が一致したのです。

法條 いや、そういうものでしょう。なぜなら、まだ技術的に電子出版というものが想定されていないときに結ばれた契約においては、当事者間で電子出版が合意されているとすることはできませんから。

菅原 山本さんは書評の中で、「相対契約への過剰な依存は不可解」と述べておられますが、このように考えてくると、著作権と契約は切っても切り離せないと思うのですが。

法條 さまざまな利用が想定され得る現在では、著作権法で定められている利用権としての著作権は、非常に限られた範囲のものだといえます。利害関係者ときちんと調整がついたものだけを立法化しているところがあって、それ以外の部分というのはおおよそ契約の範疇で議論がなされます。したがって、多くの場合、相対の契約にならざるを得ないのが実情です。そのあたりを山本先生はどのように考えておられるのかお聞きしたいですね。

法律的な考え方 再論

法條 著作権法の議論をしているのに、ひんばん

に民法に立ち返るところは、今回、この本を初めて読む人にとっては、違和感を抱いたかもしれません。

しかし、ある程度、著作権法の勉強をすれば、著作権法の民法へのフィードバックはいわば当然の前提です。ある事象にかかわる法律は、決して一つではありません。複数の法律が、いろいろな側面から関わってきます。ところが、分野が変わると別々に考える人は法律家の中にもたくさんいて、個別の専門化された法同士の関連付けは意外とできていないのです。

だから、一般の方々が著作権法だけに興味を持って、それだけに注意を向けるというのは、日本人の傾向としては非常によく分かります。それは訓練を受けた法律家さえそうなりがちなものです。

菅原 人文学もあらゆる分野で専門の細分化が進んでおり、内に閉じこもろうという研究者が多いように感じます。法條さんが言われるような日本人の傾向というのは、確かにあるのかもしれませんが。

法條 どのような研究でも、正しい判断をするためには、できるだけ多くの隣接分野・関連分野を見て判断するというのが常とう手段であって、一つの分野とか、あるいはある文書だけを見て判断するということは、通常しないだろうと思います。これは実生活でも変わらないのではないのでしょうか。

菅原 今回の法條さんとの対談を通じて、またこの本を読んだ方からの様々なリアクションを見て、法律の条文というものに対する感性の違いというものを強く感じました。繰り返しになるかもしれませんが、条文はある種の合意が取れた狭い範囲のもので、条文に書かれていないところは民法的な契約で済まざるを得ない、そういう考え方が全く違ってくのだと、考えさせられた次第です。ありがとうございました。

外国語教育ワークショップ「プチ IT を利用した外国語教育の実践」 報告

PC1 台からできる外国語教育における IT 利用

氷野 善寛（ひの よしひろ）・紅粉 芳恵（べにこ よしえ）

回 ワークショップ概要^[1]

2014 年 7 月 26 日（土）午後、関西大学千里山キャンパスにて外国語教育ワークショップ「プチ IT を利用した外国語教育の実践」と題したワークショップを開催した。このワークショップの趣旨は、e-Learning、LMS、アクティブラーニングといった多種多様な外国語教育における IT 利用がある中で、iPad、iPhone またはコンピュータ 1 台だけが利用可能、あるいは授業の準備段階でのみコンピュータを利用するという環境下における IT の利用方法について考えてみるというものである。また授業で利用する教育資源をいかに蓄積し共有するか、一つの素材をいかに多角的に利用するかといった視点について具体的な実践を通して、活用方法を考える場とするために企画したものである。ワークショップの式次第は最初に内田慶市氏（関西大学）による講演を行い、続いて氷野善寛（関西大学）、清原文代（大阪府立大学）、神谷健一（大阪工業大学）、中西千香（愛知県立大学）、紅粉芳恵（京都産業大学）、干野真一（新潟大学）、樋口拓弥（関西大学卒業生）各氏による報告、ワークショップを実施した。当日は 5 時間にも及ぶ長丁場にも

かわらず、100 名を超える参加があった。その内訳は大学関係者 61%・高校教員 12%・その他の教育機関の教員 4%・出版社 5%・学生、一般、その他 18% であった。また教員の教授言語については報告内容との関係から中国語が多かったものの、英語・ドイツ語・日本語教育関係者も 1 割以上の参加があった。このことから、この分野において大学や高校を中心とした教員の言語を超えた関心度の高さを見て取ることができた。以下に個々の発表の概要を紹介していく。

回 外国語教育に ICT 利用は有効か？

（内田慶市）

まず、内田氏が「外国語教育に ICT 利用は有効か？」と題した講演を行った。この講演では、中国語教育では古くから教育や研究においてコンピュータの利用がなされてきた点について振り返り、文字コードの話から文書解析まで幅広い話題を展開した。

次に教育についての話題に移り、中国語教育に限らず、外国語教育では「iTunesU」や「Podcast」、「反転授業」、「遠隔授業」といった様々な形態、方法

が用いられていることを指摘した。特に、最近では「反転授業」という授業形態が声高に叫ばれており、教育の形を変えるという試みがなされているが、氏は教育の形を変えることで教育の質を変えられるのかという問題が存在することを指摘し、「反転授業」は教育において非常に有効的な手段ではあるが、「反転授業」発祥の地であるアメリカの教育と日本の教育とでは根本的な授業スタイルの違いがある点について吟味しなければならないと指摘した。つまり、アメリカの大学では学生は授業の合間に分厚い本を読み、「授業内」でのディベート等に備えておかなければならず、事前に学習すべき内容を「授業外」で学習し、授業に参加する段階では内容を理解していることが当然のように求められているという事実がある。こういった土壌があることを前提として「反転授業」を進めているのがアメリカであり、この点についてよく理解した上で「反転授業」のメリット・デメリットを考え、導入すべきだとした。また関西大学におけるタブレット端末の所有率など具体的な数値を示しながら、「反転授業」や ICT を利用した授業を進める上で、学生に対して同じ環境を保証できるかどうかという点についても考慮しなければならないと指摘した。さらに ICT の利用については耳・口を利用した学習には有効であるが、一方で辞書を引く、文字を書くといった手を動かして覚える動作というものも外国語学習には必要であるとした。ICT はあくまで外国語学習の 1 つのツールであり、その開発等が「自己目的化」してはならないと結論付けた。

Web を利用した中国語教育——サジェスト型中国語オンライン辞書とコンテンツデータベース (氷野善寛)

氷野は中国語教育と学習に関する様々な情報をまとめたウェブページを作ることができないかを考えて複数の方向からアプローチをしている。本ワークショップでは現在開発中の 3 つのシステムについて開発コンセプトや概要について紹介した。

1 つめは「みんなの教材」のように教員が一緒

になって、学習教材、素材、教案などをシェアするサイトの構築である。写真、音声、動画といったデータを集める「コンテンツデータベース」と中国・中国語に関する様々な情報を発信する「中国語学習ジャーナル」の 2 つの方向からアプローチをしている。特に後者では「みんなで集めよう印刷できる学習素材」という企画を立て、印刷できる学習素材を中心に、ワンコンテンツマルチユースをコンセプトにした学習素材の収集及び公開を行っている。2 つめは共同編集を前提として開発中の「中国語辞書」についての紹介を行った。学習者の利用を想定している辞書のため、二次検索や画像検索といった検索機能やサジェスト編集機能を重視したほか、上述のコンテンツデータベースやテキストの例文を収録した例文データベースから情報を得ることができるような設計を行っている。3 つめは「形態素解析システム」について紹介した。この「形態素解析システム」は文を単純に分割するだけではなく、分割した単語を HSK などの語彙リストと比較してどのような単語が使われているかを分析し、文中で使用されている単語の難易度を確認したり、分割した単語から直接辞書データベースにジャンプしたりする機能を持たせている。これらシステムについては今後半年から一年程度の検証期間を経て順次公開する予定である。

音声付き教材を作る—— Quizlet、EPUB、ShowMe などを例に——

(清原文代)

清原氏は Quizlet、EPUB、ShowMe という外国語教育において有用なサービスやシステムについて紹介した。まず 3 つのサービスについて概要を解説した後、参加者が詳細を希望するものについて順番に具体的な操作手順を示しながら解説した。3 つのサービスについては以下の通りである。まず Quizlet は音声付きの単語カードセットを作成できる Web サービスで、音声付きカードのほか、音声を聞き取って入力する練習、マッチングゲーム、自動生成のテストなどの機能があ

る。既存のカードセットを使用するだけならアカウントは不要である。カードセットを作成するにはアカウント（無料）が必要で、カードの両面に表示する単語と訳語を入力するだけで上記の機能を持つカードセットを作成できる。iPhone や iPad、Android 向けの専用アプリがあるだけでなく、ブラウザベースでも利用でき、音声も中国語 TTS に対応している。次に、EPUB は電子書籍の規格で、Google、Apple 等が採用し、実質的に電子書籍の世界標準と言えるものである。2011 年に策定された EPUB3 では、音声や動画を内包することができ、更に音声の再生に同期してテキストがハイライト表示されるメディアオーバーレイにも対応し、音声付き外国語学習教材を EPUB で作成することが可能になった。最後に ShowMe は講義音声と板書が同期した動画を iPad を使って作成することができるサービスである。このサービスは反転授業の動画教材（講義を細かくモジュール化した予習復習用の動画教材）を作るのに適している。なお、作成した動画は Web ブラウザがあればパソコンでもスマートフォンでも視聴できる。

回 Phrase Reading Worksheet

（神谷健一）

読解・精読を中心として扱う授業などで効果的に利用できるプリント教材「フレーズ・リーディング・ワークシート」（Phrase Reading Worksheet）の効率的な作成方法について紹介した。このシートは縦方向に折って使うことで複数の学習方法から学習者に合った方法を選べることができる教材で、神谷氏が開発し無料で公開している教材作成ツールをインストールすることにより任意のテキストから簡単に作成できる。元々は英語教材の作成に利用していたが、今回、新たに中国語教材の作成ツールを開発したこともあり、本ワークショップでは中国語教材の作成手順について具体的に示した。その中で、中国語用ツールについては英語用と比べると機能面では改善の余地があるものの、3つ折りで原文・ピンイン表記・訳文を並べたレイアウトを作成することができる点は中国語教育

の現場でも有効に活用できるのではないかと結論付けた。

この取り組みはパソコン+プリンタという最低限の設備で利用できるものであり、設備的に限定される環境下であっても利用できる点が非常にありがたい。

回 レアリアを利用した授業の具体案

（中西千香）

中西氏は、レアリアを中国語教育に如何に取り入れるかを考え、授業の中で実践するという取り組みをしている。レアリアとは外国語教育の中で用いられる、その国の人が実際の生活の中で目にしている実物資料のことであり、街の看板やスーパーのチラシ、新聞雑誌、さらには HP やアプリまでとその範囲は広い。外国語教育にこれらを取り入れることで外国語学習がより実践的になり、相手の国を理解する上で非常に重要なものとなる。今回は上記のようなレアリアを紹介しながら、教える立場としてどれだけ知り、その上で、実際の授業の場でどのように使用できるかについて実践例を紹介し、参加者と考えてみるという取り組みをした。

多数あるレアリアの中からスーパーでの買い物に限定し、テキストで買い物のシーンを学び終えた後に、さらに中国の商習慣、物価などを知るためのタスクを作った。参加者は、通常の授業と同じように小グループに分かれて、スーパーのチラシを見ながら、ワークシートに沿って、それぞれのタスクを完成させていった。また、参加者にレアリアの特徴、レアリアのむずかしさ、実際の授業に導入する場合の注意点について考えてもらった。

回 YouTube を使った字幕作成と授業デザイン：インタビュー動画の利用

（紅粉芳恵）

中国語の文章を読んで、訳すだけではつまらないということで、講読クラスでインタビュー動画

画に字幕を付けるという授業実践について紹介した。まず素材として使用したのは「中国語学習ジャーナル」で公開している中国人に対するインタビュー動画「采访中国人^[2]」である。この動画は中国語学習者に等身大の中国人を知ってもらおうと企画し撮影した5分～15分程度のインタビューである。中国語字幕は付けてあるが日本語字幕はない。この動画に対して、中国語を日本語に翻訳し、グループによる協同作業で字幕として適切な日本語訳を作成し、CaptionTube という YouTube の動画に字幕を付けることができるアプリを利用して字幕作成作業を行った。字幕を作成するには単なる翻訳とは異なり、読みやすくなるための字数制限があり、一般的に1行に14文字、1秒間に4文字というルールがある。そのためインタビューに適した語彙や言葉遣いを工夫しなければならず、中国語力だけでなく、日本語力も必要となる。学生はグループワークを通じて、中国語翻訳の奥深さについて学ぶことができたと報告した。

中国語 CM を利用した教授法について (千野真一)

千野氏は、中国語 CM を活用した教授法を提案した。授業で CM を利用する利点は、音声と映像があるため印象に残りやすい、生きた言語表現に実際の速度で触れられる、短時間で有意義な気分転換ができるという3点に集約されたと考え、中国語を学ぶ上で言語習得と並んで不可欠な、言語文化への理解の助けになることが可能であるとした。CM はおよそ15～30秒、長いものでも1分程度であるが、短時間の中に出てくる表現はウィットに富み、中国語・中国文化を反映した良質なものが数多くある。CM 中には教員側が説明したい文化事象や中国語表現が複数含まれるというものもあるが、情報量が多くなり過ぎないように、紹介するポイントを絞って活用することを提案した。

報告では実際の CM を流しながら具体的にどのような手順で紹介するか、どのような点に触れると良いのか、その活用法について「文化」の紹介、「文法」の復習、成語を楽しむ、短編動画を楽しむといった観点から多角的に紹介した。

Yubiquitous Text & データセットの利用 (樋口拓弥、氷野善寛)

授業中に教室内で CD や DVD プレーヤーの代わりとして使用できる無料の iPad/iPhone アプリ “Yubiquitous Text” の開発者である樋口氏がまず本アプリの開発コンセプトから実際の細かい操作方法までを解説した。“Yubiquitous Text” は普通のプレーヤーにあるような再生・停止といった機能以外に、タッチした部分から本文を自由に再生することができ、復唱練習やシャドーイングに適した再生モード、テレビやプロジェクターなどにビデオやスクリプトを表示する機能など iPad ならではの便利な機能を備えている。続いて氷野が中国語のサンプルデータを用い具体的なデータの作成方法について解説した。さらに、アプリで利用できるデータセットを公開しているサイト「CHLANG 中国語コンテンツ検索システム」を紹介した。このサイトでは公開の許諾がとれた大学や高校向けの中国語の教科書に準拠したデータセットを提供している。

注

- [1] 本報告については当日の資料や発表内容をもとに氷野、紅粉が編集したものである。文章には報告者の資料から引用している箇所もあるが、引用元は特に明示していない。詳細については「外国語教育ワークショップ報告ページ」(<http://www.ch-station.org/ws2014/>)をご覧ください。資料と共に当日の動画も公開している。
- [2] 「采访中国人 Interview in China」<http://www.ch-station.org/movie/caifang/>

レビュー & リソース紹介

ここ数年でPCからスマホ・タブレットに情報端末の主役の座がすっかり移ってしまった観がある。そのためであろう、PC向けの面白いアプリケーションソフトやツールはすっかり数を減らしてしまった。本レビューでも、アプリケーションソフトのレビューは年々減少の一途をたどっている。

それに比して、各種コンテンツやWebサービスには勢いがある。中国学関連についても、大規模データベースのリリースは止まらず、今年はいよいよあの中華書局が『中華経典典籍庫』を発売した。また今号では中国の書籍や各種情報源を参照して、これまでに取り上げてこなかったリソースを洗い出し、瀚堂典藏・河東師範大学のコンテンツ群・農業古籍などをレビューした。これまでに取り上げたことのあるサイトについても、新たなサービスが始まっているものも多く見られ、古典籍関連サイトのレビューにはそうしたサイトが多数取り上げられている。

多言語・多漢字処理環境の一般化を背景にコンテンツの開発が進むのであれば、それは中国学の研究に従事する我々にとって、大いに歓迎すべき状況である。中国学デジタル化の流れは、今後とも広く深く、進展していくのであろう。

Contents

Microsoft Office Mobile & Online	田邊 鉄 …… 90
一太郎2014徹 & ATOK2014	山田 崇仁 …… 93
BabelPad	師 茂樹 …… 97
Windows 8.1環境におけるピンイン処理	千田 大介 … 100
When.exe Ruby版	須賀 隆 … 104
中華経典典籍庫	千田 大介 … 110
瀚堂典藏	千田 大介 … 114
蘇州地方文献データベース	佐藤 仁史 … 118
華東師範大学中国文字研究与应用中心	山田 崇仁 … 120
中国・台湾の古典籍検索	小島 浩之 … 126
「農業古籍」について	村上 陽子 … 136

Microsoft Office Mobile & Online

田邊 鉄

❖ マイクロソフト Office Mobile

この2～3年の間に、タブレット端末は新しいスタイルのコンピュータとして、すっかり定着した。パソコンに匹敵する機能・性能を備えたハードやソフトも現れ、「ノートPCの置き換え」として使える目途が立ったようにも見える。だが、筆者は未だに出張の時「ノートPCを持って行くか、タブレットで済みますか」迷うことが多い。その一番の理由は、PCでないと、マイクロソフト Office のフルバージョンが使えないことである。iOS や Android 端末には多数の「互換 Office」「オルタナ Office」アプリがあるが、文書の互換性に難があることから、マイクロソフト純正の Office ソフトが求められていた。

そんな声に応じて、2月に iPhone と Android スマートフォン向けに、Office Mobile が公開された。iPhone 用の Office Mobile は、iPad でも使える。ただし、画面は iPhone に特化されているので、作業領域が大きく広がることはなく、「iPad 用の Office ソフト」を期待していると失望するだろう。

一方、Android スマートフォン用の Office Mobile は、Android タブレットにはインストールできない。Nexus7 で Google Play にアクセスしてみたところ、「×このアプリはお使いの端末に対応していません」と表示され、ダウンロードすらできない。環境は大差ないはずなので、インストーラさえダウンロードできればインストールできるのではないかと考え、以下の方法でタブレットにインストールを試みたところ、問題なく動作した。

- ① パソコンで Microsoft Office Mobile のサイトに Web でアクセスし、URL をコピーする。2014 年 8 月現在は以下のとおり。
<https://play.google.com/store/apps/details?id=com.microsoft.office.officehub&hl=ja>
- ② APK Downloader - Evozi Official にアクセスし、上でコピーした URL を貼り付ける。Generate

Download Link をクリック。

<http://apps.evozi.com/apk-downloader/>

- ③ Click here to download com.microsoft.office.officehub と書かれたボタンをクリック。
- ④ ダウンロードされた、拡張子 apk のファイルをタブレットに転送する。
- ⑤ 「ファイルマネージャ」等のアプリで、転送したファイルを起動、指示に従ってインストールする。

インストール中に、「ブロックされました」等の表示が出て、インストールできないことがある。この場合は設定から、提供元不明のアプリのダウンロードを許可してやる必要がある。ネットの速度にもよるが、数十秒でインストールは完了する。

- ⑥ 最初の起動時に、Microsoft アカウントの設定を聞かれる。Microsoft アカウントのメールアドレスを入力、ログインすると、OneDrive 上のファイルが一覧表示される (図1)。



図1 ファイルリスト画面

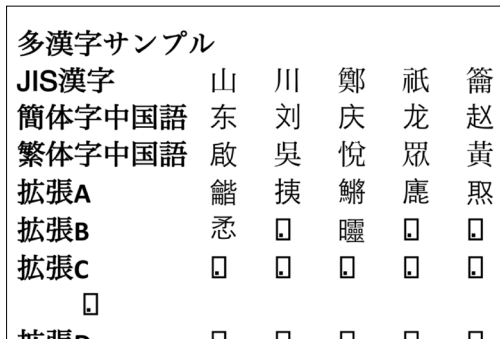


図2 Office Mobile Word の閲覧画面

図2はワードの文書を開いてみたところである。最初は閲覧モードで開かれる。画面上部の鉛筆アイコン(図3)をタップすると、編集モードに移行し、テキストや書式を変更できるようになる。



図3 ツールバーの鉛筆アイコン



図4 書式設定

PC用Officeのようなリボンインタフェースは搭載していないが、編集モードで書式設定アイコンをタップすると、太字・斜体・下線・取り消し線・背景色と文字色・文字サイズを設定できる(図4)。背景色と文字色は赤・黄・緑からの選択になる。また、行間・文字間など細かなレイアウトは設定できず、フォントの変更もできない。

現状、サードパーティから多数出ているAndroid、iOS用互換Officeの方がより高機能であり、高速で操作性も高いので、あえてAndroid端末やiPadに、ス

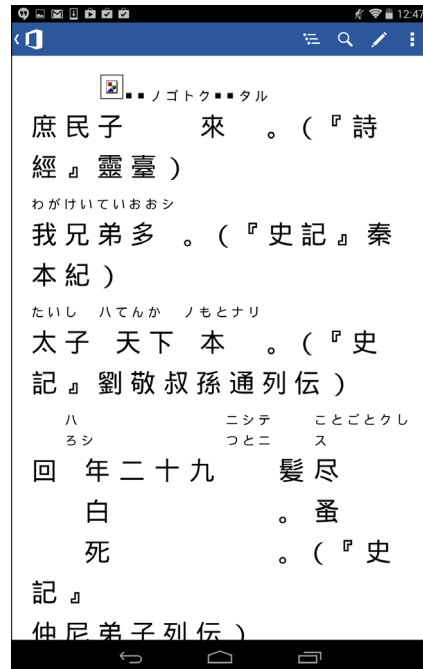


図5 縦書きは非対応

マートフォン向けのOffice Mobileをインストールするメリットはないだろう。iOS上で動くCitrixのQuick Editでは、Power Pointスライドの再現性が、図も含めて非常に高い。Google Quick Officeの「文書」アプリでは、MS Wordの変更履歴やコメント一覧を表示・更新することができる。InfrawareのPOLARIS Officeでは、縦書きにも対応している。これらは従来、Mobile Office・互換Officeの弱点として指摘されていた点であり、純正にこだわる理由はなくなりつつあるとみていいだろう。

❖ Office for iPad

Office for iPadは、iPad上でPC用Office同様のエクスペリエンスを提供するマイクロソフト純正のOfficeソフトである。PC用Officeでおなじみの、リボンインタフェースの見た目はそのままに、ボタンの間隔を変えるなど、タッチ操作に最適化されている。Word・Excel・PowerPointの文書を閲覧および編集することができる。同社のクラウドサービスOffice 365の追加機能として提供され、文書を編集するためにはOffice 365のサブスクリプションが必要である。なお、Office for iPadは現在日本のiTunes Storeでは提供されていない。

い。「アメリカ在住者の Apple ID」と「Office 365 の法人契約」の組み合わせで、日本でも利用できるが、これらの条件を満たしている本誌読者はほとんどおられないと思うので、使い方の詳細などは割愛する。

❖ Office Online

日本で Office for iPad が使えなくても、悲観するには及ばないかも知れない。iPad ではオンライン版の Office が使えるからである。本誌 14 号で紹介した通り Office Web Apps（当時）では、iOS 上の Safari や Chrome といったブラウザを用いて、縦書きや多言語対応を含む多くの編集機能を利用して文書の編集が可能である。ただし、Android 端末では、依然ただのビューアーとしてしか使えない。



図 6 縦書きが再現されている



図 7 多漢字に対応

Office Web Apps は本年 2 月に Office Online、Sky Drive は One Drive と名称を変更した。機能や性能に変更はなく、利用のための手続きも同様である。ただ、かつて「Office Online」と呼ばれていた、マイクロソフトの Office 製品ポータルページから、Office Online へのリンクが貼られ、モバイルやオンライン向けの情報が増えるなど、マイクロソフトの意図が那边にあるかは明らかだろう。

ところで、この Office Online、個人での利用は無償だが、ビジネスでの利用は制限されていることをご存知だろうか。会社の Office 文書を One Drive で家庭のパソコンにも同期させ、帰宅後仕事する、というスタイルは、私たちの社会では当たり前に見られるが、本来は One Drive のビジネス向け有料版を用いなければならない。同期して作業を行う PC も会社で管理されていなければならない。日本では仕事とプライベートの境界が曖昧になりやすく、家族共用のパソコンが時に仕事用、時に宿題用に使われることは珍しくない。無料 Office を仕事に使われたのでは、法人向けサービスは立ちゆかなくなる。

マイクロソフトは年末までに、Office 365 の個人向けサブスクリプション発売を予定している。アメリカ等で既に提供されている家庭・個人向けのオンライン Office が遅れたのは、このような日本独自の事情を勘案し、法人向けの Office 365 が浸透してから個人向けを出したい、という思惑からだろう。一方、今のタイミングで個人向けの発表をする、ということは、新 CEO サティア・ナディア氏は、先代・先々代が「Windows & Office プリインストール PC」によって築いてきた支配的な状況を、いったんチャラにしてでも、モバイルとクラウドへと大胆にシフトしようとする、並々ならぬ決意を持っている、ということだろう。

一太郎2014 徹 & ATOK2014

山田 崇仁

❖ はじめに

本年度の新バージョンのテーマは「徹(てつ)」。「徹」には「貫き通すこと、最後までする、徹底、頑固一徹」の意味が込められており、それは「文字からスタイル、罫線や画像に至るまで、一太郎史上最もこだわりが実現できる。まるで職人のような、意思を持ったこだわりの実現。」の意味が込められている(カギ括弧内は『一太郎 2014 徹バージョンアップのご案内』カタログより引用)。

それらのこだわりについては、一太郎のカタログをご覧くださいとして、本レビューでは文字に関する点を中心に、ATOK2014 共々取り上げることになしたい。

❖ 一太郎 2014 徹

■ 他ソフトへも IVS で貼り付け可能に

今回のバージョンアップの案内では、最も目立つところに「IVS 対応」と書かれていた。その文句を読んだときに、「一太郎 2008 から IVS (Ideographic Variation Sequence / Selector) での異体字(以下 IVS 異体字と表記)に対応していたはずなのでは」と不思議に思ったのである^[1]。

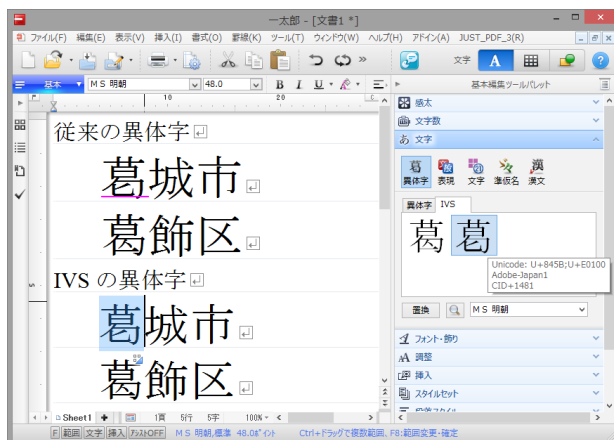


図1 一太郎で異体字を入力・表示(従来の異体字機能で変換した葛の字の下に、アンダーラインが引かれている)

しかしそれは、筆者の勘違いであった。

従来から一太郎では、ある文字を選択した後、「ツールパレット」の「文字」グループにある「異体字」ボタンをクリックして異体字設定画面を表示し、そこで(異体字が存在する場合)異体字を選択することで異体字への変換が可能だったのだが^[2]、筆者はこれを Unicode の IVD (Ideographic Variation Database) に登録された異体字を入力する機能だと勘違いしていたのだ。しかし、これは実際には OpenType フォントの異体字を入力する機能であった。

既に一太郎 2008 から IVD に登録された異体字については、他のアプリケーション経由で貼り付けることができた。それに加えて、本バージョンでは、新機能として IVS 異体字を入力するツールを実装した。

このツールを使用するには、以下のようにすればよい。まず、一太郎の「ツールパレット」から「文字」グループ→「異体字」ボタンの順にクリックする。異体字入力ツールが表示されたら、従来からの「異体字」タブのほかに「IVS」タブが追加されている。一見すると、インターフェイスや使い勝手を含めて両者とも同じようにみえるが、それぞれには機能の大きな違いがある。

それは、「異体字」で変換された異体字をコピーして他のアプリケーションに貼り付けると、異体字ではなく一般的な字体に戻ってしまうのに対し、「IVS」の場合は IVS に対応しているアプリケーションであれば、異体字をそのまま貼り付けることができるという点である。「異体字」タブが残されているのは、おそらく旧バージョンからの互換性を保つためであると思われる(「異体字」タブから異体字変換した場合、IVS 異体字ではないことを明らかにするためか、当該文字の下にアンダーラインが自動的につくようになっていく)。

筆者はこの進化に気がつかなかったために、本バージョンをそれほど評価していなかったのだが、この新機能を知ったため、それを見直すに至った。

Microsoft Office が正式に IVS 異体字の入力・

表示に対応したのは 2013 からだが^[3]、入力するには MS-IME の文字パレットを使う必要がある^[4]。その点でも、一太郎の IVS 対応は特筆すべきといえる。

無論、IVS に対応した他形式に保存した場合も、IVS 異体字の情報はそのまま保持される。筆者の試した限り、Unicode テキスト形式や Word の DOC 形式、OTD 形式などで IVS 異体字のまま保存できた。また、EPUB 形式で保存した場合、従来は「異体字」機能から入力された（Open Type に由来する）異体字については画像に変換して保存されていたのだが、これについても、IVS 異体字の情報を保持したテキストデータとして出力されている。

■ 今年のフォントは二種類追加

筆者は、一太郎プレミアム版を購入するかどうかの基準として、バンドルされるフォントに重きを置いている。本年のバージョンでは、ヒラギノ書体のデザインで知られる「字游工房フォント」が付属している^[5]。フォント名はオリジナルとなるが、その実態は游フォントファミリーと同一だと思われる。游フォントファミリーは、1 ウェイトであっても三万円以上する高価高品質なプロユースフォントだが、今回はそれがなんと 4 書体（明朝・ゴシック・かな・教科書）も提供されている。インストールが 1 台のみ制限されていることを考慮に入れても、これだけでプレミアムバージョンを購入する価値はある。

さらに加えて、今バージョンでは通常版にも IPAmj 明朝フォントが付属している。本誌読者ならよくご存じだろうが、このフォントは独立行政法人情報処理推進機構（IPA）によって配布されているフォントの一つであり、IVD の汎用電子コレクションに対応した異体字を収録している。このフォントが収録されたのは、汎用電子コレクションが我が国の行政機関向けの異体字を主に収録しており、それが未だ根強いユーザー層がいると思われる官公庁向けのアピールポイントになるためだろうが（上記字游工房フォントは、IVD の Adobe-Japan1 コレクションに対応している）、将来的には Adobe-Japan1 と汎用電子の両コレクションに対応した花園明朝を同梱してほしいと、希望するところである。

❖ ATOK2014

「速さと快適さを徹底追求。史上最強の ATOK」が今

バージョンのうたい文句だが、実装された新機能には、それほど目を見張るようなものはない。それはある意味 ATOK の機能が既に充実していることを示しているが、これらの新機能にも既存機能にプラスアルファすることで、より使い勝手をよくする方向が見て取れる。

たとえば「アクセルモード」。これは従前からある「オンメモリ辞書」（メインとなる変換辞書をメモリ上に展開して、ハードディスクからのアクセスよりも高速な変換を実現する機能）を改良したものである。筆者は以前のバージョンから「オンメモリ辞書」を使っていたので、あまりその恩恵を実感できないのだが、パフォーマンスが最高の状態ではメイン辞書以外の辞書もメモリ上に展開してくれるようなので、複数の辞書を使い分けている筆者のような人間にとっては、それなりに便利な機能である。使い方は ATOK メニューから ATOK プロパティを呼び出し、設定画面から「辞書・学習」→タブをクリックして画面を切り替え、左下の「アクセルモード設定」をクリックして画面を呼び出し、適切な設定をすればよい。

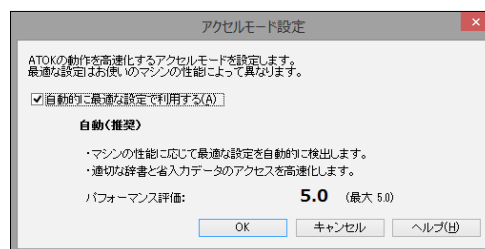


図2 ATOKのアクセルモード設定画面

それ以外にも ATOK2014 では、入力者のストレスを軽減することを目的とした機能がいくつも追加されている。いくつか紹介しておこう。

- 推測変換エンジンの強化によって、入力ミスの文字列から適切な変換候補を推測して提示する機能
- 日本語入力をオフにした状態で入力した文字を、+ キーを押すことで、再入力しないで変換可能な状態にする機能
- 日本語として理解しづらいカタカナ語を、適切な日本語表現に言い換えるための変換候補を提示する校正支援機能
- 数値を入力した際、カンマ区切りや漢字の接頭

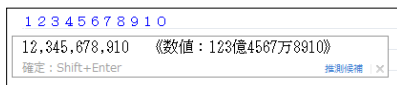


図3 ATOKで数字を入力した際、別ウィンドウで可読性向上の表示がされる

辞(千万億など)を表示して桁数を間違えないようにする機能

それ以外にも、ATOKの辞書を最大10台のデバイスで共有する「ATOK Passport(有料)」やATOK Passportの利用者に向けた大規模クラウド推測変換辞書サービスとして「ATOK クラウド推測変換サービス」が提供されるなど、Google IMEなどの競合IMEの機能を取り込もうとするかのようである。

これに加えて、Apple Computerの提供するモバイルOSであるiOSの新バージョン8では、いよいよサードパーティーによるソフトウェアキーボード+IMEの提供が可能となる予定というニュースがでたが、それに対し、ジャストシステム社としても前向きにATOKの実装を検討しているとのコメントがBLOGサイトに掲載されている^[6]。

今後も、ますますATOKの活躍の場が広がることはうれしいのだが、今年のバージョンは、時に不可解な変換ミスが見受けられるような気もする。

例えば、「いたいじきのう」と読みを入力して変換すると「痛い直納」となってしまう「異体字機能」と変換してくれないなどである。一度正しい例で確定してしまえば、それ以降は正しい候補が表示されるのだが、間違っ変換してしまった結果を確定すると、辞書をメンテナンスしない限りそのままそれが有力候補として表示されてしまう。

それ自体はATOKの正しい機能なのだろうし、筆者の日本語区切りがまずいかもしれないが、どうも今年のバージョンは、読みを文節や単語に区切る際に、区切られる形が筆者の思い通りにならないような気がする。読者諸賢においてはいかがなものだろうか。

❖ おわりに

以上、一太郎2014 徹並びにATOK2014のレビューをお送りした。

本年、一太郎とATOKの開発元である(株)ジャストシステムの名前が久しぶりに世間の耳目を集めた。

ただし、(株)ベネッセコーポレーションから流出した顧客情報に関する件という、決して喜ばしくない方面ではあったのだが。おそらくこの件は、情報セキュリティや個人情報の保護に関する事例として、後々も語られることになるだろうが、このところ好調であった業績に水を差すような事態にならないように、しかるべき対応をとるように願うところである^[7]。

実は本稿を執筆する前、上記IVSに関する誤解もあって、今年は何かしら特筆するところはあるのだろうか、やや暗い気持ちを持っていた。

しかし、いざ書き始めてみると、今までの執筆環境を一変させるような新機能はないものの、従来の機能をより充実するような形での実装に触れるにつれ、今年は書いていてなかなか楽しいレビューとなった。さて来年はどのようなレビューになるのか。今から楽しみである。

注

- [1] ジャストシステム社の元社員であり、UnicodeのCJK統合漢字やIVSに関して多大なる貢献を果たした小林龍生氏の著書『ユニコード戦記—文字符号の国際標準化バトル』(東京電機大学出版局 2011年)に、一太郎を用いてIVSのデモンストレーションを行う場面が記されている。
- [2] これについては、以前に筆者が本誌13号にて「一太郎の異体字変換機能」と題した記事として取り上げている。
- [3] Microsoftからは、それ以前のOffice2007と2010向けにIVS異体字を簡単に入力するアドオンとしてUnicode IVS Add-in for Microsoft Officeが提供されている(<http://ivsaddin.codeplex.com/>)。ただし、このアドオンは、IVS異体字を一時的にUnicodeの私用領域にコピーして外字として内部処理しており、厳密な意味でIVS異体字とはいえない(以下、本レビューで明記したURLについては、いずれも2014年8月17日の閲覧となる)。
- [4] 筆者のWebサイトにて、MS-IME 2010を使ったIVS異体字の入力方法を説明している(<http://www.shuiren.org/chuden/teach/ms-ime2010/ime2010-8.htm>)。MS-IME 2013(Office IME 2013)も同じ手順で入力可能なので参考にいただきたい。
- [5] 一覧は、以下のURLより確認可能。

<http://www.justsystems.com/jp/products/ichitaro/feature6.html>

- [6] Engadget Japanese 「ジャストシステム、iOS 8 対応 ATOK の提供に意欲。非アップル製 IME 解禁を「ずっと待ち望んでいた」」

<http://japanese.engadget.com/2014/06/02/ios-8-atok/>

- [7] ベネッセコーポレーションの個人情報漏洩の件に関するジャストシステムの対応については、以下のニュースリリースで確認することができる。

ベネッセコーポレーションの個人情報漏洩の件に対する当社の対応につきまして

<http://www.justsystems.com/jp/news/2014l/news/>

j07111.html

当社の対応についてのご報告

<http://www.justsystems.com/jp/news/2014l/news/j08071.html>

また、Microsoft 社からも、JUST Office のインターフェイスが Microsoft のリボンインターフェイスに酷似したため、販売差し止め請求がされていると報道もある。「日本 MS、ジャストシステムに JUST Office 製品の一部販売差し止めを求める通知書を送付。」

<http://itpro.nikkeibp.co.jp/atcl/column/14/346926/073100024/>（2014 年 8 月 17 日閲覧）

BabelPad

師 茂樹

❖ Unicode 7.0 対応エディタ

BabelPad は、Unicode 7.0 対応を謳う Windows 用のテキストエディタである。ウェブサイト^[1]から無償でダウンロード、利用することができる（ウェブサイトでは寄付を募っている）。

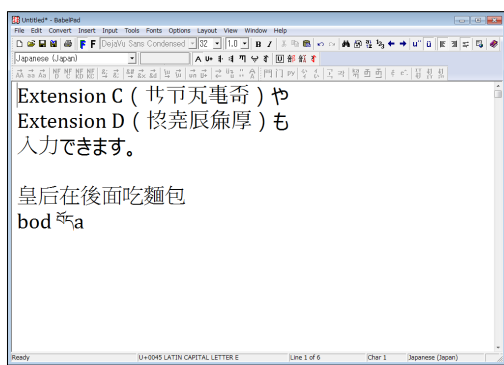


図 1 BabelPad

Unicode 対応のエディタとしては、Windows では EmEditor^[2]などが知られている。BabelPad は、一般的なテキストエディタが備えているコーディング支援（プログラミング言語の予約語などを色分けする機能）や正規表現による検索・置換などの機能は備えていないが、Zero Width Joiner のような Unicode の特殊な文字の挿入機能や、漢字⇄ピンインの変換機能など、一般的なテキストエディタには存在しない機能を備えている。Unicode に関する知識をある程度有し、Unicode の制御文字をフルに使ってプレーンテキストレベルでの多言語処理を行いたいユーザ（そのようなニーズがあるのか、評者にはわからないが）にとっては、これ以上ないエディタであると言えるだろう。

以下、漢字処理を中心に、BabelPad の機能を概観したい。

❖ 柔軟なフォント対応

文字ごとにフォントを指定できるワープロとは異なり、プレーンテキストを扱うテキストエディタは、多

く場合、すべての文字を単一のフォントで表現する。しかし、Unicode 全文字をカバーしている単一のフォントは存在しない。そこで BabelPad では、Unicode のブロックごとにフォントを指定することによって、テキストエディタであるにもかかわらず Unicode の全文字が表示できるようになっている。

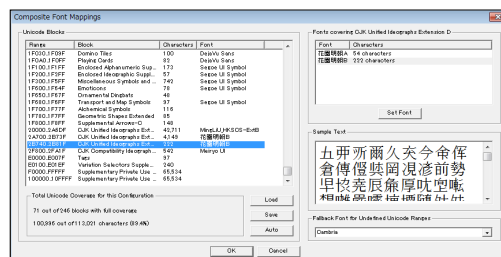


図 2 Composite Font Mappings

もともと BabelPad のデフォルト設定では、すべてのブロックについてフォントが指定されているわけではない。漢字で言えば CJK Extension C と D にはフォントが指定されていないため、デフォルトのままでは適切に表示されない。メニューの [Fonts] → [Composite Font Mappings...] でダイアログを開き、U+2A700 ~ 2B73F の CJK Extension C とその後の CJK Extension D に花明明朝 B^[3]などの対応フォントを設定すれば、適切に表示されるようになる。

ただし、図 1 のように、設定によっては文字ごとにフォントがバラバラになってしまう。一般的なテキストエディタのように単一のフォントで表示することもできるので（メニューの [Fonts] → [Single Font]）、目的に応じて使い分ければよいだろう。

❖ 様々な入力機能

BabelPad は、幅広い Unicode の文字を入力するために、IME 以外にもいくつかの入力手段が用意されている。たとえば、Left-to-Right Mark (U+200E)、Right-to-Left Mark (U+200F) など、文字の方向を制御するための制御文字や、改行しない空白である No-Break Space (U+00A0) をはじめとする様々な空白文字なども、メ

ニューから簡単に入力することができる。また、漢字の細かい異体字を表現するための Ideographic Variation Selector もメニューから入力することができる。

BabelPad は、中国を中心とした東アジアの言語について、機能が充実している。チベット語、モンゴル語、満州語、ウイグル語、彝語については、専用の IME が用意されている。Unicode 全文字をコード表から検索・入力できる Unicode Character Map (図3)、漢字の部首・画数から文字を検索・入力できる Han Radical Lookup Utility(図4)、普通話のピンインから漢字を検索・入力できる Mandarin Pinyin Lookup Utility (図5)、同様に広東語の発音から検索・入力できる Cantonese (Jyutping) Lookup Utility などがある。

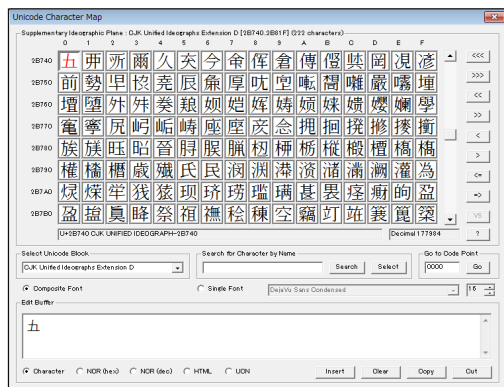


図3 Unicode Character Map

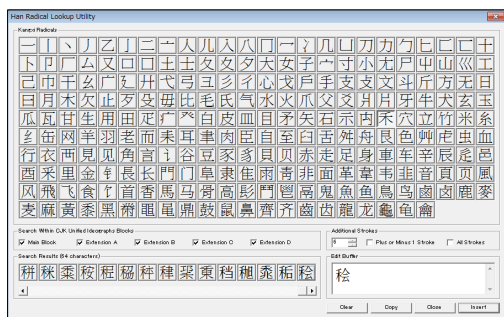


図4 Han Radical Lookup Utility

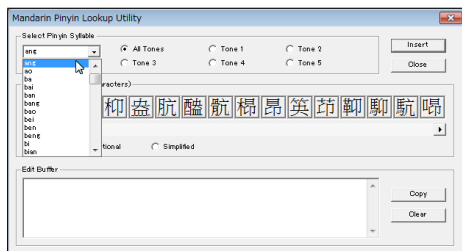


図5 Mandarin Pinyin Lookup Utility

❖ 様々な変換機能

BabelPad は、文字を変換するための機能も充実している。大文字・小文字変換のようなものから、正規形・非正規系変換のような Unicode 固有の文字変換、拡張ワイリー方式のローマ字表記からチベット文字に変換する機能など、多くの変換機能を備えているが、ここでは東アジア文字に関する変換機能をいくつか見てみたい。

■ 簡体字↔繁体字変換

簡体字と繁体字のあいだでは、相互変換が可能である（メニューの [Convert] → [Chinese] → [To Traditional Characters] など）。しかしながら、十分な変換ができるとは言えない。

たとえば、よくある例文「皇后在后面吃麵包」を簡体字に変換すると「皇后在后面吃面包」になる。これを逆に繁体字に変換すると「皇后在后面吃面包」となる。「皇后」が「皇後」にならないのは単語辞書などによってある程度文脈を判断しているからであろうが、「面包」が「麵包」にならないのは文脈把握機能が不十分であることを意味する。

このほか、日本語の新字体・旧字体を相互に変換する機能もある（メニューの [Convert] → [Japanese] → [Kanji To Old Character Forms] など）。

■ 漢字→ピンイン等変換

BabelPad には、漢字からピンインをはじめとするさまざまな発音表記に変換する機能がある。

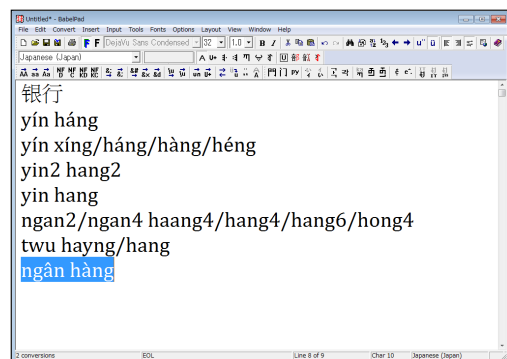


図6 ピンイン等への変換

たとえば「銀行」の場合、文字列を選択してメニュー

の [Convert] → [Chinese] → [Han Ideographs to Pinyin] を実行すれば、「yín háng」と変換される。この表記法は、[Options] → [CJK Readings] → [Use upper case letters] をチェックすれば「YÍN HÁNG」と大文字表記に、同じく [Mandarin tones as numbers] をチェックすれば「yin2 hang2」に、[Omit tones on conversion] をチェックすれば声調を無視した「yin hang」と変換される。また [Convert] → [Chinese] → [Han Ideographs to Pinyin (all readings)] とすれば「yín xíng/háng/hàng/héng」となる。

このほか、広東語、韓国語、ヴェトナム語の発音にも対応している。

- 広 東 語 ([Convert] → [Chinese] → [Han Ideographs to Jyutping]) : ngan2/ngan4 haang4/hang4/hang6/hong4
- 韓 国 語 ([Convert] → [Korean] → [Han Ideographs to Yale Romanization]) : twu hayng/hang
- ヲ ヲ ナ ム 語 ([Convert] → [Vietnamese] → [Han Ideographs to Vietnamese Alphabet]) : ngán hang

この機能も、簡体字⇄繁体字変換機能と同様、ある程度の文脈を把握しているようだが、適切に変換できない場合がある。

❖ インターフェース

BabelPad は、デフォルトで英語のメニューを中国語 (簡体字・繁体字) に変更することができる ([Options] → [User Interface Language])。また、メニューの一部はツールバー

のボタンからも使用できる。ただし、デフォルトでは表示されていないツールバーがあるので、たとえば変換に関するツールバーであれば [View] → [Convert Toolbar] で表示させなければならない。

❖ おわりに

以上、BabelPad の機能の一部を紹介した。他にもたくさん機能があるので、詳細はウェブサイトやヘルプファイル^[4]を参照していただきたい。

冒頭でも述べたように、Unicode のブロックや制御文字などについて知識があるユーザにとっては痒いところに手が届く便利なエディタであろう。また、中国語や、チベット語・モンゴル語などの中国で用いられている言語を扱うことが多いユーザにとっては、便利な機能が多い。しかし、テキストエディタとしての機能は基本的なものしか備えていないので、BabelPad だけですべてをこなそうとするのではなく、メインのエディタに対する補助のエディタとして用いるのがよいのかもしれない。

注

- [1] <http://www.babelstone.co.uk/Software/BabelPad.html>
- [2] <http://jp.emeditor.com/>
- [3] <http://fonts.jp/hanazono/>
- [4] 執筆時点、サイトで公開されているヘルプファイルは古いバージョンのもので、最新の情報は実際に使用してみて確認する必要がある。

Windows 8.1 環境におけるピンイン処理

千田 大介

❖ Win 8.1+Word2013 のピンインルビ

2013 年にリリースされた Windows 8.1 では、起動時にデスクトップが表示されるようになる、Windows ストアアプリがデスクトップのタスクバーに表示されるようになる、アンチウイルスソフトである Windows Defender が標準搭載されるなど、Windows 8 で批判されていたユーザビリティが多数改善され、セキュリティも強化されたことで、一定の評価を得ている。Windows 8 からアップデートした人も多いのではないと思われる。

しかし、こと中国語環境については問題が発生している。IME が更新されたことにともない、Word を使った中国語の漢字にピンインのルビを自動で振ることができなくなったのだ。マイクロソフトによると、これはピンイン IME の仕様変更によるもので、現時点では解決方法がないという^[1]。

Word による自動ピンインルビ機能は、IME の辞書を用いており、複数の読みを持つ漢字についても熟語単位で正しいルビを振ることができるので、Proofing Tools の単語認識機能と併せて使うことで、中国語教材を効率的に作成できる利点があった。

しかし一方で、Windows や Office のアップデートにともない、突如ピンインの自動ルビ振りができなくなる事態がこれまでもしばしば発生してきた。そこからすると、Windows 8.1 でも今後のアップデートによって問題が解消される可能性はあるものの、しばらく先のことになるだろう。

そこで、以下、その対策についていささか考えてみたい。

❖ ピンインの自動生成サービス

今回、Win 8.1 でピンインルビが振れなくなった影響が従来にも増して深刻なのは、無償オンライン中国語ツールの定番であった NAVER が 2013 年 12 月をもって、中国語辞書や中国語の単語句切り・ピンイン生成などに対応した単語分析ツールの公開を打ち切ったこ

とも起因する。このため、それら以外の中国語の漢字からピンインをできるだけ正確に生成できる別のツールを改めて探す必要がある。

実のところ、中国語の漢字からピンインを自動生成するソフト・サービスは数多い。有償ソフトでは高電社の j 北京などがあるし、Google で検索すれば、フリーのサービスも多数ヒットする。ここでは有償サービスについてはさておき、フリーサービスについて見ていくことにするが、Web 上のサービスの大半は、漢字とピンインを一対一対応で変換するもので、意味によって異なる読み方をする漢字、いわゆる破音字への対応に難がある。

■ Google 翻訳でピンイン

<https://translate.google.com/>

執筆時点において、もっとも手軽に利用できるピンイン自動生成サービスは、Google 翻訳であると思われる。Google 翻訳はその名の通り、テキストの翻訳サービスであるが、以下の手順で中国語のピンインを表示することができる。

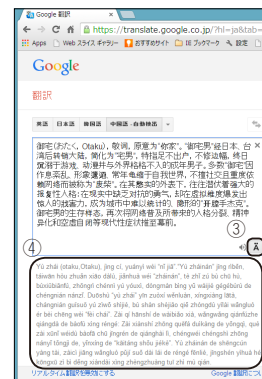


①原文のテキストボックスにピンインを生成したい文字列をペーストする。

②言語を「自動検出」、あるいは「中国語（簡体）」に設定する。

③テキストボック

スの右下に「発音を表示」ボタンが表示される



ので、クリックする。

④ピンインが表示される。

生成されるピンインであるが、辞書データが反映されているとおぼしく、例えば以下の例のように文脈に応じて「長」の字の正しい発音が選択されている。

校长等了很长时间

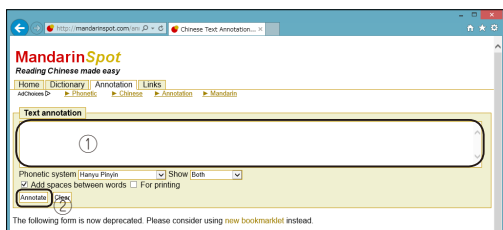
Xiàozhǎng děngle hěn cháng shíjiān

Google 翻訳のピンインは、自動で単語に句切られているが、漢字の方は句切られていない。また、漢字とピンインがそれぞれ別個のテキストデータになってしまうので、Word のルビ機能を使って教材を作るのには向いていない。

- MandarinSpot

<http://mandarinspot.com/annotate>

英語による中国語学習支援サイトである。



①「Text annotation」のボックスに中国語のテキストデータを貼りつける。

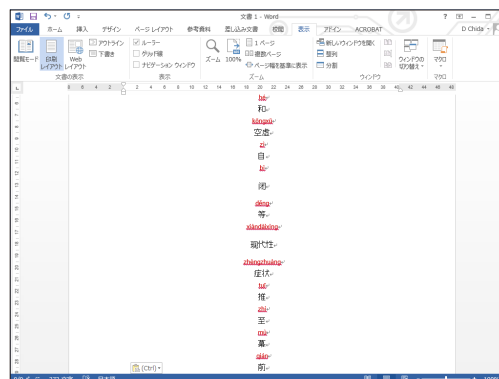
② 「Annotate」 ボタンをクリックする。



③漢字とピンインの分析結果が表示される。

ピンインのほか、注音符号・ウエード式などに変換することもできる。MandarinSpot もピンインの生成に単語辞書が使われており、熟語によって破音字を区別して変換してくれる。分析結果の各単語にカーソルを合わせると、中英辞書がポップアップ表示される。

こちらでもネックになるのが分析結果の再利用である。変換されたピンインルビ付きの漢字を Word にコピーすると、図のように、単語に句切られたピンインと漢字とそれぞれ一行ごとに表示されてしまう。



この結果を、漢字のみの文とピンインのみの文とに仕分けるには、EmEditorなどで正規表現置換する、Excelを使って奇数行・偶数行を抽出するといった方法があるが、Word でルビとして表現するには VBA でプログラミングする必要があるが敷居が高い。

■ Word でピンインルビを擬似的に表現する

以上のように、Google 翻訳にせよ MandarinSpot にせよ、Word のルビ機能でピンインを表現するように加工するのは、少々困難である。ともなれば、別の方法を考える必要がある。

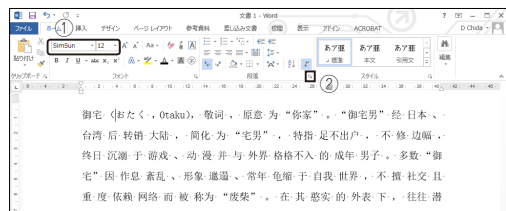
以下、そうした方法の一つとして、Google 翻訳で生成したピンインを、テキストボックスを使って、擬似的にはあるが、ピンイン付きの中国語文書に加工する方法を紹介したい。

作業に際して、あらかじめ簡体字中国語の校正ツールを導入するとともに、筆者の作成した Word 上で単語句切りするマクロ「Word Breaker」を組み込んでおく必要がある^[2]。

Google 翻訳でピンインを生成する前に、まず Word に簡体字の文を貼りつけて、Word Breaker マクロを使い、半角スペースで単語句切りし、単語句切りが正しいか、

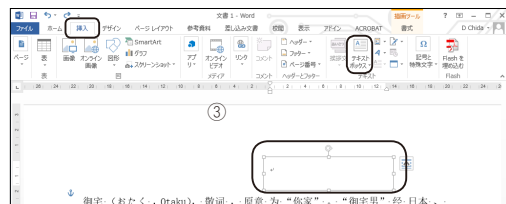
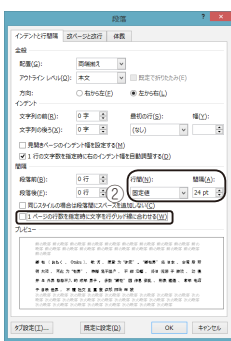
レビュー & リソース紹介

確認・調整しておく。単語句切りした中国語（簡体字）を Google 翻訳に貼りつけて、ピンインを生成しておく。

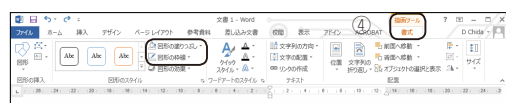


① Word で単語句切りされた簡体字の文章を整形する。全体のフォントを SimSun 12pt に設定する。

②全文を選択し、段落グループ右下をクリックして、段落ダイアログを開く。行間隔を固定値 24pt に設定し、「1 ページの行数を指定時に文字を行グリッド線に合わせる」のチェックを外す。

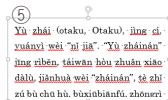


③挿入リボン→「テキストボックス」をクリックし、「横書きテキストボックスの挿入」を選択、任意の位置でドラッグする。

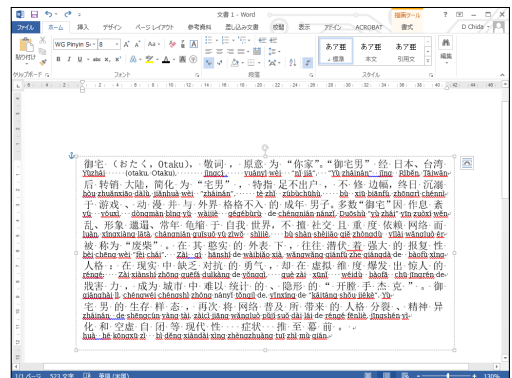


④「描画ツール」をクリックする。「図形の塗りつぶし」を「塗りつぶしなし」に、「図形の枠線」を「線なし」に設定する。

⑤テキストボックスに Google 翻訳で生成されたピンインをコピーし、テキストボックス内をクリック、ホームタブ→「貼りつけ」→「形式を選択して貼り付け」→「Unicode テキスト」を選択して貼りつける。



⑥テキストボックス内をクリックし、を押してピンイン文字列を全て選択し、ピンイン用のフォントを選択、文字サイズを 8pt に設定し、②の要領で行間隔を固定値・24pt に設定する。



⑦テキストボックスのサイズ・位置を漢字に合わせて調整する。

⑧漢字の一行にあわせて、テキストボックス内のピンインを改行し、半角スペースで漢字とピンインの位置を揃える。

行間隔は、固定値に設定することで、文字サイズの違う漢字とピンインとで揃えることができる。漢字文字サイズの倍以上に設定するとよい。ピンインのフォントサイズは、漢字の 2/3 が目安となる。

以上の方法を使えば、ピンインを文字列の上・下、どちらにも表示できる。ただし、漢字の単語間隔を調整するには、上に被さっているピンインのテキストボックスをその都度ずらさなくてはならないなど、漢字とピンインとの位置あわせにかなりの手間がかかる。

マイクロソフトでは、2014 年秋にも次期 Windows 9 のプレビュー版を公開し、2015 年中にデビューさせる予定であるという。バージョンアップを機に、Word によるピンインの自動ルビ振り機能が復活することに期待したい。

■ ピンインフォントの更新

最後にいささか手前味噌であるが、筆者が配布しているフリーのピンインフォントの更新についてお知らせしておく。従来、ピンインフォントとしては、セ

リフ系の書体である WG Pinyin Serif、サンズセリフ系の WG Pinyin Sans、およびそれぞれを簡体字中国語フォントとして設定した WG Pinyin Serif+・WG Pinyin Sans+ を配布してきたが、このほど中国語教材を作成する上で必要が生じたため、それぞれの Bold 書体を追加した。

Āiyā, wǒ diū le qiánbāo! (WG Pinyin Serif)

Āiyā, wǒ diū le qiánbāo! (WG Pinyin Serif Bold)

Āiyā, wǒ diū le qiánbāo! (WG Pinyin Sans)

Āiyā, wǒ diū le qiánbāo! (WG Pinyin Sans Bold)

Word などで Bold フォントを使う場合は、WG Pinyin Serif・WG Pinyin Sans などを選択してから、太字に設定すればよい。

これらのピンインフォントは筆者の Web サイト

(<http://wagang.econ.hc.keio.ac.jp/>) にてフリーで配布しているので、ご活用いただきたい。

注

- [1] http://answers.microsoft.com/ja-jp/windows/forum/windows8_1-ime/ime2012-microsoft/632c36e2-609e-41ed-a75a-e9f7f7157477
- [2] Office 2013 校正ツールの入手については本誌第 14 号「Windows 8 & Office 2013」レビュー p.144、Word Breaker マクロについては同「新 HSK 単語の性質について」p.10（いずれも筆者著）を参照されたい。Word Breaker の配布ページには、電腦瓦崗寨 (<http://wagang.econ.hc.keio.ac.jp/>) の左メニュー「ダウンロード」からたどっていただきたい。

When.exe Ruby 版

須賀 隆

gems https://rubygems.org/gems/when_execode https://github.com/suchowan/when_exe

❖ 概要

When.exe Ruby 版は、かつて筆者が開発した When.exe MS-DOS 版^[1]の後継版で、あらゆる暦を表現し変換することを目指している Ruby 言語のライブラリである。

■ 目的・用途

When.exe Ruby 版の基本的な目的は古今東西あらゆる文化および言語で用いられた暦日・暦法・時法・暦年代・暦注などに共有可能な名前付けを行い、統一的に扱うための枠組みを提供することであり、ISO8601・ISO19108・RFC5545 (iCalendar)・RFC6350 などの規格に対応している。結果として暦日・暦年代の変換、暦注の表示を行う Web サーバーに向くライブラリとなった。

■ リソースの所在

下記を参照されたい。

demo <http://hosi.org>

wiki <http://www2u.biglobe.ne.jp/~suchowan/>

blog <http://suchowan.at.webry.info/>

❖ Web サーバー

本稿では本誌の読者が直接利用する機会があると思われる同ライブラリを利用したデモンストレーション Web サーバーについて、主にアジアの暦日を例に解説する。ライブラリ本体については wiki の Ruby 以下の項目を参照されたい。

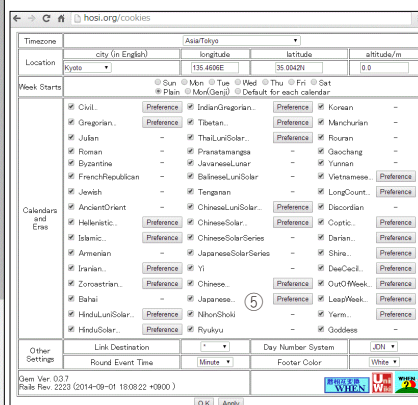
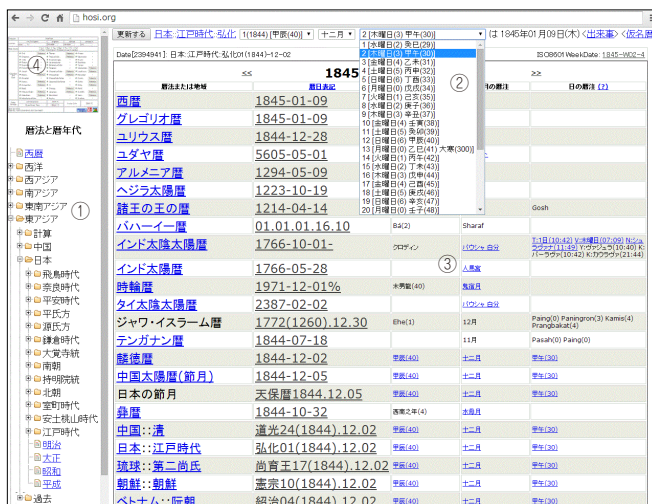
■ 画面構成

●主画面 (<http://hosi.org>)

主画面は暦法および暦年代（以下、両者を総称して単に暦法と呼ぶ）選択ツリービュー（①、左部）、暦日選択フォーム（②、右上部）、暦日表示部（③、右下部）からなる。暦日表示部には各種暦法での日付表記とその暦法での主な暦注を表示する。表示しようとする暦日の期間によって、各種暦法による一日分の表示（一日表示）、特定の暦法による一月分の表示（一月表示）、同じく一年分の表示（一年表示）になる。

●一般設定画面 (<http://hosi.org/cookies>)

主画面の左上部の一般設定画面図（④）をクリックすると一般設定画面が開く。この画面ではユーザの時間帯や経緯度、一月表示・一年表示の表示方法、表示する暦法の選択（チェックボックスで切替）、その他の設定ができる。設



定はユーザ側のブラウザのクッキーに記憶される^[2]。

●詳細設定画面 (<http://hosi.org/cookies/nn>)

一般設定画面の暦法選択行に Preference ボタンがある場合、その暦法のための詳細設定項目がある。詳細設定の内容は暦法ごとに異なる。日本暦日 (nm=Japanese) の場合 (⑤)、使用する朔閏表の選択、一月表示で表示する暦注の項目の指定、その他の設定ができる。



●仮名暦画面 (<http://hosi.org/Note/yyyyv>)

主画面の暦日選択フォームで江戸時代 (貞享年間以降) が選択されている場合、その年 (yyyy=1685~1872) の仮名暦を表示できる。

●参考情報画面

画面の各所に配置されたリンクから、必要に応じて参考情報へのリンクが張られている。例えば一日表示の2列目「暦日表記」から暦日表記の説明画面^[3]が開く。リンクの多くは、その用語を説明する Wikipedia への外部リンクである。

■ 暦日表示例

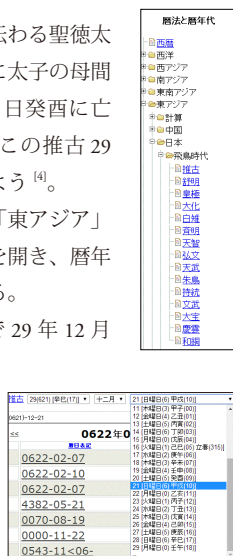
●天寿国繡帳銘

奈良県斑鳩町の中宮寺に伝わる聖徳太子ゆかりの天寿国繡帳の銘に太子の母間人皇后が推古29年12月21日癸酉に亡くなったとの記述がある。この推古29年12月21日を表示してみよう^[4]。

暦法選択ツリービューで「東アジア」→「日本」→「飛鳥時代」を開き、暦年代として「推古」を選択する。

次に暦日選択フォームで29年12月21日を選ぶ。暦日選択フォームでは年と日の干支も表示するので、この時点で12月21日が甲戌であることがわかる。

「更新する」ボタンを押す



て、この日の各種暦法での暦日を見てみよう。

麟徳暦	0621-12-22	癸巳(127)	十二月	甲戌(18)
中国太陽暦(節月)	0622-01-07	壬午(18)	正月	甲戌(18)
中国の節月	戊寅元暦0622.01.05	壬午(18)	正月	甲戌(18)
日本の節月	元嘉暦0622.01.06	壬午(18)	正月	甲戌(18)
彝暦	0622-01-31	酉辰(2年(8))	十二月	甲戌(18)
中国::唐::梁(梁師都)	永隆05(0621).12.21	癸巳(127)	十二月	甲戌(18)
中国::唐::楚(林士弘)	太平06(0621).12.21	癸巳(127)	十二月	甲戌(18)
中国::唐	武德04(0621).12.22	癸巳(127)	十二月	甲戌(18)
日本::飛鳥時代	推古29(0621).12.21	壬午(18)	十二月	甲戌(18)
(日本書紀)	推古29(0621).12.21	癸巳(127)	十二月	甲戌(18)
朝鮮::高句麗	榮留王04(0621).12.22	癸巳(127)	十二月	甲戌(18)

「中国の節月」「日本の節月」で確認できるように、当時唐では戊寅元暦、日本では元嘉暦が使われていた。東アジアの国々の日付をみると、日本は12月21日、唐は12月22日となっている (高句麗については後述)。一日表示の「<<」ボタンで前日を見てみよう。

麟徳暦	0621-12-21	癸巳(127)	十二月	癸酉(09)
中国太陽暦(節月)	0622-01-06	壬午(18)	正月	癸酉(09)
中国の節月	戊寅元暦0622.01.04	壬午(18)	正月	癸酉(09)
日本の節月	元嘉暦0622.01.05	壬午(18)	正月	癸酉(09)
彝暦	0622-01-30	酉辰(2年(8))	十二月	癸酉(09)
中国::唐::梁(梁師都)	永隆05(0621).12.20	癸巳(127)	十二月	癸酉(09)
中国::唐::楚(林士弘)	太平06(0621).12.20	癸巳(127)	十二月	癸酉(09)
中国::唐	武德04(0621).12.21	癸巳(127)	十二月	癸酉(09)
日本::飛鳥時代	推古29(0621).12.20	壬午(18)	十二月	癸酉(09)
(日本書紀)	推古29(0621).12.20	癸巳(127)	十二月	癸酉(09)
朝鮮::高句麗	榮留王04(0621).12.21	癸巳(127)	十二月	癸酉(09)

今度は唐が12月21日癸酉になる。この時期は中国で隋代まで使われていた平朔^[5]の暦法が唐の戊寅元暦で定朔^[5]の暦法に切り替わった、まさにその時期にあたっている。12月21日癸酉というのは定朔の暦法によるもので、天寿国繡帳の銘が後代に唐から定朔の暦法が伝わってから製作された可能性が高いことがわかる。

日本での定朔の暦法の

始めは麟徳暦^[6]である。

ChineseLuniSolar 詳細設定

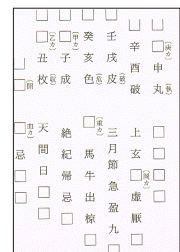
画面で「麟徳暦」を選択す

ると日本では697年~763年の期間に使われたことがわかる。暦法選択ツリービューで「東アジア」→「計算」→「麟徳暦」を開き、暦日選択フォームで621年12月21日を選ぶと確かに干支は癸酉である。

●石神遺跡出土具注暦

次に奈良県明日香村石神遺跡で出土した具注暦を再現してみよう。この具注暦の表面には、二行目に「辛酉破上玄」翌日の三行目に「壬戌(破)三月節」と書いてある^[7]。

具注暦を表示するにはあらか



レビュー & リソース紹介

じめ一般設定画面の“Week Starts”を“Plain”に設定しておく。暦法選択ツリービューで「東アジア」→「日本」→「飛鳥時代」を開き、暦年代として「持統」を選択する。次に暦日選択フォームで3年3月を選ぶ。暦日選択フォームでは二十四節気も表示するので、この時点で3月10日が清明であることがわかる。日に“-”(日付リストの最後の選択肢)を指定して一月表示を選択し「更新する」ボタンを押して、この月の具注暦を見てみよう。

／拡大／

09 (04-06)	辛酉(57) 納音是木 曜土曜日(5) 上弦 陽錯 庚辰 人神配当風
10 (04-07)	壬戌(58) 納音是水 曜日曜日(6) 清明(15) 三節 歳後 九坎

九日が上弦、十日が三月清明節になっており日の干支(辛酉・壬戌)も十二直(破)も一致する、確かに出土した具注暦と矛盾しない^[8]。

この時代は元嘉暦なので、ChineseLuniSolar 詳細設定画面で「元嘉暦」「日本暦注」を選択し、暦法選択ツリービューで「東アジア」→「計算」→「元嘉暦」を開き、暦日選択フォームで689年3月を選んで一月表示しても同じ具注暦が見える。これは中国の具注暦を表示するための実験的な機能で、朔閏を史実でなく計算によっており、また日本と中国の暦注撰日法の違いも考慮していない。

●天保暦仮名暦

最後に江戸時代の仮名暦を見てみよう。暦法選択ツリービューで「東アジア」→「日本」→「江戸時代」を開き、暦年代として「弘化」を選択する。暦日選択フォームに自動的に「弘化1甲辰十二月2甲午」が選択されるが、これは天保から弘化への改元日である。このフォームの右側のリンク<出来事>を開くと、Wikipedia 日本語版の「12月2日(旧暦)」の記事が表示される。「できごと」欄に確かに弘化改元記事がある。

暦日選択フォームの右端のリンク<仮名暦>を開くと「天保十五年きのえたつ乃天保壬寅元暦」の仮名暦が表示される。弘化元年暦は前年の天保十四年末に出版されたのである。仮名暦画面はHTMLで可能な限り忠実に実際の仮名暦にあわせたレイアウトにしており、Internet Explorer 11.0以上では縦書きで表示される。冒頭部分を下図に示す。

この年は寛政暦から天保暦への改暦が行われたため、仮名暦の前文に改暦の経緯・内容が書かれていた。仮名暦画面ではこれも忠実に再現している。現時点での実装は実験的なものであり、下段の吉事注に過不足があること、日食・月食記事の詳細説明部が反映されていないこと、弘化三年以降の暦の土用の日時にずれがあることなど制限事項がある。

❖ 実装の背景

■ 出典および精度

When.exe MS-DOS 版から継承したデータは多岐にわたるため wiki の暦説明 (特に「中国」および「参考文献」) を参照願いたい。ここでは今年に入って新規に取り入れた (あるいは取り入れなかった) データについて補足する。

●日本および中国 (唐代)

細井浩志・竹迫忍『唐・日本における進朔に関する研究』(科研 22520700) (2013)

Japanese 詳細設定画面と Chinese 詳細設定画面で朔閏表として「科研」を選択すると、科研 22520700 の成果物『唐・日本における進朔に関する研究』に主に^[9] 基づいた朔閏表を用いる。この研究は唐・日本での進朔^[10] のパターンを文献および金石文と計算を比較しつつ考証したもので8~9世紀の暦日について従来と異なる見解を提案している。研究の評価は今後であろう。実装では、すでに「規範」として定着し多

くの文献で幅広く利用されている『日本暦日原典』『二十史朔閏表』などと切り替えて使用できるようにした。

●朝鮮半島

Korea Astronomy Observatory『三国時代年曆表』『高麗時代年曆表』『朝鮮時代年曆表』(1999-2002)

朝鮮半島の朔閏表である^[11]。具体的な暦日を確認したところ、中国“正統”王朝の朔閏表をまず定義して、『三国史記』などの文献や金石文からの差分を反映したものである。このためほぼ確実に史実とは異なると考えられる部分もあるが今回の実装ではそのまま採用した。13世紀半ばよりも前に遡った暦日については、この事情を理解した上で取り扱う必要がある。たとえば、唐の成立当初から朝鮮半島で唐の暦(戊寅元暦)が使用されたことになっているが、成立直後の唐は内陸部の地方政権であり朝鮮半島には影響力がない。戊寅元暦は史上初めて定朔を採用した暦法で、それ以前の平朔の暦とは約半数の朔の日付がずれる。天寿国繡帳の例の高句麗暦日にはこのような限界がある。

●ベトナム

Lê Thành Lân “Đổi chiếu lịch dương với lịch âm dương của Việt nam và Trung quốc 2030 năm” (2007)

ベトナムの朔閏表である。西暦1543年以前は中国の朔閏をそのまま用い、西暦1544年以降についてはベトナム国立図書館所蔵の『欽定萬年書』などの史料に基づいて各政権の朔閏表を編纂している。この朔閏表は暦年代に『大越史記全書』を使用しているようであるが、ベトナムの暦年代の基本史料には『大越史記全書』と『越史略』があり、暦年代に差がみられる。今回の実装では Vietnamese 詳細設定画面で暦年代として『大越史記全書』と『越史略』のいずれを用いるか

選択できるようにしている^[12]。『大越史記全書』と『越史略』の差については天文現象の記事に着目した研究も進められているようである。

●中国戦国時代

中国戦国時代の暦は実装していない。安大玉「中国におけるおこな暦法」(朝倉書店『暦の大事典』(2014)) p.235の戦国時代の古六暦の項には、

古六暦の置閏法は、無中気月に閏月を置く方法がすでに定着していたとする見解もあるが、現存の暦日記録からみれば、閏12月、後9月など(10月を歳首とする顓頊暦の場合)とあり、歳終置閏法による可能性が高い。

とあり、専門家の見解も収束していないとみられる。この状況で秦代より前に遡ったデータを提供するのは難しいのではないだろうか。

Web サーバーの実装としては、本節に書いたような制限情報を主画面から容易に参照可能とすることが課題である。

■ 多国語対応および用語のゆれ

When.exe Ruby 版はあらゆる言語で用いられた暦を表現することを目指しているので多国語対応は避けられない。Web サーバーではブラウザの使用言語を検出して七曜やグレゴリオ暦などの月名表記、暦日選択フォームの<出来事>のリンク先ページなどを切り替えるようにしている。Chromeの言語設定をヒンディー語にして、インドの太陰太陽暦の2014年8月下旬から9月上旬に相当する半月分を表示した例を下図に示

Date[2456896]: 1936-06-~HinduLuniSolarNoteHinduNote&location[co.Indian:Ujjain]&start_month[5]&type[SESA]						
<< 1935 HinduLuniSolar 1936 1937 >>						
<< सप्तमशतकम् कालचक्रपञ्चकम् अश्विनशतकम् शतकलपञ्चकम् सप्तमशतकम् कालचक्रपञ्चकम् >>						
रवि	सोम	मंगल	बुध	गुरु	शुक्र	शनि
06: (2014-09-24)	07: (2014-09-25)	08: (2014-09-26)	09: (2014-09-27)	10: (2014-09-28)	11: (2014-09-29)	12: (2014-09-30)
T:अश्विनशतकम्(18:10:22) V:रविशतकम्(06:07:23) N:चक्र. कालचक्रपञ्चकम्(05:39:09) Y:शतकम्(14:49:06) K:अश्विनशतकम्(18:10:22)	T:अश्विनशतकम्(20:07:41) V:रविशतकम्(06:08:01) N:चक्र. कालचक्रपञ्चकम्(08:06:23) Y:शतकम्(15:50:52) K:अश्विनशतकम्(20:07:41)	T:शुक्रशतकम्(21:51:01) V:बुधशतकम्(06:08:40) N:चक्र. कालचक्रपञ्चकम्(12:39:55) Y:शतकम्(15:50:52) K:अश्विनशतकम्(21:51:01)	T:शुक्रशतकम्(23:12:41) V:शुक्रशतकम्(06:09:19) N:चक्र. कालचक्रपञ्चकम्(12:39:55) Y:शतकम्(15:50:52) K:अश्विनशतकम्(23:12:41)	T:अश्विनशतकम्(00:07:14) V:शुक्रशतकम्(06:09:58) N:चक्र. कालचक्रपञ्चकम्(12:39:55) Y:शतकम्(15:50:52) K:अश्विनशतकम्(00:07:14)	T:अश्विनशतकम्(00:31:33) V:शुक्रशतकम्(06:10:38) N:चक्र. कालचक्रपञ्चकम्(12:39:55) Y:शतकम्(15:50:52) K:अश्विनशतकम्(00:31:33)	T:अश्विनशतकम्(00:31:33) V:शुक्रशतकम्(06:10:38) N:चक्र. कालचक्रपञ्चकम्(12:39:55) Y:शतकम्(15:50:52) K:अश्विनशतकम्(00:31:33)
06: (2014-09-31)	07: (2014-09-01)	08: (2014-09-02)	09: (2014-09-03)	10: (2014-09-04)	11: (2014-09-05)	12: (2014-09-06)
T:अश्विनशतकम्(00:24:46) V:रविशतकम्(06:11:17) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(00:24:46)	T:अश्विनशतकम्(23:47:53) V:रविशतकम्(06:11:57) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(23:47:53)	T:अश्विनशतकम्(22:43:16) V:रविशतकम्(06:12:37) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(22:43:16)	T:अश्विनशतकम्(21:14:14) V:रविशतकम्(06:13:17) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(21:14:14)	T:अश्विनशतकम्(19:24:47) V:रविशतकम्(06:13:57) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(19:24:47)	T:अश्विनशतकम्(17:19:12) V:रविशतकम्(06:14:37) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(17:19:12)	T:अश्विनशतकम्(15:01:59) V:रविशतकम्(06:15:17) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(15:01:59)
13: (2014-09-07)	14: (2014-09-08)	15: (2014-09-09)	16: (2014-09-10)	17: (2014-09-11)	18: (2014-09-12)	19: (2014-09-13)
T:अश्विनशतकम्(12:37:47) V:रविशतकम्(06:15:58) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(12:37:47)	T:अश्विनशतकम्(10:11:11) V:रविशतकम्(06:16:38) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(10:11:11)	T:अश्विनशतकम्(09:00:33) V:रविशतकम्(06:17:19) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(09:00:33)	T:अश्विनशतकम्(07:46:53) V:रविशतकम्(06:18:00) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(07:46:53)	T:अश्विनशतकम्(06:23:43) V:रविशतकम्(06:18:41) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(06:23:43)	T:अश्विनशतकम्(05:01:59) V:रविशतकम्(06:19:22) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(05:01:59)	T:अश्विनशतकम्(03:50:28) V:रविशतकम्(06:20:03) N:चक्र. कालचक्रपञ्चकम्(14:49:06) Y:शतकम्(14:49:06) K:अश्विनशतकम्(03:50:28)

す^[13]。

用語のゆれに関しては、When.exe Ruby 版ライブラリ本体は対応しているが Web サーバーでは特別の対応を行っていない。このため、Web サーバーでの表記は用語の代表的な表記となる。たとえば、暦法選択ツリービューで「東アジア」→「日本」→「江戸時代」を開くと、暦年代の選択肢に「慶應」はなく「慶応」があらわれる。しかし Chrome で直接 URL “http://hosi.org/* / 慶應 3-3-1” を表示させると、特にエラーになることなく慶応 3 年 3 月 1 日の暦を表示する。これは When.exe Ruby 版ライブラリ本体が「慶応」の別名として「慶應」を認識しているためである。同じく「東アジア」→「中国」→「隋」→「梁(蕭銑)」を開くと暦年代の選択肢に「鳳鳴」のみ現れるが、Chrome で直接 URL “http://hosi.org/* / 鳳鳴 1-10-1” を表示させると鳳鳴 1 年 10 月 1 日の暦を表示する。この年号の『資治通鑑』での表記が「鳴鳳」なのである。中国に関しては本稿執筆時点では整合性を確認中であり、まだ十分ゆれに対応している状況ではない。

先の URL で鳳鳴 1 年 10 月 1 日を表示させると暦日選択フォームには「中国 :: 隋 :: 梁(蕭銑) :: 鳳鳴 1 十月 1 戊寅」が選択された状態になる。このリンク「鳳鳴」をクリックすると Wikipedia の「鳳鳴」の項が参考情報画面に表示される。これはユーザにとって便利な説明としてインタフェースを用意したというよりも、むしろ When.exe Ruby 版ライブラリが「鳳鳴」という年号をひとつの概念と捉え、その表記のゆれや意味説明へのリンクなどをひとまとまりにして“semantics”として管理していることをデモンストレーションするためのものである。

■ 共有可能な名前付け

「聖武天皇の命日、1 日間違えた」という記事が日経新聞電子版(2012/9/26)^[14]に出た。記事によれば、

「天平勝宝 8 年 5 月 2 日」は太陽暦で「756 年 6 月 8 日」に当たるのに、同庁の資料には 1 日早い「756 年 6 月 7 日」と間違えて記載しているのに気付いた。<中略>新暦と旧暦の換算は難解で、むしろ計算機がない時代に 2 つしか誤りがないことに驚いた」と福尾調査官

間違いの直接の原因は 5 月 2 日という日付を内務省

地理局『三正綜覧』(1880)の朔閏表を用いて換算したことである。しかし『続日本紀』で崩御の記事を確認すると 5 月 2 日ではなく天平勝宝 8 歳 5 月乙卯と書かれている^[15]。これを Chrome で直接 URL “http://hosi.org/* / 天平勝宝 8-5-{乙卯}” と指定して暦日を表示させると、確かにグレゴリオ暦 756 年 6 月 8 日である。

[東暦年号] 日本 奈良時代 天平勝宝 (8756) [丙申(32)] 五月		[2 金曜日の(4)乙卯(51)]		[13 756年06月08日(金) 出京事]	
Date[1997583] 天平勝宝8-5-2(日)		[1 休曜日の(3)甲寅(50) 芒種(75)]		ISO8601 WeekDate: 756-352-5	
暦法または地域名		日付の表示 (2)		日付の表示 (2)	
西暦		756-06-04		756-06-04	
グレゴリオ暦		756-06-08		756-06-08	
ユリウス暦		756-06-04		756-06-04	
ユリウス暦		4516-10-02		4516-10-02	
アルマニア暦		0205-01-15		0205-01-15	
ヘジャラ本暦		0135-03-18		0135-03-18	
諸王の王の暦		0124-12-25		0124-12-25	
インド本暦本暦		0678-04-02		0678-04-02	
インド本暦		0677-11-15		0677-11-15	
時輪		0883-06-02		0883-06-02	
中国本暦本暦		1298-08-02		1298-08-02	
新暦		0756-05-02		0756-05-02	
中国本暦(節月)		0756-05-02		0756-05-02	
中国の節月		大山暦0756.05.03		大山暦0756.05.03	
日本の節月		遠東暦0756.05.02		遠東暦0756.05.02	
新暦		0756-05-05		0756-05-05	
中国::唐		玄宗15(0756).05.02		玄宗15(0756).05.02	
中国::唐::高		景正01(0756).05.02		景正01(0756).05.02	
日本::奈良時代		天平勝宝 08(0756).05.02		天平勝宝 08(0756).05.02	

／拡大／

天平勝宝 (8756) [丙申(32)] 五月	[2 金曜日の(4)乙卯(51)]	[13 756年06月08日(金)]
	[1 休曜日の(3)甲寅(50) 芒種(75)]	
	[2 金曜日の(4)乙卯(51)]	

実は When.exe Ruby 版ライブラリはこの URL を解釈するのに朔閏表のみに頼らず五月で干支が乙卯になる日を計算しているのである^[16]。単純に 60 日周期で循環する干支で日付を指定すれば^[17]朔閏表に多少の誤りがあってもグレゴリオ暦日は影響を受けないことが多い。生データを後付けの情報で加工しないにこしたことはない。

現在インターネットで暦日情報を交換する場合 Dublin core^[18]を参照して“dc:Date”でマークアップすることが多いと思われる。これは原則グレゴリオ暦のみを扱う ISO8601 の書式を前提としている。しかし“天平勝宝 8-5-{乙卯}”や文末脚注 13 のようなオリジナルの情報をそのまま保存した書式を規格化し、マークアップの際に利用できれば、膨大な文献のマークアップの品質を向上させ、公の情報共有に貢献することができのではないだろうか。

冒頭に書いた「When.exe Ruby 版の基本的な目的は古今東西あらゆる文化および言語で用いられた暦日・暦法・時法・暦年代・暦注などに共有可能な名前付けを行い、統一的に扱うための枠組みを提供する」とは、このような問題意識によるもので、When.exe Ruby 版

はその具体的な実装例になっているのである。

❖ 現状のステータス

2014-09-07 時点では本ライブラリはβ版で rev.0.3.7 である。今後 API に非互換の変更の可能性があるため現時点では著作権を強く保留している。しかし正式版に移行するにはよりゆるやかな制限に移行する予定である。暦日データは公開してより広い範囲でレビューしていただいた方がより精度が向上すると思われるからである。

When.exe Ruby 版が提案する名前付けを公に共有するための道のりは遠い。本稿がこのような試みが行われていることを世の中に知っていただく契機となつてほしいものである。

注

- [1] <http://www.asahi-net.or.jp/~dd6t-sg/whenhome.html>
- [2] クッキーは暗号化していない。
- [3] http://www2u.biglobe.ne.jp/~suchowan/when_exe_standard_representation.html
- [4] http://suchowan.at.webry.info/201202/article_3.html
- [5] 平朔の暦では月が同じ時間間隔で新月になるものとみなして朔日を決定する。これに対して定朔の暦では天文学的に実際に新月になる日を計算し朔日とする。新月の時刻は平朔と定朔で数時間の差があり、このため朔日の日付が相違することがある。
- [6] 日本では儀鳳暦という暦法名で知られている。
- [7] 図は岡田芳朗「日本最古の暦—その年代推定と暦注解釈」『しにか』2003年8月号 pp.63-69 による。
- [8] 現時点では Web サーバーからは暦注→日付の検索機能はないので、検索には Ruby のライブラリをオフラインで使う必要があるが、本稿では詳細は省略する。日本暦日であれば竹迫忍氏の <http://wagoyomi.info> で暦注→日付の検索が可能である。
- [9] http://suchowan.at.webry.info/201401/article_25.html
- [10] 新月の時刻が夕方以降の場合に、“朔”日を新月の翌日に“進”めること。進朔有無の閾値となる時刻に曖昧性があるため、朔閏表の再現で問題となる。
- [11] http://suchowan.at.webry.info/201405/article_10.html
- [12] http://suchowan.at.webry.info/201404/article_17.html
- [13] [http://hosi.org:3000/Date/1936-06-%5e%5eHinduLuniSolar%3fnote=HinduNote&location=\(co:Indian::Ujjain\)&start_month=5&type=5BSA](http://hosi.org:3000/Date/1936-06-%5e%5eHinduLuniSolar%3fnote=HinduNote&location=(co:Indian::Ujjain)&start_month=5&type=5BSA)
一般設定画面の“Week Starts”を“Default for each calendar”に設定して、上記の URL を直接表示しても良い。“1936…”の部分は“ウッジャインの経緯度を基準にしてスールヤシッダーンタのアルゴリズムで計算したサカ暦 1936 年バードラパダ月白分の半月間”を意味する。なお、この URL に“.json”を加えて Web サーバーに送ると Web サーバーは実際に JSON フォーマットで暦日情報を応答する。
- [14] http://www.nikkei.com/article/DGXNASDG2600B_W2A920C1CR0000/
- [15] <http://kindai.ndl.go.jp/info:ndljp/pid/991092/162>
- [16] http://suchowan.at.webry.info/201209/article_30.html
- [17] http://suchowan.at.webry.info/201409/article_7.html
- [18] http://ja.wikipedia.org/wiki/Dublin_Core
プレフィクス“dc:”のネームスペースは <http://purl.org/dc/elements/1.1/>

中華經典典籍庫

千田 大介

❖ 中華書局の文献データベース

■ 待望の優良底本データベース

中華經典典籍庫は、中国で最も権威ある出版社である中華書局が発売する、中国古典文献の大規模データベースである。収録される文献は、約 300 種、2 億字に上り、全てが中華書局の標点本を底本としている。

従来の古典文献データベースは、著作権等の問題を回避するためもあるのだろう、『四庫全書』・『四部叢刊』のように抄本・版本に基づいたものが多く、底本の質にいささか問題のあるものも多かった。この点、中華經典典籍庫の底本は、二十四史と『清史稿』、十三経、新編諸子集成、『全唐詩』などなど、中国学の各分野において定番のテキストとして扱われることの多い中華書局の標点本であり、しかも整理・校訂を経ているため従来の抄本や版本に基づくデータベースに比べて異体字が減少するとともに、誤字の含まれる率が相当低くなっているものと思われる。その意味で、これまでになく信頼性の高い文献データベース製品に仕上がっていることが期待される。

以下では、筆者が 2014 年 8 月末に同データベースの開発元である北京創新力博社で行ったインタビューと、同社および中国国家図書館で試用した結果に基づき、その特色や主要な機能についてレビューしたい。

❖ 収録文献と基本的機能

■ インターフェイスと収録文献

図 1 は中華經典典籍庫のインターフェイスである。左ペインに文献の分類と名称が表示され、メインペインに検索結果・本文（電子テキスト版）・原文画像を切り替え表示する。『四庫全書』・『四部叢刊』などと基本的に同じである。

収録書籍は伝統的な四部分類でカテゴリ化されており、左ペインのリストを展開し、分類をたどって閲覧したい書籍をクリックすることで、メインペインに本文が表示される。実際には、書名検索・全文検索の結果をクリックして本文を表示する、という使い方が一般的だろう。電子テキストの本文が表示された状態で「本文画像」をクリックすると、当該箇所の影印画像を見ることができる（図 2・3）。



図 1

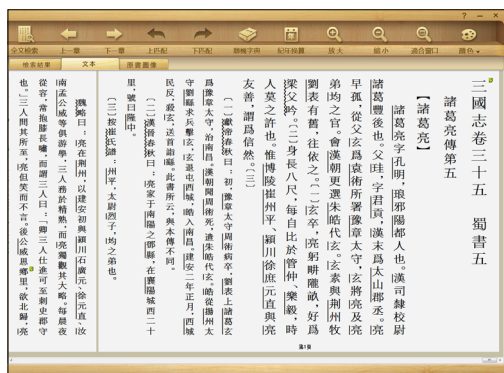


図 2

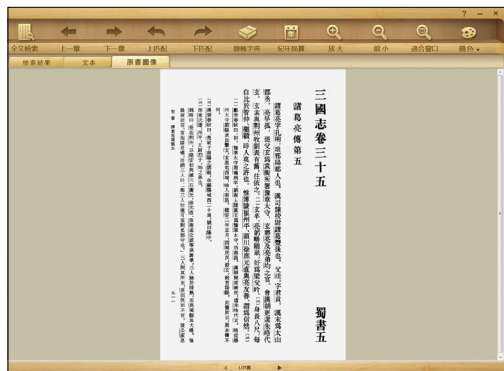


図 3

書籍の分類について、試用したところでは、『論語集解』・『孟子正義』・『四書章句集注』が子部・儒家類に分類されているのが気になった（図4）。新編諸子集成に収録される書籍を子部にまとめたのが原因であると思われるが、本来、経部に分類するべきであろう。こうした点については、今後の改善に期待したい。

主な収録書籍であるが、経部は十三經清人注疏を収録する。史部には二十四史と『清史稿』、『資治通鑑』、歴代史料筆記叢刊などが、子部には新編諸子集成・學術筆記叢刊・理學叢書などが、集部には『先秦漢魏晉南北朝詩』・『全唐詩』・『全宋詞』などの総集、中国古典文学基本叢書などが、それぞれ収録される。ただし、中国古典文学基本叢書は既刊が50種を越えているにもかかわらず収録されるのが11種に



図 4

止まるなど、各シリーズ既刊の全書籍が収録されているわけではない。

文学を専門とする立場からは、集部の収録文献がいささか少ないのが気になった。特に戯曲・章回小説は全く収録されていない。『四庫全書』・『四部叢刊』などの伝統に則ったのか、文化史的に重要な書籍を優先したのかはわからないが、いささか片手落ちの感は否めない^[1]。

ただ以上のような収録文献の問題は、今後、同データベースの続編が発売されるにつれて、改善されていくものと思われる。

■ 文字コードと外字

文字コードについては、UnicodeのExt.Bまでの漢字を使用しており、未収録文字については、Unicode Plane15のF0001～F24A6に大量の外字を登録している。小学関連の書籍が収録されているため、外字には金文なども多く含まれており、それらを収録した専用フォント、「方正宋体超大字庫_中華書局版」が添付される。ユーザーは、フォントを含むクライアントプログラムをインストールして利用することになる。

❖ 検索と引用

■ 基本的検索機能と検索条件式

検索ダイアログボックスは、ツールバーの「全文検索」ボタンをクリックして開く（図5）。複数の検索語を指定する場合、検索語を以下の符号で句切ること



図 5

とで検索条件を指定できる。

and 検索：半角スペース、+

or 検索：|

not 検索には対応していない。検索語は30語まで指定できる。また、「?」で任意の一字を指定することができる。

検索結果は、10件ごとにヒット箇所の原文の抜粋とともに表示される（図6）。ある語句の用例を収集す



図 6

るような場合は、この検索結果画面をコピーして集積することで効率的に用例・事例集を作成することができるだろう。ただ、検索結果が表組みで表示されているため手間かけないと Excel などにコピーしにくく、検索結果が 10 件ずつしか表示できないためヒット数が多ければ多いほど作業が繁雑になるのがいささか残念である。そうした使い方を考えた場合、台湾中央研究院漢籍電子文庫の「検索報表」のような、ヒット箇所一覧の CSV 出力機能を追加してもらえると嬉しいのだが。

異体字処理については、常用漢字体や Unicode に収録された僻字にも対応した一括検索テーブルを備えている。また、同音字の同一視検索にも対応しており、それらの ON・OFF も選択できる。

■ 固有名詞の検索

全文検索ダイアログで、検索語を入力してから「同義詞」ボタンをクリックすると、入力された語句の同義語が表示される。人名を入力した場合は、その人物の異称が表示される(図7)。これは、中華書局の正史の人名索引などをもとに作られた人名ソーラスに基づいており、中華書局製品ならではの機能であるといえよう。一覧表示された同義語にチェックを入れて検索すると、指定した語句を or 検索することができる。

中華經典典籍庫の従来のデータベースにない特色が、書名・固有名詞の指定検索である。検索ボックスの「検索範囲」で「書名」あるいは「専名」を

チェックして検索すると、検索範囲をそれぞれ書籍タイトル、人名・地名などの固有名詞だけに絞り込むことができる。



図 7

中華書局の標

点本では書名に波線、人名・地名に傍線が施されているが、それをマークアップして検索対象として指定できるようにしたものである。例えば「曹」を検索するとき、そのまま検索すると官職名の「曹」などに大量にヒットするが、「専名」をチェックしておくと「曹」を含む地名・人名のみを検索することができる(図8)。さらに、前述の同義詞検索と組み合わせることで、効率的かつ網羅的に固有名詞を検索することができる。

従来の大規模文庫データベースでは、異称を全て拾っている人名索引に比べて、人名の検索効率が落ちることが指摘されていた^[2]。中華經典典籍庫がその問題をかなりの程度改善していることは、特筆に値しよう。理想的には、人名・地名・書名のマークアップとあわせて ID を振っておけば、検索語にかかわらず全ての固有名詞出現箇所を検索できるようになるのだが、しかし 2 億字規模のデータベースにそこまで手間のかかる加工を求めるのは酷であり、傍線のマークアップというこの方法が現実的な最適解であるといえよう。



図 8



図9

このほか曖昧検索は、フレーズ・文などを入力して検索する機能で、あるフレーズと似た表現や、うろ覚えの語句や、誤りを含む可能性のある引用の出典などを検索する際に役にたつものと思われる。

■ 引用機能

文献データベースでは、必要なフレーズを見つけたら、それを論文・レポートなどにコピー＆ペーストして引用することになる。中華經典典籍庫には、その作業の効率を飛躍的に高める「引用」機能が実装された。本文の必要箇所をドラッグして選択した後で右クリックして「引用」を選ぶ(図9)。すると以下のように、引用箇所の書誌情報付きでペーストされる。

(魏氏春秋曰：朗字元明，新興人。獻帝傳曰：朗父名宜祿，爲呂布使詣袁術，術妻以漢宗室女。其前妻杜氏留下邳。布之被圍，關羽屢請於太祖，求以杜氏爲妻，太祖疑其有色，及城陷，太祖見之，乃自納之。)

〔晉〕陳壽撰〔南朝宋〕裴松之注：《三國志・卷三 魏書三 明帝紀第三》，中華書局，1982年7月，第二版，第100頁

従来の文献データベースでは、ヒットした箇所の巻数・章・節や篇名などがわかりにくく、時にページを遡って探す必要があったが、この「引用」機能ではそ

のような手間をかけることなく正確に必要な箇所を引用できる。他のデータベースにも同様の機能を追加してもらいたいものだ。

また他のデータベースでは検索結果を引用するにあたり、それこそ中華書局の標点本などの信頼性の高いテキストにあたって確認する必要があったが、中華經典典籍庫は収録文献の大半が学界で認められたテキストなので、他のテキストで確認する必要なくそのまま引用できる。その意味でも、ユーザビリティの高いデータベースに仕上がっていると評価できる。

◇ おわりに

以上、中華經典典籍庫の機能について検証してきたが、従来の文献データベースには無かった便利な機能を搭載しており、検索の精度やデータベースの利用効率がきわめて高くなっている。それは手間暇かかった質の高いテキストを底本とし、データベースの利用目的を考えて機能設計されているがゆえであり、期待通り、現状で最もクオリティの高い大規模古典文献データベースに仕上がっているといえよう。

中華經典典籍庫は現在発売中の第1輯に続き、今後とも続々と発売され、工具書・論著なども提供される予定であると聞く。それらとあわせて、今後、中国学の研究・教育に必須な定番ハイクオリティ・データベースとしての地位を確立していくと思われる。また、上海古籍出版社などライバル出版社にも、自社出版物のデータベース製品化を進めてもらいたいものだ。

我が国の大学・機関での導入にあたり問題になるのは、中国国内価格15万元、輸入価格約300万円という値段と、Windows Server対応のネットワーク版のみの提供でサーバ管理の必要がある、という点であろうか。これらの点については、我が国の輸入代理業者の努力とサービスに期待したい。

注

- [1] もっとも、戯曲・章回小説については、本データベースに収録するのにふさわしい校訂本を中華書局があまり出版していない、という事情があるのかもしれない。
- [2] 陳弱水「中央研究院歴史語言研究所漢籍全文自動化計画の発展、現状、未来」(野村英登訳、『漢字文献情報処理研究』第2号、2001年)

瀚堂典藏

千田 大介

❖ 後発大規模データベース

書同文・愛如生・中易中標・創新力博などの文献情報処理企業と、それら企業が開発した『四庫全書』・『中国基本古籍庫』などの大規模全文データベースが林立する中にあって、瀚堂典藏は後発のデータベースである。王宏源が北米留学から帰国した後、2004年に会社を立ち上げて構築を開始したもので、公式サイト、<http://www.hytung.cn/>の説明によると、2014年8月現在で登録済みデータは「古典籍 13,000 種余り、古典籍の書影 500 万枚、新聞雑誌の明瞭な画像 100 万枚、計 2100 項目のデータと画像を直接対応させており、漢字の合計文字数は 25 億を超える」という。

以下、2014年9月初旬に上海図書館において瀚堂典藏を^[1]、8月末に中国国家図書館において瀚堂近代報刊データベースを、それぞれ試用した結果に基づき、その機能・特色と学術研究における有効性について考えてみたい。

❖ 瀚堂典藏の特色

■ 文字コードとインターフェイス

瀚堂典藏の特色について、公式サイトなどでは「目下唯一の7万字以上の漢字のUnicode 4 バイト文字コードと自然言語全文検索、通用ブラウザモデルに基づく古典文献データベース」と説明される。確かに、専用クライアントソフトを必要とせずブラウザから利用できるのは、非常に便利である。もっとも、2000 年代半ば

においてこれらは確かに独自の特徴であったかもしれないが、現在では文淵閣『四庫全書』v.3・中華經典典籍庫など創新力博系の製品が Unicode の拡張領域に対応しているし、ブラウザによる検索は書同文の『四部叢刊』オンライン版なども実現しているので、個々の特色は決して「唯一」とは言えない。それらを兼ね備えているものは確かに瀚堂典藏のみ、ということにはなろうが。

瀚堂典藏の Web ページでは、旧世代 OS 向けに、SimSun・MingLiU の Ext.B 対応フォントを配布しているが、Windows Vista 以降はそれらのフォントを標準搭載しており、また Windows XP はサポート期限が終了しているので、我が国の大半のユーザーはフォント追加の必要がないものと思われる。

また、中華經典典籍庫が Ext.B に対応しながら数千字の外字を使用しているように、Unicode の拡張漢字を使ったとしても中国古典籍の全てをテキスト化できるわけではないが、瀚堂典藏は外字の不使用をうたっている。具体的に Unicode Ext.B 未収録文字がどのように処理されているのかについては、未確認である。

瀚堂典藏は図 1 のようなインターフェイスを持つ。主ペインの上部に検索ボックスが置かれ、ペインには文献を四部分類したツリーが表示される。この分類をクリックすることで展開できるし(図 2)、また分類や



図 1

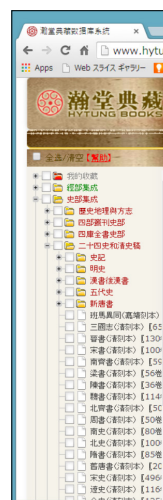


図 2

文献名の前のチェックボックスをクリックして検索範囲を指定することもできる。

■ 収録文献

瀚堂典藏には収録文献の一覧ページが見あたらず、データベースの紹介ページにアウトラインが説明されているだけである。それによると、四部の各集成庫に登録される代表的文献には、それぞれ以下のようなものがある。

- 経部：十三經・十三經注疏・皇清經解
- 史部：二十四史・清史稿・明實錄・清實錄・史料筆記
- 子部：佩文韻府・古今圖書集成・太平御覽・初學記・冊府元龜・永樂大典（殘卷）・翰苑新書・北堂書鈔・本草綱・龍龕手鏡・一切經音義・続一切經音義・可洪音義・丁福保『仏教大詞典』・正統道藏・統道藏
- 集部：全唐詩・全唐文・全宋詞・皇明經世文摘・晚晴移詩匯

このほか、各部に個別の文献が相当数登録されており、たとえば子部には漢方医学関連書籍が、集部には別集などが、それぞれ相当量収録されている。また、四部とは別立ての專題資料庫には、台湾文献・漢訳世界名著・出土文献・敦煌文献などが収録されている。

また、四部分類ごとに『四庫全書』・『四部叢刊』の当該分類が登録されている。このため、総合文字数25億字をうたっているものの、異本が重複登録されている文献も相当数にのぼるものと思われる。

収録文献であるが、ごく大まかに確認したところでは、底本に叢書に収録された版本を使っているものが大半であった。おそらく、著作権などの権利問題の発生を回避するために標点本を避けているのであろう。画像および説明を見る限り、それらも原本を直接撮影したのではなく、影印本に基づくものが多いようである。

筆者の専門である集部の戯曲を見たところ、収録文献が少なくあまりに不完全であった。例えば、古本戯曲叢刊の初集～五集というカテゴリが立っているが、四集の雜劇がそれなりに登録されているものの、他はいずれも1～2種の伝奇が登録されているだけだった。瀚堂典藏はユーザーのニーズに応じて入力すると

うたっているが、全体のバランスを見て代表的な総集・選集や叢書などを揃えることも必要であろう。今後の改善が待たれる。

■ 検索と閲覧

瀚堂典藏では、主ペインのテキストボックスに字句を入力して検索することになる（図3）。

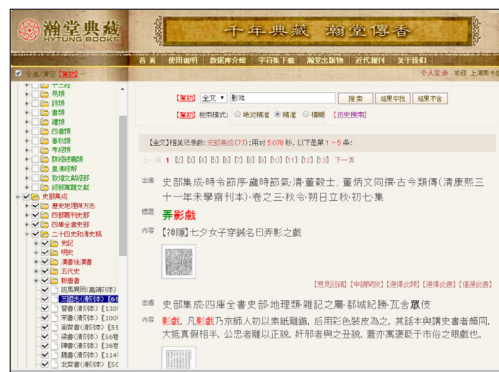


図3

検索結果は、1ページあたり5件で30ページまで、すなわち150件しか表示できない。このため、ヒット数が多い場合は、検索範囲を絞り込むか、二次検索を行う必要がある。データベースの用途が語句や事例の調査に絞られており、語彙の抽出・分析などが考慮されていないことがわかる。

検索時に「模糊」をチェックしておく、異体字を一括検索することができる。試みに「總」とその簡体字・異体字で史部を対象に検索したところ、ヒット数は下表のようになった。

検索文字	ヒット数
總	139,183
總	140,792
总	140,792
總	139,215

いずれもヒット数が14万前後であることから、異体字の一括検索が機能していることがわかる。特に、簡体字と繁体字は全く同じ結果になっている。一方、異体字の場合は、なぜか結果に微妙な差異が見られる。

検索結果の版本サムネイル画像をクリックすると、当該箇所本文が表示される（図4）。

左ペインに当該ページの電子テキスト、主ペインに

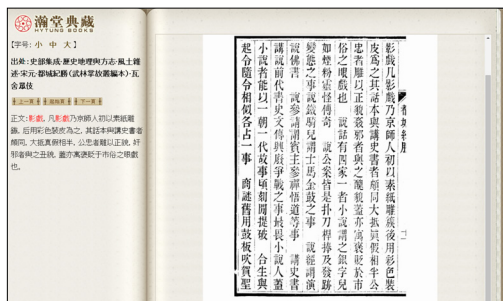


図 4

書影が表示される。左ペインにはこのほか版本情報・巻数・章節などが表示され、また「起始頁」をクリックして検索結果一覧ページに戻ることもできる。書影は JPEG 画像で、保存することもできる。

書籍を閲覧する場合は、検索対象を「書目」に設定して検索してもよいし、左ペインの分類をたどってもよい。書籍によっては、まず当該書籍の紹介と目次が表示される。

❖ 瀚堂近代報刊データベース

中華人民共和国成立前の新聞・雑誌史料を収録するデータベースは、「瀚堂近代報刊データベース」のタイトルで別立てになっている。URL は、<http://www.neohytung.com/> になる。

インターフェイス (図 4) は、瀚堂典藏と大差ないが、分類が地域や時代ではなく、刊行物の名称のピンイン順になっている。



図 5

サイトの解説は収録する代表的定期刊行物として、以下を掲げている。

- 香港『遐邇珍珍』

- 北京『順天時報』
- 天津『大公報』・『益世報』・『北洋画報』
- 上海『上海新報』・『申報』

実際に収録されている資料は以上にとどまらない。例えば左ペインの D をクリックすると、図 6 のような D で始まる新聞・雑誌が収録されていることがわかる。



図 6

検索や閲覧の方法は瀚堂典藏と同じで、左ペインおツリーで検索範囲を指定し、主ペインの検索ボックスに語句を入力して検索し、結果のサムネイルをクリックして影印を表示する (図 7・8)。



図 7



図 8

清末から民国時期にかけての新聞・雑誌資料のデータベースとしては愛如生の『申報』データベースなどがあるが、その『申報』をも収録している瀚堂近代報刊は、同種のデータベースとして現時点最大規模であると思われ、近現代の文化・社会などを研究する上で大いに役立つことだろう。

❖ おわりに

以上のように瀚堂典藏・瀚堂近代報刊は、収録文献の量が膨大であり、しかも他のデータベースには収録されていないコンテンツをも擁している。

また、瀚堂典藏の収録文献には、底本の多くが叢書などの版本であるという特色があるが、これは、中国基本古籍庫全文検索の「通行本」、抄本の大規模叢書である『四庫全書』、中華書局の標点本を集めた中華經典典籍庫とも異なっており、『叢書集成』的な趣を持つ古典文献全文データベースであるといえよう。

一方、検索機能は複雑な条件設定に対応していない上に検索結果の表示件数が限られるなど、中国の他の大規模データベースと比べても貧弱であるといわざるをえない。このためデータベースの用途が、出典・事

例の発見のみに限られてしまい、語句の用例を大量に集めるといった使い方ができない。ただ、Web ベースのデータベースであるので、収録文献や検索システムなど、今後さらなる発展や改善が期待できる。

気になるデータベース導入のコストであるが、公式サイト上には価格に関する情報が見あたらないし、日本からの個人登録を試みたが、反応は得られなかった。Google で検索すると6 万元・8 万9 千元といった数字が出てくるが、もしも10 万元弱で導入できるのであれば、中国基本古籍庫の10 分の1 程度ということになり、データベースの規模に対して、比較的リーズナブルであるといえよう。今後の日本国内での正式発売が待たれる。

注

- [1] 上海図書館では、持ち込み PC を WiFi 接続し、同図書館サイトのトップページから「e 卡通」をクリックして認証することで、各種データベースを利用することができる。また閲覧証は、身分証明書がパスポートであっても自動で作成することができる。

蘇州地方文献データベース

佐藤 仁史

❖ はじめに

中国方志庫や中国基本古籍庫などの登場によって、民国期までの地方志や古籍の利用の便が飛躍的に向上したことは贅言を要しまい。しかしながら、これらの商用データベースは非常に高価であるため、購入できた大学・研究機関は非常に限られており、地域によっては近隣に購入した大学がないところもある。このような状況に置かれた研究者や学生は国内外の所蔵機関に見に行くほか術はない。朗報なのは、地域によっては現地の図書館が独自に構築したデータベースが公開されていることであり、具体的な地域に即した研究には特に裨益するところが多い。そのような事例として以下では蘇州図書館と呉江図書館の取り組みを紹介する。

❖ 蘇州図書館古籍庫

http://fzk.szlib.com/AncientBook/Main/Ancient_Index.html

本データベースは蘇州図書館が構築した蘇州の古籍に関するデータベースである。「古籍庫」と銘打ってはいるものの、多くは各種の地方志が占めている。収録されている文献の内訳は、府県志、郷鎮志、人物志、園林名志、館藏珍貴古籍、其他となっている。検索機能には、「書名」「作者」「年代」「簡介」「全文」があり、目的に応じて検索機能を適宜使い分けることができるようになっている（図1参照）。ここでは、筆者が実際に現在調べている事柄に即して検索機能を使ってみよう。

筆者は数年来近代における蘇州周辺農村の民間信仰について調べている。現地調査の中で、改革開放政策が始まって以来廟会が「復興」しており、そのいくつかについては実際に参観する機会を得た。蘇州市東南部にある上方山の廟会もそのうちの1つである。上方山には五通神廟があり、そこには五通神が祀

られてきた。これを淫祠とした江蘇巡撫湯斌が康熙24年（1685年）に神像を破棄した経緯があることで有名である。後にここに祀られた上方山老爺と上方山姆姆が江南一帯で広く信仰を集め、旧暦8月18日に開催される廟会は近隣一帯から多くの参拝客で賑わう。

この廟の実態や変遷について調べるため、試みに「上方山」というキーワードを入力し、全文検索機能を用いて検索してみたところ、67種類の文献に「上方山」を含む用例があることが示された。最も多い検索結果が出たのは同治『蘇州府志』（19例）、次いで道光『蘇州府志』（17例）、民国『吳県志』（14例）であり、比較的新しい時期には同一の文献に用例が多いことがわかる。時期で見た場合、乾隆『蘇州府志』（4例）などがあり、長いスパンでも追跡が可能となっている。府県志以外でもいくつかの郷鎮志において用例が見られ、例えば蘇州中心部にほど近い木瀆鎮の郷鎮志『木瀆小志』に3例あった。また、本データベースの特徴は、地方志とその他の文献とを分割せず、古籍を広く包括している点である。いくつかの文集からの用例もえられ、面白いところでは『吳門竹枝詞』（5例）などがあつた。

地方志利用の便からみれば、1949年までに編纂・刊行された旧地方志に加え、1980年代以降に陸続と刊行されたいわゆる新編地方志との横断検索ができればより便利であろう。蘇州の新編地方志は蘇州市地方



図1 蘇州図書館古籍庫のトップ画面

志弁公室の Web ページにおいて既に大多数をブラウズすることができるようになっている^[1]。データベースを構築している機関は異なるものの、これら 2 種のデータベースが連携すれば、相乗効果によってより多くの知見をもたらすことができるであろう。

本データベースの使用には蘇州図書館において閲覧証を申請し、ID とパスワードを取得する必要がある。館内では全文データを、館外からは用例の一部を閲覧することができる。閲覧証がない場合には館外から全文を閲覧することはできないが、関連キーワードが含まれている史料の書誌情報を知ることができるので、関連史料に「あたり」をつけておき、原文を図書館で探すという使い方もできる。

❖ 吳江古代地方志データベース

http://www.wjlib.com/show_news.asp?id=1453

県レベルの図書館においても地方文献のデジタル化が積極的に進められている。吳江図書館の「吳江古代地方志データベース」もそのような取り組みの顕著な例である。

ここに含まれているコンテンツは 2 つに分かれ、1 つは明代から民国期の吳江、震沢両県の県志 10 種であり、もう 1 つは両県下で編纂された郷鎮志など 14 種である。後者には郷土教科書である民国『同里郷土志』が含まれている点がユニークである。ここに収録された地方志は蘇州図書館古籍庫収録の地方志と重複が見られるが、両県の地方志をほぼ網羅しており、当該地域について調べる場合には非常に便利である。ただ、郷鎮志では『同里郷土志』を収録しているにもかかわらず、同県で編纂された『吳江県郷土志』（范烟橋、1917 年）を収録していないのは不思議なことである。これを収録すればより網羅的なデータベースとなったことであろう。

使用には 2 点注意する必要がある。1 つは閲覧に Rtbooks というプラグインをインストールする必要があることである。もう 1 つは基本的にこのデータベースは館内での閲覧が前提となっている点である。

館内でのみの閲覧ではあまりにも不便であるが、本データベースはデジタル化の完成に合わせて DVD 版『縦覧吳江——吳江 500 年古

代地方志匯編』江蘇電子音像出版社、2006 年）も比較的廉価で販売されており、手元に置いておきたい研究者は DVD 版を入手することもできる（図 2 参照）。とはいえ、オンライン上で広く公開されることが望まれる。

❖ おわりに

地方志をはじめとする地方文献は地域史研究のみならず、様々な領域においても有用な史料群であるものの、存在が知られていないことによって必ずしも有効に活用されてこなかった。商業データベースとは異なる関連機関ごとの地道な作業が集積していくことによって利用が広がり、様々な価値が見いだされることを期待したい。

注

- [1] http://www.dfzb.suzhou.gov.cn/zsbl/index_1.htm

なお、上海市地方志弁公室の web ページでは上海市下で編纂された新編地方志が公開されており、全文検索をすることもできる。<http://www.shtong.gov.cn/>

目次	全文検索	書名	著者	巻数
1	吳江縣志	吳江縣志	吳江縣志	1
2	震澤縣志	震澤縣志	震澤縣志	2
3	震澤縣志	震澤縣志	震澤縣志	3
4	震澤縣志	震澤縣志	震澤縣志	4
5	震澤縣志	震澤縣志	震澤縣志	5
6	震澤縣志	震澤縣志	震澤縣志	6
7	震澤縣志	震澤縣志	震澤縣志	7
8	震澤縣志	震澤縣志	震澤縣志	8
9	震澤縣志	震澤縣志	震澤縣志	9
10	震澤縣志	震澤縣志	震澤縣志	10
11	震澤縣志	震澤縣志	震澤縣志	11
12	震澤縣志	震澤縣志	震澤縣志	12
13	震澤縣志	震澤縣志	震澤縣志	13
14	震澤縣志	震澤縣志	震澤縣志	14
15	震澤縣志	震澤縣志	震澤縣志	15
16	震澤縣志	震澤縣志	震澤縣志	16
17	震澤縣志	震澤縣志	震澤縣志	17
18	震澤縣志	震澤縣志	震澤縣志	18
19	震澤縣志	震澤縣志	震澤縣志	19
20	震澤縣志	震澤縣志	震澤縣志	20
21	震澤縣志	震澤縣志	震澤縣志	21
22	震澤縣志	震澤縣志	震澤縣志	22
23	震澤縣志	震澤縣志	震澤縣志	23
24	震澤縣志	震澤縣志	震澤縣志	24

図2 『縦覧吳江——吳江 500 年古代地方志匯編』のトップ画面

華東師範大学中国文字研究与应用中心

山田 崇仁

❖ はじめに

華東師範大学中国文字研究与应用中心（以下、センターと略す）（<http://www.wenzi.cn/web/Default.aspx>）は、中華人民共和国教育部が認可する人文社会科学重点研究基地の一つとして、2000年に設置された。その目的として、「情報科学の研究手法を導入したコーパス文字学体系の確立」が唱われており、具体的な項目として古文字学・今文字学・少数民族文字学・域外漢字学、漢字認知科学および完全な漢字伝播システムが計画されている（以下、URLはすべて2014年8月9日の確認）。

その成果物については、これまでも書籍媒体やアプリケーションソフトなどの形で公開されている。筆者もこれらセンターの成果物について、本誌第6号（2005年）にて「商周金文数字化処理系統 戦国楚文字数字化処理系統」と題したレビューを行っている。

今回は、センターが設置するWebサイトで利用可能な、いくつかのデータベースについて採り上げることとする。

■ 利用にあたっての諸注意

まず、利用にあたって注意したい点が三つある。

一点目は、センターが公開するデータベースについては、基本的にInternet Explorerでの閲覧を前提としており、それ以外のWWWブラウザを利用した場合は、データベース接続ができなかったり、一部の情報が閲覧できなかったりする可能性があることである。

二点目は、データベース関連は基本的に無料で利用できるが、それ以外のダウンロードサービスなどの中に、会員登録が必要なものがあることである（会員登録は無料で可能）。

三点目は、ページ作りのずさんさに起因するのか、データベースエラーが表示される場合がある、また閲覧に必要な外字ファイルがダウンロードできない場合があることである。

❖ 数字化資源

センターの公開するデータベースは、サイトの「数字化資源」の項目に登録されている。その中には更にいくつかの細目に分かれて複数のデータベースが存在する。以下、項目毎に内容と簡単な使用方法と特徴を紹介する。

データベースの中には、甲骨文や金文といった現在の漢字では既に使われていないものも使用されるため、センターが独自に作成した外字フォントをダウンロードしてインストールを必要とするものがある。具体的な方法については、各データベースのWebページに書かれた指示に従うこと。

■ 関連書系

この項目では、本センターの活動の一環として編纂された書籍や、関連する書籍の索引データベース



図1 華東師範大学中国文字研究与应用中心トップページ

などが登録されている。

●《古文字考釋提要総覧》(第一冊) 関連検索

<http://www.wenzi.cn/web/content.aspx?moduleid=21&parentid=20>

2008年に上海人民出版社から発行された同名書籍のデータベース。当該書は、古文字に関する専門家の学説を整理したものである(『説文解字』～2007年9月末まで)。当該書にも見出し索引は収録されているが、本データベースでは、見出し字以外に文字構造や研究者、収録書名・誌名などを指定して検索することもできる。その点では、当該書をより便利に使いこなすことができるはずなのだが、実際には検索、特に文字入力の入力インターフェイスの作りにやや癖がある。

例えば、見出し字の入力は中国のGBKコードの範囲で繁体字入力をする必要がある。GBK未収録字については、アルファベットを用いた独自の漢字構造記述式が用意されているが、その記述方法に慣れないと検索が難しい。これについては、GBK未収録字と記述式情報の一覧をPDFファイルなどで公開した方が利用者にとっては便利ではないだろうか。

また、実際には存在するはずの文字を入力しても結果が表示されないことがあった。これはひとえに筆者の利用方法の理解不足に起因するものなのだろうが、なぜそれが起きるのかがよくわからないという点も、インターフェイスの使いにくさに一因があるのかもしれない。

さらに、閲覧には外字の利用が必須となるが、肝心の外字ファイルへのリンクが正しく設定されておらず、ダウンロードすることができない。一応、他のデータベースからリンクされている外字を適用することが可能なので(例えば甲骨文ならば、後述の花園荘東地甲骨検索系のページから外字ファイルをダウンロードできる)、そちらの外字ファイルを使えばよい^[1]。

关键字	考釋字	文字类型	作者	篇名/题目	书名/刊名	出版时间	页码
18046	𠂔	甲骨文	白川静著、鄭清茂譯	作考	中國文字	1971年	307
19375	𠂔	金文	白川静	西周彝器斷代小記	“中央研究院”歷史語言研究所集刊	1963年	538
19446	𠂔	金文	白川静著、鄭清茂譯	作考	中國文字	1971年	351

図2 《古文字考釋提要総覧》(第一冊) 関連検索にて、[考釈者]の項目に「白川静」を指定して検索した結果

●《中国異体字大系一篆書編》 関連検索之原始材料検索

http://www.wenzi.cn/pages/guanlianshuxi/YT_Index.HTM

●《中国異体字大系一篆書編》 関連検索之考釈材料検索

<http://www.wenzi.cn/daxikaoshi/index.asp>

2007年に上海書畫出版社から発行された同名書籍のデータベース。本書は、篆書の実際の用例を文字毎に一覧表示する形式で編纂されたものだが、本データベースはそれを作成する際に資料となった実物の用例や論文などを検索するものである。

ただし本データベースは、検索のキーワードとして当該書のページ数・行(段)数・文字数などを指定する必要があるため(実際の文字をキーワードとして指定できない)、当該書を利用するための限定サービスに位置づけられていると思われる。

●《中国異体字大系一楷書編》 索引

<http://www.wenzi.cn/kaishubian/index.asp>

●《中国異体字大系一隸書編》 索引

<http://www.wenzi.cn/lishubian/index.asp>

こちらも上記と同じシリーズの書籍の索引データベース。ただしこちらは、文字の検索に対応したものである。文字ならびに筆画での検索が可能。検索結果で得られる情報は、部首・読音・筆画・書籍での表示位置である。こちらも当該書利用者のための限定サービスと思われる。

●《新金文編》 対応銘文及拓片関連検索

<http://www.wenzi.cn/jinwenbiannew/index.asp>

作家出版社から2011年に発行された同名書籍のデータベース。検索キーワードとして指定可能なのは、『殷周金文集成』(1984-1990 中華書局/修訂増補版 2007)の青銅器の整理番号のみである。検索結果としては、銘文全体や断代、器種、国別などの情報が表示される。

また検索結果から一応拓本画像へのリンクが設定されているが、権利の関係で実際に表示されるのは一部分に留まっている。使いようによっては、論文などで引用されている『殷周金文集成』の番号から銘文を調査する場合に役立つかも知れないが、『新金文編』のデータベースとして有用かといえ、今の実装状況では否とせざるを得ない。

本誌第6号のレビューを執筆したときにも感じたのだが、センターの成果物の中には未実装のものをいかにも実装されているような形でリリースするケースが多い。そのため、羊頭を掲げて狗肉を売る状態のものが散見される結果になってしまう。本データベースもその一つとならないように祈っている。

- 《中国出土簡帛文献引得総録・郭店楚簡巻》情報検索—《郭店楚簡電子誌林》

<http://www.wenzi.cn/guodianyinde/index.asp>

- 《中国出土簡帛文献引得総録・郭店楚簡巻》情報検索—《郭店楚簡電子類書》

<http://www.wenzi.cn/guodianyinde/indexlei.asp>

2012年に上海人民出版社から発行された同名書籍の索引。こちらも当該書利用者のための限定サービスと思われる。

後者に関しては、項目別に絞り込むことで検索結果を得ることができる。

- 《中国漢字文物大系》関連検索

<http://www.wenzi.cn/web/content.aspx?moduleid=21&parentid=20>

2013年に大象出版社から発行された同名書籍の索引データベース。キーワードは、文字・ピンイン・筆画を指定可能。本書は甲骨・青銅器・印璽・古陶・貨幣・簡帛・石刻・紙といった漢字が実際に使用された実例を、文字毎に整理して収録したものである。本書はその目録検索をオンラインで提供する形となるが、このような大部(15冊)の書物を検索する際には、このような検字タイプのデータベースの存在はありがたい。

■ 古文字資源庫

ここでは、統一秦以前に使われていた文字に関連するデータベースを登録している。

- 商周金文検索系統

<http://www.wenzi.cn/pages/jwyd.asp>

センターの成果物である上記『商周金文数字化処理系統』(2003年発売)を作成した際に使用した資料から、高頻度で使用される10字の用例を検索可能にしたもの。本データベースの紹介文を読む限り、将来的には『商周金文数字化処理系統』発売後に明らかとなった新出金文を含めた総合金文データベース

スとしてバージョンアップしたいようなのだが、現状ではコンテンツの内容が題目に追いついていない。

- 戦国楚文字検索系統

<http://www.wenzi.cn/pages/czjs.asp>

こちらも上記『戦国楚文字数字化処理系統』(2003年発売)をベースに開発された、戦国楚簡を検索対象としたデータベース。上記『戦国楚文字数字化処理系統』発売以降に出土した新出楚簡(『上海博物館蔵戦国楚竹書』3~5・『新蔡葛陵楚簡』)も収録されている。

本データベースでは、キーワードに文字を指定することができる。検索結果としては、著録書名・篇名・簡番号・実際の字形がそれぞれ表示される。この手の

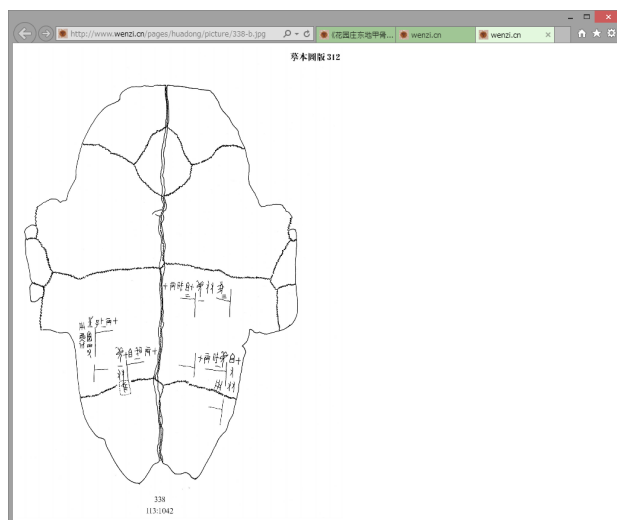


図3 花園莊東地甲骨検索系統で「上」をキーワードに検索した結果(上画面)と甲骨の模本画面を表示したもの(下画面)

文字篇のデータベースとしては、基本的な機能を押さえているが、頻用字であるはずの用例を入力しても結果が表示されない、あるいは一部の用例しか表示されないなど問題が多い。

●花園莊東地甲骨検索系統

<http://www.wenzi.cn/pages/huadong/index>.

HTM

雲南人民出版社より 2003 年に出版された『花園莊東地甲骨』のデータベース。甲骨文の釈文を対象に、単字を指定して検索するシステムを持つ。検索結果はキーワード・条号・篇号・拓本・形態別に閲覧できる。篇号では原拓本と摹本へのリンクが設定されており、それぞれのリンクをクリックすると、『花園莊東地甲骨』の当該ページが表示される(何故か表示されないものもある)。書籍の掲載の許諾など、どのような権利処理をしているのか不明ではあるが、出土文字資料では釈文と写真(拓本)を見比べることが研究上必要な作業であるため、利用者にとっては便利。

●《古文字詁林》電子検索

[http://www.wenzi.cn/gulin0715/index](http://www.wenzi.cn/gulin0715/index.asp).

asp

上海教育出版社より 2011 年に出版された『古文字詁林』のデータベース。本書は、古文字の実例や解釈についての事例を網羅した辞典である。無論、研究に際してはそこから典拠に当たって再調査する必要があるが、当該字についてどのような学説が展開されたのかわかる上では重宝する書籍である。

キーワードとしては、部首・親字・ピンイン・反切・作者・題目・典拠を指定することができる。内容としては、整理番号・部首・見出し字・《説文解字》の巻数・ピンイン・大徐本の反切・(本書の)冊数・ページ・(言説の)作者・論題・出典・出版時期ならびに本書の画像へのリンクで構成される。こちらもどのような許諾や権利処理をしているのか不明なので、利用に際しては注意を必要とする。

●《金文文献集成》電子検索

<http://www.wenzi.cn/jinwenwenxianjicheng/index.HTM>

2005 年に線装書局から出版された同名書籍の作者索引データベース。当該書は金文に

関する論文や単行本を収録した一大叢書であり、このような索引を利用できることは非常にありがたい。キーワードは姓名のそれぞれ一部でも検索できる。また、検索結果は作者・題目・出典・出典情報・集成の収録冊数とページ数がそれぞれ表示される。ただし、典拠情報が曖昧なものもあるため、他の書籍・論文データベースを併用するべきである。

●常用古文字字形検索

<http://www.wenzi.cn/guwenzizixingjiansuo/guwenzizixingjiansuo.HTM>

現在我々が使用している文字が、古文字でどのよう

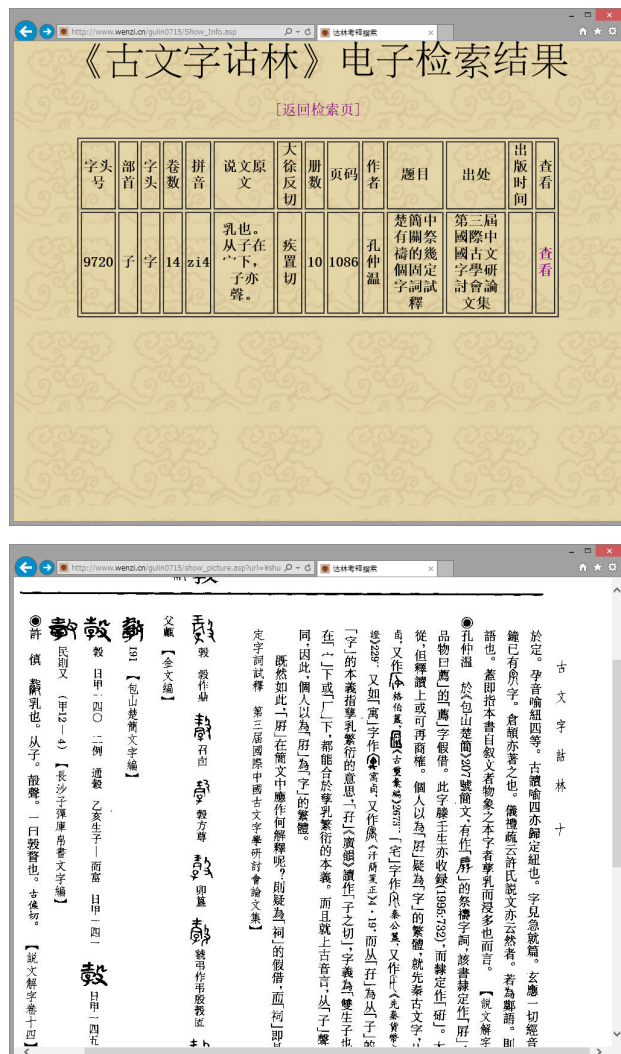


図4 《古文字詁林》電子検索で「字頭(見出し字)」の項目に「字」を指定して検索した結果(上画面)と、冊子の当該部分を表示したもの(下画面)

な字形であったかを検索・表示するデータベース。

キーワードを入力して検索を実行すると、古文字の字形と甲骨文や金文などの類別が併せて表示される。字形の変遷を見るのに便利ではあるが、典拠情報が一切表示されないのが難点。

■ 今文字資料庫

この項目については、執筆時（2014年8月9日）現在、公開されていない。

■ 民族文字資源庫

ここでは漢字以外の文字についてのデータベースを収録する。

ただし筆者の環境では、説明に提示されている用例を入力しても、いずれもデータベースエラーが返ってきてしまい、検索できなかった^[2]。

● 東巴文検索系統

<http://www.wenzi.cn/minzuwen/dongba/index.asp>

ナシ族が使用するトンパ文字を対象としたデータベース。漢字の字義や類別、所謂六書による分類別検索ができる。

● 古壮文検索系統

<http://www.wenzi.cn/minzuwen/zhuang/index.asp>

チワン族が使用する文字を対象としたデータベース。漢字の字義や類別を指定しての検索ができる。

● 水文検索系統

<http://www.wenzi.cn/minzuwen/shui/index.asp>

スイ族が使用する文字を対象としたデータベース。漢字の字義や類別を指定しての検索ができる。

● 瑪麗瑪莎文検索系統

<http://www.wenzi.cn/pages/smwz.asp>

トンパ文字から派生したとされるマリマサ文字のデータベース。

● 傣僇文検索系統

<http://www.wenzi.cn/minzuwen/lishu/index.asp>

リス族の使用文字のうち、1920年に汪忍波によって考案されたリス文字を対象としたデータベース。漢字の字義や類別を指定しての検索ができる。

● 中国民族古文字字表

<http://www.wenzi.cn/minzuwen/15guwenzi.pdf>

● 中国民族文字術義対照詞典（様例）

<http://www.wenzi.cn/minzuwen/tongyiduizhao.pdf>

それぞれのタイトルを持つ文書をPDFで出力したものが表示される。

■ 歴代辞書資源庫

● 《説文解字》検索五種

<http://www.wenzi.cn/shuowen/index.htm>

中華書局《説文解字》、上海古籍出版社《説文解字注》、中華書局《説文解字義証》、中華書局《説文解字詁林》、上海教育出版社《古文字詁林》の文字を検索することができるのだが、筆者の環境ではこちらも民族文字資源庫と同様に上手く動かなかった。

■ 域外漢字資源庫

● 韓国学者漢字研究資源

<http://www.wenzi.cn/yuwai/hanguoyanjiu/O8he.asp>

● 中国学者韓国漢字研究情報庫

<http://www.wenzi.cn/yuwai/chinayanjiu/O8hechina.asp>

いずれも、大韓民国の研究者による漢字研究の成果を一覧表示したもの。域外と銘打ちながらも、実際には韓国のWebサイトのみが挙げられており、この分野の研究実績が少なからず存在する日本や台湾・香港が全く挙げられていない。本項目の一層の充実に期待したいが、我が国の側からも韓国のような積極的な対外交渉・宣伝活動をすべきだろう。

● 漢字認知習得資源庫

児童が漢字学習をするのに便利なサイトが二つほど紹介されている。

※ おわりに

以上、センターが公開する各種データベースを紹介した。この記事を担当することになった当初は、どれだけ役立つものが含まれているのか楽しみに調べ始めたのだが、実際に使用してみるうちに、みるみる初めの高揚が失われることとなった。

それは各データベースの題目から期待される内容と実際に検索可能な内容との乖離が甚だしいこと、簡体字中国語圏以外の検索に不便なこと、インターフェイスに癖があることなどが挙げられる。ただ、データベ

中国・台湾の古典籍検索

小島 浩之

❖ はじめに

漢籍目録や漢籍善本画像のデータベース化は、わが国の東京大学東洋文化研究所と京都大学人文科学研究所が中国に先行してきた^[1]。

しかし、近年は中国・台湾の図書館でも古典籍の各種データベースを公開する動きが加速してきている。その全体像を余すところなく確認することはとうていできないが、主要なものを中心にレビューを試みたい。

❖ 中国の古典籍検索

本稿では、中国で公開されている古典籍 DB を、公開内容や公開方法から以下の 4 分類 5 種に分けてみていこうと思う。

(1) 目録検索型

① OPAC 組み込み系

② 独立検索系

(2) 目録検索＋書影公開型

(3) 電子図書館型

(4) その他

ところで、中国語で日本語の古典籍に該当する言葉は古籍である。古籍の中でも特に質の良い重要なものを善本と呼び、転じて貴重書の意味にも使われる。ただし、現在の中国の図書館では古籍の中に、古典を活字化したもの（排印本や点校本）や、複製したもの（影印本や影照本）などを含んでいることもある。後述する CALIS や CADAL のように古籍の定義を明示しているところは僅かで、その範囲は様々である。

■ 目録検索型

目録検索型とは、古典籍の書誌・所蔵データのみを公開しているものである。このタイプはさらに、OPAC の検索メニューの一つとして古典籍が含まれているものと、OPAC とは別に独立した検索システムを構築しているものに二分される。

① OPAC 組み込み系

●首都图书馆馆藏书目

<http://www.clcn.net.cn/>

首都図書館（北京市立図書館）の OPAC（馆藏书目）^[2] は、Web サイトトップページに検索ボックスが設置されている。古籍に限定して検索する場合は、検索ボックス下部のオプションボタンから、「古籍文献」を選択する。



図1 首都図書館トップページ



図2 検索結果の詳細

検索すると検索結果一覧が別ウィンドウで立ち上がる。一覧から必要なものを選択して「詳細内容」をクリックすると、簡略な書誌情報と馆藏信息（所蔵情報）が示される。さらに詳細な書誌・所蔵データは「更多信息」をクリックすることで表示される。

興味深いのは、詳細書誌データ下部に「载体形态」という項目があることである。载体形态とは、媒体の型式、フォーマットというようなニュアンスであろう。ここでは图片（写真・図・拓本などの）資料・视频资料（映像資料）・PDF 文本・普通文本の四種に区分され、それ

ぞれにリンクが張ってある。このリンクは現状では生きていないが、近い将来、デジタル化された資料をWeb配信できるよう設計されているのかもしれない。

ところで、現代の目録法には「転記の原則」という考え方があり、タイトルや著者名は本のタイトルページや奥付に記されているままに記録するよう定められている。中国の図書目録もこういった近代的な目録法が浸透しつつあり、書名が簡体字であれば目録も簡体字、逆に書名が繁体字であれば目録も繁体字となっている。このため図書館のOPACは簡体字・繁体字の異体字テーブルを作成し、どちらの検索も可能になるよう工夫されている場合が多い。

だが、首都図書館では、「漢書」「汉书」の双方で検索してみると、ヒット数に差が生じてしまう。簡体字で検索しても繁体字の書誌がヒットするので、何らかの異体字処理はなされているようだが、管見の限りではヒット数に差が生じる理由はわからなかった。

●武漢地区公共図書館聯合目録

<http://219.140.69.151/opac/index/advance>



図3 武漢地区公共図書館聯合目録トップページ

武漢地区公共図書館聯合目録は、湖北省武漢市の17図書館の蔵書をまとめて検索できる総合目録である。やはり、検索ボックスの下部にオプションボタンがあるので、「古籍」を選んで検索する。

一図書館の所蔵だけを検索できる目録を個別目録というのに対し、複数の図書館の所蔵を一度に検索できる目録のことを総合目録という。前述の首都図書館OPACが、個別目録の例であるのに対し、武漢市のものは総合目録の代表例である。

●湖南図書館 OPAC

<http://opac.library.hn.cn/opac/>

湖南図書館のOPACでは、「高級検索」(詳細検索)画面から、古籍に限定した検索ができる。高級検索画

面へは、簡易検索画面の検索ボタン右にある「高級検索」をクリックすればよい。

湖南図書館OPACは武漢地区公共図書館聯合目録と同一システムであり、高級検索画面は、武漢地区公共図書館聯合目録の検索画面と同一のインターフェイスとなっている。

●广东省立中山图书馆目录检索特藏文献

<http://opac.zslib.com.cn:8991/>

[F/?func=file&file_name=find-b-tcwx](http://opac.zslib.com.cn:8991/?func=file&file_name=find-b-tcwx)

广东省立中山図書館の古籍検索は、一見するとOPACから独立しているように見える。しかし、古籍検索は中国文献全体の特殊コレクション(特藏文献)中に分類されており、OPAC検索メニューの一つである。



図4 广东省立中山図書館目録検索特藏文献

通常のOPAC画面(<http://opac.zslib.com.cn:8991/>)では古籍に絞り込んだ検索はできないので、検索ボックス下の「査特藏文献」をクリックすると、当該検索画面が立ち上がる(図4)。特藏文献のメニューの中には、「古籍善本」・「线装书库」・「影印古籍」・「古籍丛书」といった下位カテゴリがあるので、必要なものにチェックを入れれば、OPACを古籍に特化して検索することができる。

②独立検索系

●上海图书馆古籍书目查询



図5 上海図書館古籍書目查詢検索結果一覧

<http://search.library.sh.cn/guji/>

上海図書館所蔵の古籍（活字本や影印本などを含む）129,660 件の目録データ。検索は、題名・著者名・分類・叢書 [名] から検索したい項目を選択し、検索ボックスにキーワードを入力して検索ボタンをクリックすればよい。

データは繁体字に正規化されているので、簡体字では検索できない。

● 辽宁省图书馆古籍目录

<http://lnlib.com/gj/index.htm>

遼寧省図書館のトップページ (<http://lnlib.com/>) からの場合は、上部のメインメニュー>书目检索>古籍目录の順にたどる。



図6 遼寧省図書館古籍目録

伝統的な漢籍目録の手法による目録で、これまでに述べた図書館系の目録とは一線を画している。漢籍目録に慣れていたれば、現代風の構造化された書誌よりこちらの方が解りやすいかもしれない。

経部 2,297 件、史部 9,322 件、子部 5,158 件、集部 7,208 件、叢書部 732 件を著録しており、データは全て繁体字である。

検索は、キーワード検索と階層検索が用意されている。階層検索は四部分類が、「部—類—属」のツリー構造で表示され、属ごと（一部は類ごと）に一覧表示できる仕様になっている。

● 南京市古籍普查联合目录

<http://www.1jlib.cn/lib/lhml/gj>

南京市公共図書館の古典籍総合目録で、収録範囲は民国以前の版本・鈔本に限定されているようである（サイトには明示されていないが、筆者が複数回の検索を試した限りでは、点校本等は含まれていなかった）。

検索ボックスがあるだけの簡便なインターフェイスで、キーワードは簡体字、繁体字のいずれでもよい。

■ 目録検索＋書影公開型

日本でも諸事情から典籍の全文をデジタル化できない場合、とりあえず封面や序跋の画像を目録とともに公開して、利用者の参考に供することが行われてきた。一般に書影とは本の概観の画像という意味であるが、ここでは、上述のような最低限度の書誌情報源画像といった意味で用いる。

● 秘籍琳琅：北京大學数字圖書館古文獻資源庫

<http://rbdl.calis.edu.cn/aopac/indexold.jsp>



図7 秘籍琳琅トップページ

秘籍琳琅は、北京大学図書館が所蔵する 150 万冊の古典籍、3 万余種の金石拓本をデジタル化し、Web 配信するプロジェクトの試行版サイトである。

検索画面に入るには、トップページでゲストログインする。この際には、表示された数字 4 桁の「驗證碼」(確認番号)を入力した上で、「登録」をクリックする。

最初は「瀏覽」(閲覧)画面で、主ペインがサイトの使用説明、左ペインは階層検索用にカテゴリフォルダがツリー構造で表示される。カテゴリは大きく古籍と拓片(拓本)に分かれ、古籍の方は版本類別(版本の種類)、出版年代、出版地点の 3 種からブラウジングできる。なお、拓本類のブラウジング機能はゲストログインでは使用できない。



図8 秘籍琳琅「瀏覽」画面

古籍の検索は「瀏覽」画面上部にあるメインメニューから「検索」を選択する。初期画面では検索メニューのうち、古籍と拓本の両方にチェックが入っているが、拓本のチェックを必ず外すこと。拓片にチェックが入ったまま検索すると、ヒット数が表示されるだけで、それ以上はエラーとなる。

検索ボックスへの入力、古籍は繁体字・簡体字のどちらでも検索結果は同じとなる^[3]。またタイトル・著者・主題についてはピンインでの検索もできる。この場合は漢字一字ごとに分かち書きしなければならない。たとえば「漢書」のピンイン入力ならば、hanshuではなく han shu と語と語の間をスペースで区切って入力する。

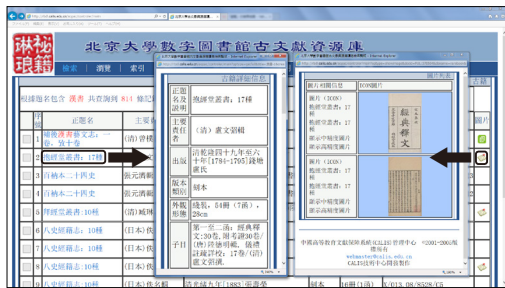


図9 検索結果一覧画面

検索結果は10件ごとの一覧画面で表示され、「正題名」覧のタイトルをクリックすると当該書籍の詳細な書誌が、「圖片」覧のアイコンをクリックすると書影が、それぞれ別ウィンドウで表示される。ただし、残念ながら現段階では、検索結果は最初の10件しか表示できない。

ところが、メインメニューの「索引」をクリックした「索引」画面からは、古籍・拓本全ての目録を閲覧することができる。「索引」画面は古籍・拓本をタイトルもしくは著者名のピンイン順に一覧したものが表示され、必要なタイトルや著者名をクリックすると詳細書誌や図版へのアイコンが表示される。検索も可能で、検索結果一覧の表示制限もない。ただし、拓本類の画像は見ることができない。

試行段階であることを考えると、こういった不完全な部分はやむを得ないのかもしれないが、詳細な説明もないので、利用する側は試行錯誤しなければならない。この点は改善の余地がある。

● 學苑汲古：高校古文獻資源庫

<http://rbcs.calis.edu.cn/aopac/index.htm>

學苑汲古は前述した北京大学の秘籍琳琅を含む25大学図書館（以下の参加大学図書館を参照）の古典籍総合目録データベースで、目録データの検索に加えて一部の書影が閲覧できる。

参加大学図書館：北京・中国人民・清華・北京師範・南開・内モンゴ・遼寧・吉林・復旦・華東師範・南京・蘇州・南京師範・浙江師範・廈門・山東・中国海洋・鄭州・河南・武漢・中山・四川・寧夏・澳門・香港・中文の各大学図書館



図10 學苑汲古トップページ

學苑汲古を運営するのは、^{キャリス}CALIS（高等教育文献保障システム=China Academic Library & Information System）という中国の大学図書館コンソーシアムである。敢えて日本の組織にたてれば、中国版のNII（ただしCALISはコンソーシアムなので組織的にはより緩やかなもの）と考えていただければよいだろう。

CALISでは活動の一つとして、1911年以前に発行された古典的装幀の書籍を古籍と定義し、書誌データの標準化による総合目録の作成を目指し、学術資源の共有化に取り組んでいる^[4]。學苑汲古や秘籍琳琅はいずれもその成果である。

このため両者のインターフェイスは共通であり、操作方法や仕様については秘籍琳琅の項目を参照されたい。秘籍琳琅とは異なり試行版ではないため、検索・表示機能に制限はない（ただし、画像の閲覧については制限のかかっているものもある）。

■ 電子図書館型

● CADAL

<http://www.cadal.zju.edu.cn/index>

キャダル

CADAL（中国数字图书馆国际合作计划=China Academic Digital Associative Library）は、アメリカ、インド、エジプト、中国で行われている国際プロジェクト Million Book Project の中国版名称である。2013年末までのデジタ



図 11 CADAL トップページ

ル化総数は約 275 万点にのぼる^[5]。

Million Book Project とは、大学が主導する世界最大規模の電子図書館プロジェクトで、参加機関がデジタル化した資料をプロジェクト全体で共有するものである。したがって、大学などの機関が資料を提供することを条件に加入でき、参加機関の構成員が利用できるというのが大原則になっている。

残念ながら日本の参加機関はなく、国内で CADAL が自由に利用できる場所はない。ただし、著作権保護期間の満了した書籍については、個人でも登録して利用することができる。CADAL は 1911 年以前に出版された書籍を古籍と定義しているため、古籍の多くは個人で登録すれば自由に閲覧が可能となる。2013 年度末の段階で、236,581 点の古籍が登録されている。ただし、データのダウンロードはできない。

登録はトップページ上部のメインメニュー「注册」(図 11 の○部分)をクリックする。登録方法については湯野基生氏の丁寧な解説^[6]が公開されているので、詳細はそちらをご覧ください。ここでは最低限の解説と留意点について述べる。

注册画面では、以下の項目(*は入力必須項目)につ

図 12 CADAL の利用登録画面

いて入力し、「注册」ボタンをクリックする。

- ①* 用户名: 6 ~ 20 文字の英字・数字・記号(アンダーバーのみ) でユーザー名を指定する。
- ②* 邮箱: メールアドレス
- ③* 登录密码: パスワード (6 文字以上)
- ④* 确认密码: ③のパスワードを確認のため再入力する。
- ⑤ 专业: 専門分野をプルダウンメニューから選択する。
- ⑥ 性别: オプションボタンで選択する。
- ⑦ 生日: 生年月日
- ⑧* 所在单位: 所属機関をプルダウンメニューから選択する。日本に加盟している機関はないので、原則は最下段の「其他」を選ぶ。
- ⑨* 常住地: 居住地。中国国内が基準となっており、省名や市名をプルダウンメニューから選択するようになっているので、日本からの利用は全て最下段の「其他」を選ぶ。
- ⑩* 兴趣: 関心のある分野を記述する。

その後、登録したメールアドレスに確認のメールが届くので、記載の URL をクリックすると登録が完了する仕組みになっているらしい。

ここで、「らしい」と推測表現になっているのは、実際に URL をクリックしてもスクリプトのエラーが出てしまうからである。

ただし、エラーが出て、URL をクリックすることで登録は完了する。再度、トップページに戻って、メインメニュー「注册」の左側にある「登录」(登録)をクリック、先ほど登録したユーザー名(用户名)とパスワード(密码)を入力して「登录」ボタンをクリックすれば、晴れて登録利用者としての利用が可能となる。

古籍の検索は、トップページの検索ボックスに適当なキーワードを入力して検索するか、検索ボックスの下にある古籍のカテゴリから適宜選択してブラウジングする。なお、キーワードは繁体字と簡体字で検索結果が異なるので、双方で検索を試してみる方がよい。

本文画像の閲覧には Adobe Flash Player が必要となる。閲覧用のインターフェイスには、様々な機能が用意されているが、いざ使用しようとするときスクリプトエラーとなる場合もあり、あまり使い勝手はよくない。

また、版本の原所蔵者に関するデータがないため、データ素性の透明性という点では少々問題がある。この点、自己責任において利用されたい。

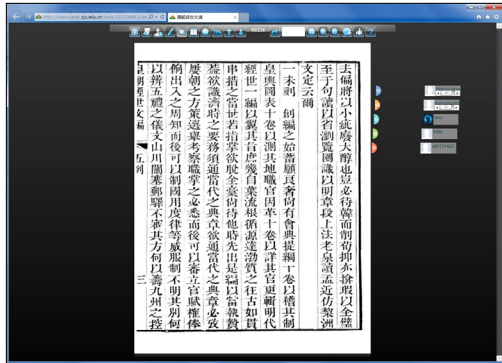


図 13 CADAL での公開画像例

●天津図書館 縮微文獻影像數據庫

<http://swyx.tj.tj.cn/>

民国期刊（雑誌）63 誌、民国報紙（新聞）1 紙、民国図書 176 タイトル、漢文古籍 132 タイトルをデジタル公開している。古籍の多くは清末から民国の版本だが、一部に抄本もあって興味深い。

縮微とはマイクロフィルムのことなので、天津図書館所蔵資料を一度マイクロフィルム化したものをデジタルし、公開したものと思われる。

●大連図書館 特色館蔵

<http://www.dl-library.net.cn/book/>

大連図書館では古籍目録や明清小説のデジタルアーカイブが公開されている。大連図書館トップページ上部のメニューからの「特色館蔵」をクリックすると、メインページに入ることができる。



図 14 大連図書館特色館蔵メインページ

ページ左側には、大連図書館所蔵の特殊コレクションがメニュー化されている。この中には館内限定公開のものもあるので、内容が古典籍かつフリーアクセス

できるものは以下の三つになる。

- ①館蔵明清小説全文データベース：デジタル画像を 40 点公開
- ②館蔵古旧籍书目データベース：8,910 点の目録データを公開
- ③館蔵善本书目データベース：1,711 点の目録データを公開

●中国国家数字图书馆 読者门户

<http://mylib.nlc.gov.cn>



図 15 中国国家数字图书馆 読者门户トップページ

本誌および『電腦中国学入門』では、中国国家図書館の各種リソースについて、折に触れて紹介や解説を試みてきた。これまでに採り上げたような各種データベースは、現在では中国国家数字図書館（以下、数字図書館）として集約されている。

数字図書館はユーザー登録が必要であるため、これまでフリーで利用できていたコンテンツがフリーで使えなくなり、嘆いた方々も多かったと聞く。ただし、本誌で千田大介氏が詳しく解説しているように、外国人でも簡単にユーザー登録ができる¹⁷⁾。

ユーザー登録すれば、中国国家図書館だけでなく、提携の東京大学東洋文化研究所、ハーバード大学イェンチン図書館が所蔵する多くの善本古籍をデジタル画像で閲覧することができる。

このほかにも、古籍のカテゴリに分類されている「数字方志」・「西夏文献」・「宋人文集」や、地方館資源のカテゴリにある「地方志」・「家譜」といった歴史資料の多くをデジタルで閲覧できる。

数字図書館は複数のカテゴリに膨大な量のデジタルコンテンツを有しているので、機会を改めて紹介できればと考えている。

■ その他

その他に分類したものは、各図書館が所蔵する代表的な古典籍の書誌と書影を公開するタイプで、デジタル展示館といった趣が強い。例として以下の2つのサイトを挙げておく。

● 重慶図書館特色館蔵・珍貴古籍

<http://www.cqlib.cn/gczy/guji/>

● 天津図書館特色館蔵

<http://www.tjli.tj.cn/jsp/tsgc/index.jsp?sort2id=30>

❖ 台湾の古典籍検索

● 古籍文献资讯网

<http://rarebook.ncl.edu.tw/rbookod/>

台湾の国家図書館が運営するサイト「古籍文献资讯网」には、①古籍與特藏文獻・②古籍影像検索系統・③明人文集資料庫・④中文古籍書目資料庫・⑤台灣地區家譜聯合目錄・⑥金石拓片資料庫・⑦古籍全文検索系統の7つのコンテンツがあり、大変充実している。

これらはトップページの「資料庫查詢」からリンクしている。



図 16 古籍文献资讯网トップページ

① 古籍與特藏文獻

②・④・⑤・⑥の内容を統合検索できるほか、「特藏電子書」というコンテンツがあり、明清の版本・抄本などが電子化されている。多くは Flash Player があれば閲覧できるが、一部に EPUB 型式のものもある。

② 古籍影像検索系統

ここには、複数のコンテンツがあるので順に解説する。

② -1 特藏珍品選介

国家図書館所蔵資料の中から貴重な典籍 19 点を全文公開している。ただし、閲覧には HyView という専用ソフトをインストールする必要がある。HyView は「ダウンロード」からダウンロードできる。インストールの際に言語を英語、中国語（簡体字もしくは繁体字）から選択できる。中国語を選択した場合は、日本語環境であるとインストールを指示するダイアログが文字化けするので、Windows のシステムロケールを予め中国語環境に変更した上でインストールするのが望ましい（詳しくは本書 138 頁を参照）。

画像はカラーで美しいが、以下の点で学術的な利用の際は慎重な態度をとるべきだろう。たとえば、公開されている典籍中の『永樂大典』は、表紙画像と巻頭画像の間にあるべき表紙裏の画像や、最後にあるべき裏表紙の画像を欠いている。『永樂大典』の表紙裏や裏表紙には、四庫全書編纂時に『永樂大典』を利用した記録が残されていることが多く、この存否は『永樂大典』じたいの書誌学的研究や、さらには真贋判定にとっても欠かせない材料となる。したがって、この部分を欠いているということだけで、学術的価値は半減するのである。

古典籍のデジタル画像は、全てをそのまま公開するのが最も正しいあり方である。少なくとも画像の省略がある場合は、省略した葉（枚）数や省略部分の文字記録の有無や形態情報などについて、何らかの形で明示しておく必要があろう。

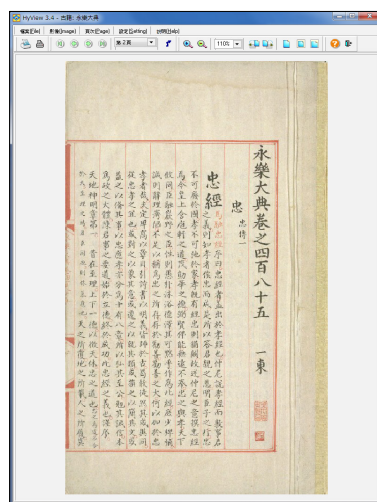


図 17 HyView による閲覧の例

②-2 善本資料查詢

国家図書館の善本・線装本はもちろんのこと、アメリカ議会図書館・ワシントン大学図書館・カリフォルニア大学バークレイ校東アジア図書館所蔵の古典籍を検索できる。

国家図書館内ではデジタル画像で全文を閲覧できるものもあるようだが、館外は書影の公開までに留まっている。

②-3 影像瀏覽

画像公開されている資料だけを、キーワードで検索したり、四部分類や版本の種類などから階層検索できる。ただし前述のように、国家図書館外では書影が見られるだけである。

■ ③ 明人文集資料庫

国家図書館が所蔵する 17 種類の明人の文集について、目録と目次、さらに一部については JPEG 画像で本文を公開するもの。

■ ④ 中文古籍書目資料庫

中国・台湾・韓国・日本・アメリカ・イギリス・ドイツの大学図書館やアジア研究機関、合計 42 機関の古典籍（630,758 件）を検索できる。2000 年以降、研究・討論が重ねられ、各機関からの協力により 2004 年 5 月から公開されており、わが国からは東京大学東洋文化研究所が参画している。

インターフェイスは、繁体字版のほかに簡体字版と英語版も用意されている。簡体字版は簡体字で、英語版はピンインで、キーワードを入力して検索することができる。

■ ⑤ 台灣地區家譜聯合目錄

台湾地区の図書館など 11 機関が所蔵する家譜（約 28,846 種）の総合目録。インターフェイスは繁体字版のほかに簡体字版と英語版も用意されている。英語版は説明言語が英語であるだけで、検索キーワードは繁体字で入力しなければならない。

■ ⑥ 金石拓片資料庫

国家図書館所蔵の金石拓本 6,462 件の目録を検索できる。国家図書館内では拓本の画像を閲覧できるようになっているという。

②と同様に内部コンテンツとして「珍品選介」が用

意されており、文物 8 点の拓本画像が閲覧できる。

■ ⑦ 古籍全文検索系統

「摩訶般若波羅蜜經」と「金剛般若波羅蜜經」の版本画像および対応するテキストが閲覧できる。テキストは全文検索が可能である。また詳細画像の閲覧には Java が必要となっている。

残念ながら、依拠した版本に関する説明はみられない。そこで②-2 の善本資料查詢で検索し、版本の巻頭にみられる「福縁蓮社蔵印」および「佛弟子陳廷照敬蔵印」という蔵書印などを手懸かりにあたりを付けてみた。結果、該当する版本は、「摩訶般若波羅蜜經」が明万曆三十三年（1605）至三十四年（1606）金壇于玉立刊本、「金剛般若波羅蜜經」が明万曆辛丑（二十九年，1601）金壇王肯堂刊本（いずれも国家図書館所蔵）であると推測される。

● 國立故宮博物院 基本古籍資料庫

<http://npmhost.npm.gov.tw/tts/npmmeta/RB/RB.html>

故宮博物院は、清朝までの王宮にあった皇室蔵書や、楊守敬が日本で蒐集した佚存書（観海堂蔵書）など、貴重な善本を多数所蔵することで知られている。

このサイトでは、これらの古典籍の目録が検索できるとともに、書影の閲覧が可能である。



図 18 國立故宮博物院 基本古籍資料庫トップページ

● 古籍善本數位化資料庫

<http://rarebookdl.ihp.sinica.edu.tw/index.html>

上記の URL は正確には「古籍善本數位化資料庫國際合作建置計畫」というプロジェクトのサイトである。このプロジェクトは、中央研究院歴史語言研究所傅斯年図書館・ハーバード大学イェンチン図書館・アメリカ議会図書館アジア部・プリンストン大学東アジア図書館の 4 機関が、古典籍の書誌情報やデジタルアーカイブの共有化をめざすものである。

当該サイトでは、これら4機関の古典籍の目録検索ができ、一部について書影が公開されている。トップページの「資料庫選單」をクリックすると、検索のメインメニューが表示される。



図 19 古籍善本數位化資料庫検索メインメニュー

メインメニューは大きく「資料庫検索」と「資料庫著録」に分かれており、一般の利用者は「資料庫検索」(図 19 の白枠で囲った部分)を使うことになる。「資料庫著録」の方は図書館職員の業務用システムへのポータルで、一般には利用できない。

「資料庫検索」には、さらに、①古漢籍善本數位化資料庫検索系統、②古漢籍善本數位化資料庫一人名資料検索、③古漢籍善本數位化資料庫一印記資料系統の3つのメニューがある。

①は最もオーソドックスな目録検索で、書名・著者名のキーワードからの検索や、四部分類からの階層検索が用意されている。

②は、図書館情報学でいうところの「著者名典拠情報」を検索できる。現代の図書館では、書誌や所蔵のデータベースのほかに、著者に関するデータベースを作成・保有している。これは、書誌には著者の名称しか記録されないため、別途詳細なデータベースを構築して情報を管理する必要があるからである。たとえば、書誌データだけでは、同姓同名の著者の区別がつかなくなるが、書誌データベースの裏側で、著者名典拠情報データベースが機能してくれるおかげで、その区別をつけることができる。ほかにも、改姓・改名や、複数のペンネームを持つ人物についての情報をコントロールしたり、団体名であれば、合併や分離、改組などによる名称の変遷をコントロールしたりしている。OPAC の著者名検索による精度を高め、円滑に情報伝達ができるよう、こういった仕組みが構築されているのである。

ただ、通常は「著者名典拠情報」は業務用であり、これがユーザーの目に直接触れることはほとんどない。ところが、本サイトではこのデータにユーザーがアクセスできるよう整備しているのである。

このデータは非常に詳細なもので、^{あざな}字・号といった異名、生卒年、籍貫(本籍地)、略伝、官僚としての履歴、当該データベースに収録されている著作へのリンクの順に記されている。京都大学人文科学研究所で公開されている、唐代人物知識データベース(<http://tkb.zinbun.kyoto-u.ac.jp/pers-db/>)に近い感覚で利用でき、前近代の中国知識人の履歴を知るための工具書として活用できるだろう。

図 20 人名資料検索で検索した著者名典拠情報

③は典籍に捺された蔵書印影(印記)のデータベースである。印記を集成したものを印譜と呼ぶが、このデータベース版だと考えていただければよい。印文の録文、印主はもちろんのこと、印影の形状や寸法、印の書体、当該印が捺されていた版本など非常に細かく記録されている。これまでに出版されている各種印譜への参照もあり、書誌学研究にとってはこの上なく便利である。

❖ むすびにかえて

デジタルアーカイブといった場合、テキスト化と画像化の二つの方向があることはつとに指摘されてきた。本誌の特集1でも師茂樹氏がこの点について論じているように^[8]、デジタルアーカイブについていえば、中国・台湾はテキストデータが、日本は画像データが主流となってきたのである。

しかし、本稿で採り上げたように、中国でも大規模なデジタル画像によるアーカイブ公開も進捗しつつあり、漢文で書かれた古典籍の画像データは、日中双方を合わせると相当な数にのぼることがわかるだろう。

歴史学の立場からするとテキストデータだけでなく、古籍、なかでも善本の画像データによる公開が進むことは大いに歓迎したい。ともすると我々は便利なテキ

ストデータだけを礼賛しがちであるが、テキストデータにも素性のよくわからないものや、一癖も二癖もあるものもある。

漢文資料のテキストデータの多くは既存の点校本に頼りがちであるが、点校本から作成したデータが必ずしも最良だとは限らない。

たとえば、唐・杜佑の『通典』は唐代後半期の制度の研究には必須の資料であり、政書と呼ばれるジャンルを成したことで目錄学上でも有名である。中華書局が点校本を出してから、もっぱらこれが通行本となっているが、この校訂に問題が多いことは唐代史研究者の間では周知の事実で、最終的に論文などに使用する場合には、北宋版や十通本と対照するのは自明となっている。しかし、テキストデータになって流布してしまった今日、恐らく学生はわざわざ版本に立ち返ろうとはしないであろう。もしかすると、いま使っているテキストデータが成り立つまでに、版本以来どのような系譜があるのかすら意識していないかもしれない。

この意味で、テキストだけでなく善本とされるものの画像公開が進むのは、少なくとも、歴史学の立場からすれば良いことだと考える。

また『唐六典』は、唐代の政治制度の書、さらには中国古代の律令体制の集大成の書としてよく知られている。これも中華書局の校訂本が通行本となっており、この点校は内外で高い評価を得ている。しかし、そうした評価の高い点校本をもってしても、東京大学東洋文化研究所でデジタル画像が公開されている明・正徳本と対照すると、どうしても訂正せざるを得ない部分が見つかるのである。

『六典』は、江戸時代に近衛家熙が校訂した所謂「近衛本」も高い評価を得ている。校訂に際して家熙は正徳本の入手に難渋したという。現代に到っても、10年ほど前であれば、気軽に正徳本を見られるなどということはあり得なかった。しかし、今はググれば正徳本でも嘉靖本でもデジタル画像がヒットする。

逆に言えば、10年前ならば、見られなかった、存在が確認できなかったということで逃げ切れたが、今やそれでは済まされなくなっているのである。便

利になった反面、歴史学者は自身の資料操作に言い逃れをしにくくなってきている。

デジタルアーカイブがどんなに発達しようとも、テキストの素性を理解することや、複数のテキストを比較・検討することは学問の発展の上で、避けては通れないことなのである。これは産業でいえば、どうしても機械化できず、技術者に頼らざるを得ない部分があるのと似通っている。こういった部分と情報化との折り合いをどうつけるのか、これもまた中国学情報化の抱える課題の一つだといえよう。

注

- [1] 橋本秀美「電子版漢籍 2011 年の圧倒的浸透」『明日の東洋学』27, 2012 <<http://ricas.ioc.u-tokyo.ac.jp/pub/pdf/nl027.pdf>>
- [2] このほか、首都図書館には「首都図書館公共検索」(<http://query.clcn.net.cn/Index.htm>) もあり、同様に古籍が検索できる。ただし、こちらは首都図書館 Web サイトからはリンクがなく、正式のサービスではない可能性も高い。このため本稿では註で言及するに留めておきたい。
- [3] 拓本データの検索は繁体字だけに対応するという。
- [4] 詳しくは、樊普「基于 CALIS 高校古文献资源库的图书馆古籍开发与利用」『山西档案』2012(4)などを参照。
- [5] CADAL および Million Book Project の現状に関しては、CADAL の Web サイトおよび、篠田麻美「中国における Million Book Project : 中国の大学図書館の資料電子化戦略」『カレントアウェアネス』298, 2008 <<http://current.ndl.go.jp/ca1678>>、湯野基生「中国の資料デジタル化プロジェクト・CADAL の利用と参加について」『アジア情報室通報』12(1), 2014 <<https://rnavi.ndl.go.jp/asia/entry/bulletin12-1-1.php>>などを参考とした。
- [6] 前掲注 5 湯野論文参照。
- [7] 千田大介「中国の論文を探す」『漢字文献情報処理研究』15, 2014。
- [8] 師茂樹「学術データベースの構築と公開」『漢字文献情報処理研究』15, 2014

「農業古籍」について

村上 陽子

❖ はじめに

歴史学において史料を講読するには、まず複数のテキストを入手して史料批判するのが必須である。それは中国農業史においても同様であり、例えば西山武一・熊代幸雄『校訂譯註齊民要術』では、20 種を超えるテキストを挙げている^[1]。また元・王禎『農書』でも嘉靖九年本が3 種、乾隆三十九年本が2 種というように、同一の刷年で複数の異なるテキストが存在したりする。こういった多種のテキストを収集して比較することが研究のスタートとなる。

❖ 中国農業科学院農業信息研究所と「農業古籍」

テキストを収集するのが、作業の第一歩である。とは言うものの、中国にしかないようなテキストの入手は、困難である場合が多い。この意味で、中国農業科学院農業信息研究所の「农业古籍（農業古籍）」のようなサイトは、これからの研究に有益である。

● 农业古籍

<http://211.147.18.86/>



図1 「農業古籍」トップページ

まず、このサイトを作成した中国農業科学院農業信息研究所について述べておこう。

中国農業科学院は、中華人民共和国農業部に直属する組織である。1957 年3 月に設立され、現在ではその傘下に40 を超える研究所を中国各地に配しており、農業信息研究所はその一つである。そこでは業務の一環として、農業関係情報のデータ蓄積とその応用研

究、そして収集した古籍のデジタル化をしている。これまで研究所（国家農業図書館）で収蔵された古籍は約15,000 冊、3494 種^[2]。『農政全書』の明刻本などの貴重な版本も含まれている。それらを研究者だけでなく、外部の一般の利用者、さらには海外の利用者も想定してデジタル化し公開した。それが「農業古籍」である^[3]。

■ 「農業古籍」：目録検索

それでは検索方法をみていきたい。まず「農業古籍」のトップページから「目録（目録）」または「全文」を選ぶ（図2）。「目録」は、農業信息研究所が収蔵した古籍の目録を検索することができる。一方「全文」は、デジタル化した資料を検索・閲覧できる（このほか「词典（辞書）」もあるが、現在は利用できない^[4]）。



図2 トップページ：部分

「目録」「全文」ともに、それぞれ、所蔵資料の全体を包括した「总库（総庫）」と、資料の内容によって「古农书（古農書）」「史书（史書）」「地方志」「类书（類書）」の4 つに区分したページ構成になっている（図3）。

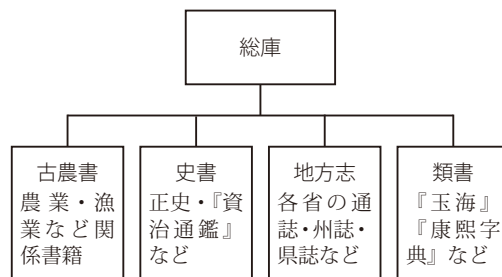


図3 「目録」と「全体」それぞれの構成

例えば「总库」の検索ページは、図4 のようになる。「古农书」から検索するには、○印で示した部分をクリックする。

また画面左の主題別分類フォルダーをクリックすると、主題毎に収録されている資料の一覧を見ることが



図4 「目録」中の「総庫」の画面

できる。

では、検索してみよう。まずウェブブラウザは、IEかGoogle Chromeを利用する^[5]。

次に検索ボックスに検索語を入れる。このとき入力モードを中国語（簡体字）に切り替えておく必要がある。例えば「齊民要術」「齊民要術」ではヒットせず、「齐民要術」ではヒットする。次にプルダウンメニューから、検索対象を書名・作者・分類・出版年・キーワード・「所有字段（ここでは書誌中のあらゆる語彙）」から選び、「模糊」（部分一致検索）または「精确」（完全一致検索）を選択して検索する。

例えば「目録」の下位にある「总库」のページで、登録されている書誌全体の中から、農書の『齊民要術』を<「齐民要」+書名+模糊>で検索すると、図5のような検索結果が表示される。目録は1冊毎に作成されており、34冊がヒットした。具体的には、光緒紀元年・石印本（4巻本）、光緒二十二年・刻本（4巻本）の版本や影鈔本等と、『齊民要術校釈』『齊民要術今釈』などの注釈本である（ちなみに精确で検索すると、当然ヒット0となる^[6]）。

また4区分でそれぞれ検索すると、「古農書（古農書）」では30冊、「史書」「地方志」では0冊、「类书（類書）」では4冊ヒットした。これは「总库」でヒットした内の『津逮秘書』に引用された『齊民要術』4冊が「类书」の区分であったため、「古農書」30冊、「类书」4冊となったのである^[7]。

検索結果には、書名・作者・出版年・版本項・ページ数・キーワード（キーワード）・書籍の概略が記されている。（概略が長文の場合はすべて表示されず、赤字の「…」にカーソルを合わせると、後半が表示される仕組みである。）

ここから検索対象で「キーワード」を選んだ場合、その検索対象となるのは、書誌にあらかじめ付与しているキーワードのことであることがわかる。キーワードは



図5 「目録」での検索結果

『農業科学叙詞表』を利用して付与されている^[8]。例えば図6の3つの目録を見てみよう。

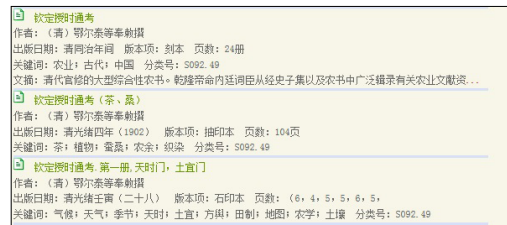


図6 『欽定授時通考』の書誌検索結果の一部

一見して分かるように、同じ『欽定授時通考』であっても、付与されているキーワード（キーワード）は異なる。例えば3冊目のように『農業科学叙詞表』から選択した語彙が入っているものがあれば、1冊目のように「农业」「古代」だけ、とキーワードの付与に一定のルールを見いだせない。このため書名・作者で検索するのが妥当であろう。

■「農業古籍」：デジタル画像検索

では、次にデジタル化された資料をみてみよう。まずトップページから「全文」ページに入る。

「目録」ページを同様に、「总库」から<「齐民要」+書名+模糊>で検索すると2件ヒットする。この結果から、農業情報研究所で所蔵している『齊民要術』34冊のうち、現在デジタル化されて閲覧できるのは2冊（同じ影鈔本。収録巻数が異なる）だということになる（図7）。

検索結果からデジタル化された資料を見るには、図7の検索結果に表示された各書名をクリックすればよ

い(図8)。



図7 「全文」での検索結果

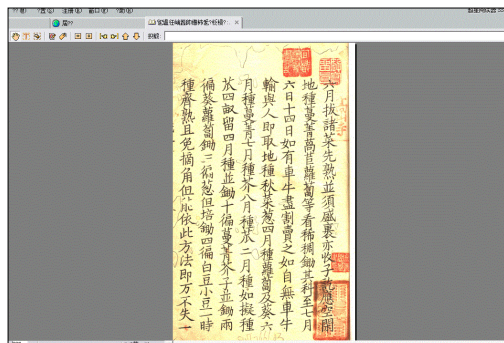


図8 検索結果(図7)から、SSReaderを使用してデジタル化された『齐民要術』を開いたところ

■ デジタル画像を検索する前に

これまで、サイトの解説と検索および画像の閲覧方法を述べてきた。このように書誌やデジタル化資料を検索するだけなら、通常の日本語版 Windows 環境で問題はない。しかし、デジタル化した資料を見るためには、超星網の SSReader を利用する必要がある。(例えば PDF を見るためには、アドビ社のプラグインが必要のように。)

また SSReader をインストールするためには、システムロケール(動作言語)を中国語(中国)に切り替えておかなければならない。Windows のシステムロケールは、以下の手順で切り替える^[9]。

- ①コントロールパネル→時計・言語および地域→場所の変更→管理タブ(Windoes8.1では、コントロールパネル→地域→管理タブ)
- ②システムローケルの変更をクリック→中国語

(簡体字、中国語)を選択→OK

③再起動

そのあとで SSReader をインストールする^[10]。

SSReader は図4・5・7の画面右上にある「阅读图书请先下载安装浏览器」からダウンロードできるが、対応 OS が Windows XP までと古いものなので、Windows7以降に対応する最新バージョン(ver.4.1.3)を、各種ダウンロードサイトから落としてくるほうがよい。またダウンロード完了後、手順にしたがってインストールすると、デスクトップに「超星阅读器」のアイコンができる。これをクリックすると、SSReader が立ち上がりパスワード入力を求められるが、これは超星数字図書館の利用パスなので、無視しても「農業古籍」の利用はできる。

ここまで済ませて SSReader 上の URL 欄に、「農業古籍」の URL を入力すれば、「全文」ページの検索結果から画像を見ることができる(図8)。

閲覧環境設定に多少手間がかかるが、一度インストールしてしまえば、Windows のシステムロケールを日本語に戻しても閲覧することができる。SSReader の簡体字中国語メニューは文字化けしているが、「農業古籍」の検索・閲覧には問題ないようだ。

◇ デジタル化された資料の閲覧について——高山寺本『齐民要術』を主に

最後に、当サイトでどれだけのデジタル化資料を閲覧することができるのか述べてい。

先に検索したように古農書の筆頭ともいえる『齐民要術』は、1種の資料それも巻第一の一部分と巻第五・巻第八しか見られない。しかし目録には北宋の影鈔本(透き写しをした本)とあり、そのうえ「高山寺」の朱印が見られる(図8・9)。

わが国の高山寺に残存するのは、全十巻のうち巻第五と巻第八のみである^[11]。石塚晴通氏の調査によると^[12]、高山寺本は、料紙が漉返し(再生紙)であり、竹を主な繊維とし、楮が少量混ざっているとある。石塚氏の報告では巻第五・巻第八の全文の画像が掲載されているので見比べると、わが国所蔵の資料には界線があるが、農業情報研究所蔵本には見られない。また朱印の押印数の違いがある。図9右の巻第五の「高山寺」の押印は、日本のものは巻数「巻第五」下だけ

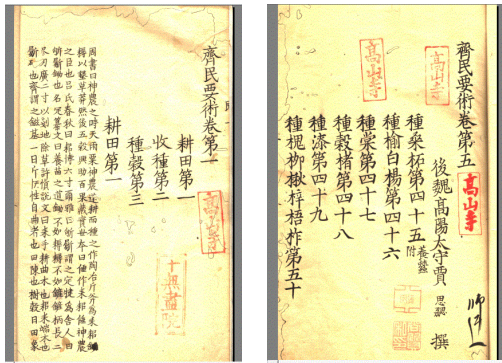


図9 左：巻第一 頁目 右：巻第五 1 頁目

にあり、画面上部の2か所の押印はない。さらに農業
 情報研究所蔵本にのみ押印されている朱印もある
 (図9右画面下の2つの朱印)。

何より「農業古籍」によって分かったのは、農業信
 息研究所には巻頭雑説の最後の部分と巻第一の一部
 とが残存していることである(図8・図9左)。4ペー
 ジ分とはいえ、それを見られるのは非常に貴重である。
 わが国の高山寺本を写し取った後、いかなる経緯を
 もって中国へ伝わり、これまで残ってきたのか、考え
 ると真に興味深い。

しかし、その他の基本的な農書、例えば王禎『農書』
 『農桑輯要』や本草書の『本草綱目』など、「目録」検
 索で清代に出版された資料の所蔵を確認できるが、「全
 文」検索では該当なしとなる。まだデジタル化事業が
 進んでいないのか、またはアップされていないよう
 である。そもそも現在 12,000 冊を超える目録登録数
 のうち、デジタル化されて検索できる資料は 27 冊と、
 ごく一部である。サイトが開設されて以後、どれだけ
 更新されているのか一抹の不安が残る。

目録検索結果とデジタル化画像検索結果とのヒット
 数のあまりに大きな開きは、このサイトを利用するも
 のにとっては非常に残念なところである。今後、デジ
 タル化された資料が増加することを切に望む。

注

- [1] 西山武一・熊代幸雄『校訂譯註齊民要術』(上巻 1957 年・
 下巻 1959 年、のち 1969 年にアジア経済出版会より合
 冊本で出版された)。
- [2] 農業情報研究所の国家農業図書館サイトによる。
<http://www.nais.net.cn/site/default.aspx> 2014.9.1 参

照

- [3] 常春ほか「基于図像のデジタル化農業古籍全文検索方案」
 (『情報検索』No.6, 2005 年)、常春・張桂英「農業古
 籍数字図書館項目評価方案」(『現代情報』2005 年 11 月、
 第 11 期)
- [4] 2014.9.1 現在
- [5] 検索時のウェブブラウザで Firefox を利用すると、検
 索結果がすべて文字化けするので、表示されるたびに
 エンコードを中国語(簡体字)に変更しなければなら
 ない。IE・Google Chrome では簡体字のままでも検索結
 果が表示された。また Firefox を利用した場合、「书名」
 以外の検索対象たとえば「キーワード」(キーワード)で
 検索しても、結果画面の検索対象表示は「书名」になっ
 てしまう。ただ検索結果はキーワード検索で選ばれた
 内容なので、検索機能は正確に動いているようだ。
- [6] 「齊民要術」で検索すると 14 冊。差の 20 冊は、書名に「第
 1 冊」「今釈」等が入っていたために外れた資料である。
- [7] 『津逮秘書』は叢書であるが、同じく叢書の『説郭』も「類
 書」に区分されている。また『津逮秘書』はキーワ
 ードに「丛书(叢書)」とあり、『説郭』は「雑著(雑著)」
 とある。「類書」に区分されている資料は幅広いようだ。
- [8] 盧文林「農業古籍データベース建立和著録実践」(『農業図書
 情報学刊』第 23 巻第 11 期、2011 年 11 月)
- [9] 中国語版 Windows 環境にするための詳細は、漢字文献
 情報処理研究会『電腦中国学入門』(好文出版、2012 年)
 p.62-63 も参照されたい。
- [10] 超星網(超星数字図書館 <http://www.chaoxing.com>)
 と SSRider については、注 9 の『電腦中国学入門』
 p.182-183 のほか、安藤一博「中国の電子図書館「超
 星数字図書館」」(カレントアウェアネス No.273, 2002
 年)が詳しい(<http://current.ndl.go.jp/ca1472>)。
- [11] 注 1 および注 12 参照
- [12] 石塚晴通「高山寺本宋版齊民要術巻第五」、同「高山
 寺本宋版齊民要術巻第八」(高山寺典籍文書総合調査
 団『高山寺典籍文書総合調査団研究報告論集』平成 23
 年度、平成 24 年度)

<補記>

中国農業科学院農業遺産室のサイトに
 「中華農業文明網」
<http://www.icac.edu.cn/home.asp>
 がある。こちらは学会動向などが随時更新され、また
 伝統農具デジタル博物館などもあり、研究者だけでなく
 初学者にとっても有益なサイトである。

年会費の納入についてのお知らせ

年会費が未納の方は、下記口座にお振り込み下さい。年会費は BBS 会員が 1,000 円、一般会員が 3,000 円になります。

郵便振替

口座番号：00120-1-91107

口座名称：漢字文献情報処理研究会

三菱東京 UFJ 銀行 京都駅前支店

口座番号：普通 4571788

口座名称（漢字表記）：漢字文献情報処理研究会代表 師茂樹

（機械振込等カナ入力）：カンジブンケン モロシゲキ

ゆうちょ銀行

ゆうちょ銀行 〇一九（ゼロイチキュウ）店

（金融機関コード：9900 店番：019）

預金種目：当座

口座番号：0091107

カナ氏名：カンジブンケンジョウホウショリケンキュウカイ

漢字文献情報処理研究会彙報

2013.10～2014.9

2013 年 10 月 1 日

会誌『漢字文献情報処理研究』第 14 号刊行

2013 年 12 月 22 日

第 16 回大会、2013 年度総会開催

2014 年 2 月 28 日

『人文学と著作権問題—研究・教育のためのコンプライアンス』刊行

2014 年 8 月 2 日

2014 年度夏期公開シンポジウム「情報化時代における中国学研究基盤を考える」開催

◆総会

- 2013 年度事業報告、会計報告（会計監査：石井公成・矢野正隆）
- 執行部改選
代表：師茂樹
副代表（兼会誌編集局長）：千田大介
副代表（会計・名簿担当）：小島浩之
幹事（兼サーバ管理担当）：上地宏一
幹事：佐藤仁史・田邊鉄・二階堂善弘・山田崇仁
- 2014 年度事業計画、予算案承認
- 会誌・書籍販売促進のお願い

第16回大会・2013年度総会

日時：2013 年 12 月 22 日（日）13:30 ～ 17:30

会場：花園大学・拈花館 202 教室

◆第 16 回大会

研究発表

- 『中国古今地名大辞典』のデジタルテキスト作成と Wordpress を利用した Web 公開について
山田崇仁（立命館大学）
- 日中両言語混在テキストデータリソースの継承について
二階堂善弘（関西大学）
- フコト点の電子化と移点ツール
田島孝治（岐阜高専）・高田智和（国立国語研究所）
- デジタルヒューマニティーズ／デジタルヒストリーの情報源
菊池信彦（国立国会図書館）

2014年度夏期公開シンポジウム

題目：情報化時代における中国学研究基盤を考える

日時：2014 年 8 月 2 日（土）13:00 ～ 17:30

会場：関西大学 CSAC4 階セミナースペース

- 学術情報収集のスキルとインフラ整備
千田大介（慶應義塾大学）
- 学術データベースの構築と公開
師茂樹（花園大学）
- 学術情報共有プラットフォームとしての CiNii
大向一輝（国立情報学研究所）
- フリーディスカッション

執筆者紹介

石岡 克俊 (いしおか かつとし)

1970年北海道生まれ。現在、慶應義塾大学大学院法務研究科准教授。経済法、知的財産法、消費者法が主な専門。著作物再販制度についての検討以来、著作権法などの知的財産制度と独占禁止法の交錯領域を主なフィールドとしてきた。最近では、コンテンツに対する関心をさらに発展させ、電気通信事業や放送事業などメディアやコンテンツ流通に対する規制と競争の問題に関心の対象を広げている。

大向 一輝 (おおむかい いっき)

1977年京都生まれ。2005年総合研究大学院大学博士課程修了。博士(情報学)。2005年国立情報学研究所助手、2009年同准教授。セマンティックウェブやソーシャルメディア、オープンデータの研究とともに、学術情報サービス CiNii の開発に携わる。著書に『ウェブがわかる本』(岩波書店)、『ウェブらしさを考える本』(丸善出版)がある。

川幡 太一 (かわばた たいち)

NTT 未来ねっと研究所・研究員。本業の傍ら、漢字・文字符号やテキスト処理等に関心を持つ。国際文字符号標準化委員会：ISO/IEC JTC 1/SC 2 の国内対応委員会の委員。

小島 浩之 (こじま ひろゆき)

1971年岐阜県生まれ。東京大学大学院経済学研究科講師・経済学部資料室室長代理。専門は東洋史学および歴史資料の保存と活用に関する研究。最近の業績に「唐代後半期の官僚人事と八儁」『明大アジア史論集』18、「『大唐六典』の構造と史料的性格」『『唐六典』巻六・尚書刑部訳註稿』下など。

佐藤 信弥 (さとう しんや)

1976年兵庫県生まれ。関西学院大学大学院文学研究科博士課程後期課程満期退学。博士(歴史学)。現在は立命館大学白川静記念東洋文字文化研究所客員研究員。専門は中国殷周史。著書に『西周期における祭祀儀礼の研究』(朋友書店、2014年)がある。

<http://blog.goo.ne.jp/xizhou257>

Twitter ID:satoshin257

佐藤 仁史 (さとう よしふみ)

1971年愛知県生まれ。一橋大学大学院社会学研究科教授。専門は中国近現代社会史。著書に、『近代中国の郷土意識——清末民初江南の在地指導層と地域社会』研文出版、2013年、『嘉定県事——14至20世紀初江南地域社会史研究』(共著) 広東人民出版社、2014年、がある。サイトは、<http://www.soc.hit-u.ac.jp/~satojin/>。

須賀 隆 (すが たかし)

1958年愛媛県生まれ。株式会社メビウス基礎技術研究室所属。博士(工学)。基本的な関心事は「ものごとに目盛りをつけてはかること」。その延長線上で時間に目盛りをつけたもの一層—に興味を持ち現在に至る。<http://www.asahi-net.or.jp/~dd6t-sg/>

田邊 鉄 (たなべ てつ)

1963年京都府生まれ。大阪外国語大学大学院外国語学研究科東アジア語学専攻修士課程修了。現在、北海道大学情報基盤センター准教授。中国語・インターネット・ガジェットの「三結合」により、(自分が) 飽きない授業コンテンツの開発を行っている。最近は教科書のダイアログでどうすれば「キャラが立つ」のかについて、考えているところ。

千田 大介 (ちだ だいすけ)

1968 年東京生まれ。慶應義塾大学経済学部教授。当初は明清小説を研究していたようだが、伝統芸能や現代文化・情報関係など守備範囲が広がりすぎて説明が面倒なので、学生にはここ千年くらいの中国のサブカルチャーが専門と自己紹介している。今年度こそは好文出版から『中国同時代文化小事典』を刊行できるはず。

脳瓦崗寨：<http://wagang.econ.hc.keio.ac.jp/>

二階堂 善弘 (にかいどう よしひろ)

1962 年東京の下町生まれ。東洋大学文学部卒業、早稲田大学大学院文学研究科博士課程退学。現職は関西大学文学部教授。博士（文学）・博士（文化交渉学）。専門は中国の民間信仰研究。最新の著作は『アジアの民間信仰と文化交渉』（関西大学出版部）。サイトは、<http://www2.itc.kansai-u.ac.jp/~nikaido/>

氷野 善寛 (ひの よしひろ)

1980 年大阪生まれ。関西大学博士後期課程単位取得満期退学、博士（文化交渉学）。現在は関西大学アジア文化研究センター PD、京都女子大学、関西大学非常勤講師。『官話指南』を中心とした近代における中国語教材の研究の傍ら、様々な媒体を用いた中国語デジタル教材の開発を行っている。また所属センターでは近代中国語コーパスをはじめとして東アジア研究に関わるデータベースやアーカイブズの構築及び設計を担当している。

紅粉 芳恵 (べにこ よしえ)

1965 年兵庫県生まれ。関西大学大学院外国語教育学研究科後期課程修了、博士（外国語教育学）。関西大学、神戸市外国語大学や一般企業で広く中国語を教えた経験を生かし、2012 年度より京都産業大学全学共通教育センター講師。中国語語彙教育を専門とし、共著で『キクタン中国語』シリーズ(アルク)、『中国語成語ハンドブック』（白水社）などを出版。中国語教育のレベル向上を目指す「中国語教授法研究会」を立ち上げ、阿部慎太郎氏、氷野善寛氏とともに幹事を務める。

村上 陽子 (むらかみ ようこ)

1969 年広島県生まれ。上智大学大学院後期博士課程を単位取得満期退学後、博士（史学）を取得。現在は東京外国語大学 AA 研共同研究員。専門は中国食生活史。農業を含め食に関わる領域を研究対象としている。最新の業績は「何を食べていたのか」（中国出土資料学会編『地下からの贈り物 新出土資料が語るいにしへの中国』第 1 章 7 節、東方書店、2014 年）。

師 茂樹 (もろ しげき)

1972 年生まれ。花園大学准教授。博士（文化交渉学）。ここ数年、学生を連れてフィールドワークをする機会が増えてきた。最近書いたものとしては、“Xuanzang's proof of idealism（真唯識量）and Silabhadra's Teaching”（『2013 年第一屆慈宗國際學術論壇論文集』、2013 年 8 月）など。Twitter ID: @moroshigeki

矢野 正隆 (やの まさたか)

1972 年福岡県生まれ。東京大学大学院経済学研究科助教。専門はベトナム前近代史およびアーカイブズ資料の保存。最近の論考に「MLA におけるメディアの特性とアクセスに関する試論」（『アーカイブズ学研究』20）など。

山田 崇仁 (やまだ たかひと)

1970 年愛知県生まれ。立命館大学院文学研究科修了。博士（文学）。大阪大谷大学・滋賀大学・立命館大学非常勤講師。最近は、「漢字」にまつわる言説がどのような過程で形成されてきたのかについて、あれこれ調べている最中。<http://www.shuiren.org/>

編集後記

『漢字文献情報処理研究』第15号をお届けする。

今号は例年に比べていささかページ数は少なめであるが、しかし、夏期シンポジウムの報告や、これまでまとまった形で採りあげてこなかった論文検索の特集、さまざまなWebサイトのレビューなど、記事の充実においては既刊各号に遜色ない仕上がりであると自負している。

記念論文集の原稿募集案内でも触れたように、本誌は次号が1つの区切りとなる。漢情研第一期の締めくくりにふさわしいものにすべく、編集日程も前倒しで進行することになるだろう。漢情研内外の諸賢には、原稿執筆などご協力をお願いすることもあるかと思われるが、ご協力賜れたら幸いである。

近頃、大学教育の英語化をめぐる動きが慌ただしさを増しており、スーパー・グローバルなる20世紀ロボットアニメ風ネーミングのプロジェクトもスタートする。こうした動きの中で気になるのが、英語化＝国際化という昔ながらの混同である。現在、OS・アプリケーションソフトの国際化、i18nがかなりの程度実現され、1つのシステムでさまざまな地域の言語に対応できるようになっているが、それは英語を媒介として実現されたわけではなく、あらゆる言語をフラットに扱うための枠組みを作りあげた結果である。そうしてみれば、英語さえできれば国際化というのは、あたかもWindows 95/98やMS DOSを国際化OSというようなものではないか。事実上の国際語としての英語の重要性は否定しないが、各国・地域の文化的多様性を尊重した本当の意味での国際化教育にも十分に力が注がれるべきであると考えている。

本誌が本年も滞りなく刊行できたのは、ひとえに好文出版の尾方社長、編集委員各位、会員諸賢、および関係各位のご援助の賜物である。ここに記して感謝申し上げる。(★)

漢字文献情報処理研究 第15号

発行日 2014年10月2日
定価 本体2,000円＋税
編集 © 漢字文献情報処理研究会
<http://www.jaet.gr.jp/>
編集委員 ○千田 大介 金子 真也
上地 宏一 小島 浩之
佐藤 仁史 田邊 鉄
二階堂善弘 師 茂樹
山崎 直樹 山田 崇仁
デザイン & DTP 電腦瓦崗寨：
<http://wagang.econ.hc.keio.ac.jp/>
発行人 尾方敏裕
発行所 株式会社好文出版
〒162-0041
東京都新宿区早稲田鶴巻町540
林ビル3F
TEL: 03-5273-2739
FAX: 03-5273-2740
URL: <http://www.kohbun.co.jp/>

◎本誌に関する訂正・補足情報は、漢字文献情報処理研究会サイト (<http://www.jaet.gr.jp/>) に掲載します。

◎本誌の定期購読をご希望の方は、以下の項目につき明記の上、好文出版まで、書面・FAXもしくは電話にてお申し込みください（住所・FAX・電話は上記奥付参照）。

○送付先住所 ○氏名 ○年齢 ○職業
○勤務先 ○必要部数

◎漢字文献情報処理研究会への入会をご希望の方は、<http://www.jaet.gr.jp/guiding.html>の趣意書および規約をよくお読みの上、同ページにリンクが掲載されている入会フォームよりお申し込みください。書面での申し込みは受け付けておりません。

ISBN978-4-87220-183-3

C3004 ¥2000E

