

漢字文献 情報処理研究

第4号

漢字文献情報処理研究会 編

日本中国語 CAI 研究会 編集協力

好文出版

漢字文獻情報処理研究 第4号

目次

論文	4 篆隸万象名義データベースの改訂	池田 証寿
	12 『老子』傳奕本来源考 テキスト処理による「項羽妾本」介在の検証	秋山 陽一郎
特集 1：人文学情報処理教育	21	
	22 文学部における情報処理教育 ——大谷大学人文情報学科における試み——	松川 節
	28 花園大学「情報歴史学コース」の現状と課題	佐藤 誠、田中 正流
	32 オンライン討論 人文学情報処理教育における教える側の問題	
漢情研 2003 年夏期公開講座 東洋学情報化と著作権問題 報告	39	
	40 「東洋学情報化と著作権問題」参加レポート	師 茂樹
	43 東洋学情報化と法律問題——第1回 所有権の行使と無体財産の法的保護 判例の分析と解説	石岡 克俊
	53 著作権についての知識を深めよう 東洋学のための著作権サイト・ページ指南	小島 浩之
特集 2 漢字処理技術の最新動向	57	
	58 CHISE Project	守岡 知彦、江渡 浩一郎、苔米地 等流、宮崎 泉、師 茂樹
	70 文字コード外フォント処理	上地 宏一
	74 Embedding Glyph Identifiers in XML Documents	Christian Wittern
	80 JIS X 0213 の改正と UCS との関係について	川幡 太一
	86 Unicode 4.0	師 茂樹
中国語 CAI 実践レポート	89	日本中国語 CAI 研究会
	90 市販教材ソフトは使えるか 中国語入門／初級授業での実践	田邊 鉄
	96 Blackboard Learning System について バーチャルユニバーシティ構築の可能性	高橋 由利子

ソフトウェア レビュー 103

- | | | |
|-----|-----------------|---|
| 104 | OS | 総括 / Linux
二階堂 善弘 / 師 茂樹 |
| 108 | ワープロソフト | 一太郎 13 & ATOK16 / EGWORD
山田 崇仁 / 二階堂 善弘 |
| 111 | IME と中文ツール、フォント | cWnn (楽々中国語) / Chinese Writer 7 /
フォント / 翻訳ツール
二階堂 善弘 / 千田 大介 |
| 121 | ネットワーク | WWW ブラウザ / その他
秋山 陽一郎 / 千田 大介 |
| 127 | その他 | テキストエディタ / その他 / 中国のソフトとネットの動向
山田 崇仁 / 千田 大介 |
-

学術リソース レビュー 137

- | | | |
|-----|----------|---|
| 138 | 学術サイト | 図書館と OPAC / 東アジアの言語に関連したサイト /
中国語教育サイト / 中国古代史 / 中国近現代史 /
中国古典文学 / 中国現代文学 / 仏教
小島 浩之 / 山崎 直樹 / 小川 利康 / 山田 崇仁 /
佐藤 仁史 / 千田 大介 / 師 茂樹 |
| 165 | 学術ソフト・製品 | 国学の CD-ROM / 中国の文献データ製品 /
電子辞書 2 機種 / 字通
千田 大介 / 田邊 鉄 / 山田 崇仁 |
-

書評 178 『青森県史資料編古代 1 補遺』/『文字コード「超」研究』/『電腦日本語論』

コラム 136 勝村哲也先生への追悼文 石川 正敏

180 漢字文献情報処理研究会彙報 / 著者紹介

◇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、™・®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。

篆隸万象名義 データベースの改訂

池田 証寿 (いけだ しょうじゅ)

回 多漢字文献としての漢字字書

Unicode に代表されるように、文字コードの拡張は一段と進んでいる。また、今昔文字鏡も登録の文字を拡大し、従来よりもはるかに、多漢字文献の情報処理は容易になった。篆隸万象名義という日本の平安時代に編纂された漢字字書を例として、多漢字文献処理の実情を報告したい。

回 篆隸万象名義の研究の現状と問題点

篆隸万象名義（高山寺本、全六帖）は、弘法大師空海（774-835）が撰述した、部首分類体の漢字字書である。その内容は、中国・梁の顧野王が大同二年（536）に撰述した玉篇三十巻からすべての掲出字を取り出し、玉篇の長大な注文を節略したものである。玉篇（後の大広益会玉篇と区別して原本系玉篇、または原本玉篇と呼ぶ）が広範に利用されたことは、日中の諸書に引用された逸文の存在によって明白であるが、玉篇の現存本は、巻八、九、十八、二十二、二十四、二十七が知られるに過ぎない^[1]。玉篇の現存量は全体の八分の一であり、玉篇の内容を忠実に伝える篆隸万象名義の存在価値は高い。現存の古写本は高山寺本が唯一といってよく、江戸時代に高山寺本を転写した以外の写本は知られていない。

玉篇はその後、蕭愷による刪改、孫強による増

字（唐・上元元年（674））を経て、宋の大中祥符六年（1013）、陳彭年等による大広益会玉篇に至る。大広益会玉篇は、三十巻三冊だが、原本系玉篇にあった長大な注文は簡略化されている。

原本系玉篇と篆隸万象名義は、明治十三年（1888）、清国から来日した楊守敬が日本に伝存する古逸書を探索した成果として、両書の真価を中国に知らしめたものである。楊守敬による原本系玉篇残巻の発見・紹介と篆隸万象名義の高山寺本及びその転写本作成の事情から、最近に至るまでの研究の成果は、高田時雄「篆隸萬象名義解説」（『定本弘法大師全集』第九巻、密教文化研究所、1995年）^[2]に詳しい。

篆隸万象名義全六帖のうちの前半の四帖は空海自撰であるが、後半二帖は「惹囊三佛陀」なる人物による続撰である。前半の自撰部は、全体では百巻構成となるうちの五十巻までを収め、後半の続撰部は玉篇どおりの三十巻構成の後半の十五巻を収めている。

篆隸万象名義が原本系玉篇を抄録している点を考慮して、篆隸万象名義と原本系玉篇残巻とを比較してみると、玉篇残巻の巻二十七と篆隸万象名義とは誤りを含めて一致しているのに対して、巻九、十八、十九は不一致が多く、巻二十二は不一致が少ない。これは貞苺伊徳「玉篇と篆隸万象名義について」（『国語学』第三十一輯、1957年）による研究の成果である。上田正「玉篇残巻論考」（『神戸女学院大学論集』17-1、1970年）は貞苺説を推し進め、反切用字の分析から依拠玉篇の新

古を追求している。

中国語音韻史の研究では、周祖謨「萬象名義中之原本玉篇音系」（『問学集』）、河野六郎「玉篇に現れたる反切の音韻的研究」（『河野六郎著作集』二、平凡社、1979年）がある。

国語学、特に辞書史研究では、篆隸万象名義を玉篇の代用として利用し、新撰字鏡と玉篇との関係を解明した貞苺伊徳「新撰字鏡の解剖（要旨）——その出典を尋ねて——」（『訓点語と訓点資料』第十二輯、1959年）^[3]などの成果があり、篆隸万象名義と図書寮本類聚名義抄との関係に関して、宮澤俊雅「図書寮本類聚名義抄に見える篆隸万象名義について」（『訓点語と訓点資料』第五十二輯、1973年）、池田証寿「図書寮本類聚名義抄と篆隸万象名義との関係について」（『人文科学論集』第二十七号、信州大学人文学部、1993年）、同「図書寮本類聚名義抄の単字書的性格」（『国語国文研究』第94号、1993年）^[4]などがある。

国文学では、小島憲之「上代に於ける学問の一面——原本系玉篇の周辺——」（『文学』39-12、1971年）、同「日本書紀の「よみ」——原本系玉篇を通して——」（『文学』41-8、1973年）があって、上代の訓詁学における玉篇の利用を説く。

玉篇の逸文の収集と整理は、岡井慎吾『玉篇の研究』（東洋文庫、1933年）を嚆矢とし、これを補正した馬淵和夫「玉篇佚文補正」（『東京文理科大学国語国文学会紀要』第三号、1952年）があり、個別の文献については、井上順理「令集解引玉篇佚文考——「孟子伝来考」附論——」（『鳥取大学教育学部研究報告人文・社会科学』通号17、1966年）をはじめとして、その成果は多い。

宮澤俊雅「篆隸万象名義掲出字一覧表」（『高山寺古辞書資料第一（高山寺資料叢書第六冊）』東京大学出版会、1977年）は、篆隸万象名義と玉篇残巻・大広益会玉篇とを対照し一覧した表であるが、それまでに報告された逸文の有無と所在が検索可能となっている。その後、玉篇逸文に関しては、上田正「玉篇逸文論考」（『訓点語と訓点資料』第七十三輯、1985年）と同『玉篇反切総覧』（私家版、1986年）が出て、玉篇反切の総合的考察を達成したが、「玉篇逸文論考」で義注を含めた

逸文の集大成は井野口孝に依頼したい旨の発言がなされた。井野口孝は、「法進『沙彌十戒并威儀経疏』にみえる『玉篇』佚文について」（『京都府立大学学術報告人文・社会』53、2001年）などの成果を積み重ねている。

高山寺本の篆隸万象名義は、中国語音韻史、日本の辞書史・学問史を研究する上で大きな価値を有するものであるが、誤写が少なくない。白藤礼幸「篆隸万象名義解説」（『高山寺古辞書資料第一（高山寺資料叢書第六冊）』東京大学出版会、1977年）が

「万象名義」の写本としての価値は、誤写の多いことで若干割引かれねばならないであらう。漢字の字形について、現在の字形意識をそのまま「万象名義」にあてはめる訳には行かないが、それにしても誤写と考へられるものが多い。しかし、高山寺本は、唯一の写本であり、本書の原となつた原本玉篇を大部分佚してゐる現在、この「万象名義」は極めて貴重である。ただその利用においては、本文に対して慎重な検討が必要である。

と述べて解説を締めくくり、高田時雄「篆隸萬象名義解説」（前掲）も解説の最後に

しかし、どのような利用が為されるにせよ、この写本は、楊守敬が早くに「此の書を鈔せし者、草率の極みにして、奪誤満紙なれば、此は則ち小学に深き者の理董を待つ有らざる能わず」と指摘したように、「玉篇」から引用するに際しての誤字、脱字がすこぶる多く、このことが資料としての「萬象名義」の弱点になっている。有効な利用の前提としては、「玉篇」残巻や逸文の存するものは当然これらを利用し、また「説文」を初めとする中国字書に助けを求めながら、これらの誤謬を正す厳密な校定の作業が必要であらう。

と述べるように、厳密な校訂が大きな課題である。しかし、誤写の判断はなかなか難しい。

回 篆隸万象名義データベースの作成とその問題点

篆隸万象名義の本文研究の成果は、先に掲げた宮澤俊雅「掲出字一覧表」と上田正『玉篇反切総覧』の二つの業績に加えて、白藤禮幸「掲出字索引」・同「校勘篇」（『高山寺古辞書資料第一』前掲）の成果に集約される。筆者は、約十年前、これらの成果を踏まえた篆隸万象名義データベースを作成することを思い立ち、一応の完成を見た。池田証寿「篆隸万象名義データベースについて」（『国語学』第百七十八集、1994年）^[5]は、その報告であり、およそ次のように述べた。

このデータベースは、篆隸万象名義の本文研究を主要な目的とするものであり、篆隸万象名義に掲出される一万六千余字の所在、諸橋「大漢和辞典」番号、玉篇卷数・部首番号、JIS 区点番号等の情報を得ることができる。

作成上問題となるのは、第一に JIS に無い字の扱い方。JIS に無い字は篆隸万象名義の掲出字の七割、ISO/IEC 10646-1 でさえ三割に及ぶ。第二に掲出漢字の同定。これは、篆隸万象名義の本文研究に役立つようなデータの扱い方が必要である。

漢字字書データベースとして見た篆隸万象名義データベースの特色は、漢字字書の採録字やその配列の規準を解明するのに有効であることと、JIS や ISO/IEC 10646-1 等の漢字コード表に採用された漢字の典拠を示すことである。

この論文では、篆隸万象名義第一～四帖（目録部分を含む）の範囲の掲出字について、JIS と ISO/IEC 10646-1:1993 の漢字でどれだけ処理できるかを調査した結果を示した。現在は、篆隸万象名義全六帖のすべての掲出字について対応を調

査し終えているので、それを表 1 に示そう。JIS X 0208 で篆隸万象名義全体の約三割、ISO/IEC 10646-1 で約七割が処理可能であることが分かる。

規格	篆隸万象名義
JIS X 0208-1990	5125/16901(30.3%)
JIS X 0212-1990	3511/16901(20.8%)
ISO/IEC 10646-1	11252/16901(66.6%)

表 1 JIS X 0208 等で処理可能な篆隸万象名義の漢字数

この篆隸万象名義データベースは、1994 年に一応の完成を見たもので、暫定版として公開してきたが、現時点で見ると、さまざまな課題が残る。

- (1) JIS X 0208 に無い漢字は、諸橋番号を示すのみで、具体的な字形を示していない。
- (2) 原本系玉篇残巻、原本系玉篇逸文及び大広益会玉篇の所在情報の入力がされていない。
- (3) JIS X 0208:1997 による包摂規準と文字同定の成果が参照できていない。
- (4) Unicode で新たに規定された拡張漢字 A 及び B が参照されていない。
- (5) 反切・義注のデータベースをどのように構築していくか、検討していない。

まず、(1) は、二戸麻砂彦「パソコンにおける漢字処理／試論」（『山梨県立女子短大紀要』28号、1995年）、同「国語研究と JIS 外漢字について」（『国語研究』59号、国学院大学、1996年）^[6]、瀬間正之「『原本系玉篇佚文集成データベース』の構想と今昔文字鏡」（『文字鏡研究会会報』第二号、1999年）^[7]などで指摘された問題である。

私のとった JIS 外字の処理は、次のような方法である。批判は、要するに分かりにくいという点に尽きる。

掲出字のうち JIS 漢字に無いものは以下の方法で示す。

M2746600, 見 χ	糸偏に見
M3545200, 洲 ω	言偏に州
M3542500, 察 λ	察の古字（通用・譌字等）

M0487300, 経 μ 経の傍の部分の字
 M2273500, 百 2 百を二つ並べた字
 M1082600, 心 3 蕊の草冠の無い字
 M2806300, ▼ 以上で駄目な時

この方法は、かつて岡田希雄が「類聚名義抄の研究」（一條書房、1944 年）で難字を×（糸偏に見）のように示したことや、豊島正之「JIS に無い字」をめぐって」（しにか 3-2、大修館書店、1992 年）に示された考え方を参考にしたものである。

たとえば、M2746600 は、諸橋番号の 27466 番の漢字であることを示す。この番号の下 2 桁は諸橋『大漢和辞典』（大修館書店）に掲載のない漢字のための予備として設けたものである。

そもそも、私の「篆隸万象名義データベースについて」という論文は、平成 5 年度国語学会秋季大会（北海道大学、1993 年 10 月 30-31 日）での研究発表会分科会 C セッション「電子化テキストの国際的共有」（コンヴィナー：豊島正之）で口頭発表された四つの論文

- ・豊島正之「電子化テキストの国際的共有」
- ・家辺勝文「電子化テキストと書物との接点——テキスト・データとページ構成——」
- ・池田証寿「篆隸万象名義データベースについて」
- ・芝野耕司「国際符号化文字集合（ISO/IEC 10646-1）の JIS 化と今後の拡張」

の一つであった。これらの内容は『国語学』第百七十八集に収められている。この研究発表会では、国際符号化文字集合（ISO/IEC 10646-1）の JIS 化が議論の中心で、芝野耕司の熱弁に圧倒される国語学者が多かった。その席で芝野は、ISO/IEC 10646-1（≡ Unicode、以下では単に Unicode として説明していく）が実用されるようになる時代はすぐそこに来ているという見通しを述べていた。そうした事情もあり、Unicode で使えるようになる 20,902 字の統合漢字について、わざわざ苦労して作字する必要はないし、諸橋番

号さえ入力しておけば変換は容易だから諸橋番号の入力ミスを防げる程度の精度でよいだろうと判断したのであった。Unicode の実用はやや遅れたとの印象を私は持ったが、これは情報処理のエキスパートと、一般の、人文系のユーザとの認識の差なのだろう。

その後、Windows 2000 と Windows XP の普及により、Unicode を日常的に使用するレベルに到達した。Unicode の漢字の拡張はその後進められ、拡張漢字 A として 6,582 字、拡張漢字 B として 42,711 字の追加があり、総計 70,195 字の漢字が使用可能となった。

Unicode に ない 漢 字 は、1999 年 1 月 に Windows 用の TrueType Font を備えた今昔文字鏡 ver2.0 の刊行があり、単漢字 10 万字を中心に約 12 万字の文字が使用可能となっている。篆隸万象名義の全掲出字を処理する多漢字環境は整ったと言える。しかし、後述するように、篆隸万象名義の掲出字であって今昔文字鏡に未登録の漢字は少なからず存している。

(2) は、原本系玉篇残巻、原本系玉篇逸文及び大広益会玉篇の所在情報の入力となされていないという問題だが、入力の難易度から言えば、原本系玉篇残巻＜大広益会玉篇＜原本系玉篇逸文の順に困難の度合いは大きくなるであろう。原本系玉篇残巻は、特定の巻に偏るから入力は容易である。大広益会玉篇の所在情報は、大広益会玉篇と篆隸万象名義との配列順に一致しないところがあり、徹底網羅するのは容易ではない。原本系玉篇逸文は、瀬間正之『原本系玉篇佚文集成データベース』の構想と今昔文字鏡（前掲）がその問題点を説くように、極めて多数の典籍文書に原本系玉篇が利用されており、新出の逸文の報告もある。しかも、被注字の同定には慎重を期すことが必要であって、これを徹底網羅するのは、相当に困難であると考えられる。筆者は大広益会玉篇の所在情報の入力を最優先して行い、一応の作業が終了したので、その結果をのちほど報告することにした。

(3)(4)(5) の問題は、ほとんど手付かずであり、今後の大きな課題であるが、(3) の JIS X

0208:1997 による包摂規準と文字同定の成果は参照できていない点について、のちほど簡単に触れておくことにしたい。(4) と (5) は作業時間と作業量の問題が大きいが、方法論的な問題は少ないであろう。

回 今昔文字鏡への対応

私が現在使用している今昔文字鏡は、version3.0 の単漢字 10 万字版である。篆隸万象名義の全掲出字 16,901 字のうち、今昔文字鏡に未登録の漢字は、46 字であった。この 46 字のうち、36 字は大広益会玉篇に見えぬ漢字である。一例を挙げて見よう。

図 1 には篆隸万象名義（高山寺本）の第六帖 135 丁裏の写真を掲げた。上段の五行目に見える「𠂔」が文字鏡に登録されていない。この字を原本系玉篇では「𠂔」に作っているが（図 2 参照）、大広益会玉篇には見えない（図 3 参照）。

図 4 は宮澤俊雅「掲出字一覧表」（前掲）の該当部分だが、ここでは「𠂔」と判読している。掲出字の上に付された○は、原本系玉篇残巻が存すること、△は大広益会玉篇に存しないことを示している。

篆隸万象名義の全掲出字のうち今昔文字鏡に未登録の漢字は 46 字であり、この 46 字のうち、36 字は大広益会玉篇に見えず、10 字が大広益会玉篇に見える漢字である。大広益会玉篇にあっても、文字鏡にない漢字がある、つまり諸橋『大漢和辞典』にない漢字があるというのは意外であった。字形は省略し、篆隸万象名義と大広益会玉篇の所在を記しておく。

1/015+34	c003b073
1/023-54	a009a052*
2/34-42	a059a082*
3/72+22	b007a103*
5/032+41	b055a043
5/032+53	b055a042
5/036-33	b056b082
6/050+22	c030b022*

6/102+12	c046b091*
6/166-62	c068b103

たとえば、最初の行の 1/015+34 は篆隸万象名義の第一帖 15 丁 3 行 4 字目を示す（+ は表、- は裏の意）。次の c003b073 は、大広益会玉篇下巻 3 丁裏 7 行 3 字目を示す（上巻を a、中巻を b、下巻を c とし、中ほどの a と b で表裏を示す）。大広益会玉篇の所在の後に付した * は、同定に問題のあるもので、事実上△としてもよいものである。

次の段階としては、Unicode の拡張漢字 A、拡張漢字 B によって、篆隸万象名義の全掲出字の何パーセントが処理可能であるかを調査すべきかと思うが、果たしていない。

回 大広益会玉篇との照合

大広益会玉篇（宋本玉篇）は、澤存堂本（『大広益会玉篇』中華書局、1986 年）を用いることとし、宮内庁書陵部の宋版も適宜参照した。

大広益会玉篇の所在は、前述した方式で入力していたが、これは極めて困難な作業であった。篆隸万象名義と大広益会玉篇とは、配列の順序にずれがあり、また、字形の差をどう処理するかが難しい。

ひとまず調査し得たところでは、篆隸万象名義にあつて大広益会玉篇に見えぬ掲出字は、篆隸万象名義の全掲出字 16,901 字のうち 481 字であった。この 481 字のうちの 99 字は、宮澤俊雅「掲出字一覧表」（前掲）で大広益会玉篇にあるとされている字であり、今回の筆者の作業によって、大広益会玉篇にないことを確認したものである。つまり、宮澤俊雅「掲出字一覧表」（前掲）で△を付け落としたことを確認したものであつて、篆隸万象名義の本文研究を一步前進させたと思ふものである。

一方、大広益会玉篇にあつて篆隸万象名義にない掲出字がどれくらいあるかも興味を引く問題である。しかし、正確なデータは持っていない。

大広益会玉篇の全掲出字 22,804 字を諸橋

熙字典体に直して、それを活字化する方法がとられていた。Unicode や今昔文字鏡において、拡張の作業が進められており、古版本・古写本に見える俗字・古字などの異体字を処理できる日も近い。しかし、これは喜んでばかりいられる事態ではない。古版本・古写本の字体を活字体として翻刻するための一定の規準を設けなければ、「見たままの活字化・符号化」が横行し、結果的に、学術的な資料として使えなくなることを警戒しなければならない、と考えるのである。

回 今後の課題

現在の篆隸万象名義データベースは Microsoft Excel で操作している。今昔文字鏡のフォントを組み込んであるので、JIS にない漢字も表示されており、操作が非常に容易になった。Unicode への対応も課題なのだが、Unicode で篆隸万象名義のすべての掲出字を処理できる見込みはない。Unicode を使うとしても、結局は今昔文字鏡との併用となってしまう。現在の掲出字の表示は、JIS 漢字（JIS X 0208:1997）にあるものはそれを用い、ないものだけを今昔文字鏡で処理している。これが JIS X 0208:1997 に加えて、Unicode さらに今昔文字鏡となると、データの管理がより面倒になるのでは考え、Unicode 対応を躊躇している。

もっとも、篆隸万象名義や大広益会玉篇にあって、今昔文字鏡にない漢字は確実に存在するのであるから、Unicode の拡張に備えて、データを整備するのが急務であると言えるであろう。

注

- [1] なお、高田時雄「玉篇の敦煌本」（『人文』33、京都大学教養部、1987 年）は、従来未紹介の敦煌本の報告である。
- [2] <http://www.zinbun.kyoto-u.ac.jp/~takata/Banshomeigi.pdf> でも閲覧できる。
- [3] 貞苺伊徳『新撰字鏡の研究』（汲古書院、1998 年）に再録。
- [4] <http://member.nifty.ne.jp/shikeda/930731.html> でも閲覧できる。

- [5] <http://member.nifty.ne.jp/shikeda/kkg178.html> で閲覧できる。
- [6] <http://www.yamanashi-ken.ac.jp/%7Emnito/Research.html> 参照。
- [7] <http://www.mojikyo.org/html/institute/kaiho/02/020406j.htm> 参照。
- [8] 笹原宏之「JIS 漢字と位相」（『日本語学』12 巻 8 号、明治書院、1993 年 7 月）など参照。
- [9] これに関しては、次の研究が参考になる。横山詔一・笹原宏之・野崎造成・エリク＝ロング『新聞電子メディアの漢字——朝日新聞 CD-ROM による漢字頻度表——』（国立国語研究所プロジェクト選書 1、三省堂、1998 年）、豊島正之「書評 横山詔一・笹原宏之・野崎造成・エリク＝ロング編『新聞電子メディアの漢字——朝日新聞 CD-ROM による漢字頻度表——』（『日本語科学』6、国書刊行会、1999 年）、高田智和「電子化辞書とねじれの漢字」（『計量国語学』第 23 巻第 5 号、2002 年）。

（付記）

本稿は、平成 15 ～ 16 年度文部科学省科学研究費補助金特定領域研究「東アジア出版文化の研究」の公募研究「『玉篇』写本・版本の漢字字体規範に関する研究」による成果の一部である。図版の掲載については、高山寺、宮内庁書陵部、東京大学出版会に格別のご許可を頂戴した。記して感謝したい。

『老子』 傳奕本来源考

テキスト処理による「項羽妾本」介在の検証

秋山 陽一郎（あきやま よういちろう）

回 はじめに

儒家以外の先秦諸子の中で、『老子』ほど長きにわたって様々なかたちで好んで読まれた書物はない。そのためか、伝承されている今本『老子』には各時代の様々な改変を被っている。たとえば字数だけを取り上げてみても、最も字数の少ない本で 4,999 字（後述の敦煌唐鈔本）、逆に今本中で最も字数が多いとされる本で 5,539 字（傳奕本）と、わずか 5,000 字という短編にもかかわらず 500 字以上も開きが見られる。

1973 年に中国湖南省の馬王堆 3 号前漢墓から、帛書——つまりシルクに書かれた、甲・乙 2 種類の『老子』が出土した時には、そうした後代のフィルタを幾枚も突き破った、漢人が見ていた『老子』の貌が明らかになったということで注目されたが、このとき、この帛書本との字句の近似性の面から注目されたテキストが本稿で取り上げる「傳奕本」である^[1]。しかし、この傳奕本については、帛書本との近似性が指摘されて重視されるようにはなったものの、その来源については不明な点が多く、傳奕本自体に分析と評価を加えた研究もほとんど見ない^[2]。本稿では、NGSM (N-Gram based System for Multiple document comparison and analysis) の手法を用いて傳奕本を他本と比較対照し、その特質を明らかにしつつ、傳奕本がいかなる本に基づいて制作されたのかという疑問に

ついて再評価したい^[3]。

回 傳奕本とは

いわゆる「傳奕本」というのは、唐の初め頃（7c 前半）の道士・傳奕が校訂したテキストで、現在、明の正統道藏纂字号に収録されている『道德經古本篇』のことをいう。傳奕の『老子』関係の著作については、『旧唐書』経籍志に「老子二卷、傳奕注」が、『新唐書』芸文志には「傳奕注老子二卷」と「傳奕老子音義（卷亡）」がそれぞれ著録されているが、実際に道藏本を見てみると、ごく簡単な音注がまれに散見するだけで基本的には無注本である。傳奕の注本については、残念ながらすでに亡びてしまっていて、今日では見るができない。

実はこの傳奕本の来源については、ほとんど判っていない。『旧唐書』・『新唐書』の傳奕の伝記を参照しても、彼が『老子』に注したことは言及されているものの、それ以上の踏み込んだ説明はされていないし、肝心の道藏の傳奕本にも参考とすべき序文・跋文のたぐいが見えない。唯一、この傳奕本の詳細について言及しているのは、道藏與字号に「紹熙二年（1191）」の上表と共に収録されている、南宋・謝守灝の『混元聖紀』（以下、『聖紀』）である。この書物は老子に関する伝承を詳細にまとめたもので、同じく南宋の彭耜の『道德真經集注雜說』（道藏長字号所収）ではこれを「老

君実録」と呼んで引用している。傳奕本の来源を窺う唯一詳細な資料なので、以下にその該当する文を引用しておく。本稿で扱う問題の焦点は、この謝守灝『聖紀』の記述の妥当性の検証にあると換言しても良い。

唐の傳奕、衆本を考覈し、其の字を勘数して云ふ、「項羽妾本は、斉の武平五年、彭城の人、項羽の妾の塚を開きて之れを得たり。安丘望之本(*)は、魏の太和中、道士 寇謙之 之れを得たり。河上丈人本は、斉の処士 仇岳 之れを伝ふ。三家の本（項羽妾本・安丘望之本・河上丈人本）、五千七百二十二字有りて、《韓非・喻老》と相参し、又た洛陽官本は五千六百三十五字有り、王弼本は五千六百八十三字、或ひは五千六百一十字有り、河上公本は五千五百五十五字、或ひは五千五百九十字有り」と。

これによれば、傳奕本は「項羽妾本」・「安丘望之本」・「河上丈人本」の三種類の本を基礎に、それを『韓非子』喻老篇中の『老子』句の引用部・「洛陽官本」・河上公本二種・王弼本二種によって対校しているということになる。このうち帛書本との関係で注目されているのが、北齊の武平五年に彭城の人が項羽の妾の塚をあばいて発見したという「項羽妾本」で、これが事実であるとすれば、帛書甲本とほぼ同じ鈔写年代のテキストを使っていることになる。しかし、この記述については幾つかの問題がある。第一に傳奕の時代から 500 年も隔たった謝守灝の『聖紀（老君実録）』にしは見えない記述をそのまま妄信することはできないこと、第二に「項羽妾本」の詳細（例えば竹簡・縑帛などの出土原本の書写媒体）に関する言及がないこと、第三に『聖紀』に記されている『老子』諸本の字数が、傳奕本も含めた現存諸本に比して明らかに多いこと。特に第三の点は、同じ唐代で 8c 前半の遺物である景龍二年に建てられた易州龍興観碑や、敦煌出土の唐鈔本に見られる字数が少ないという事象と全く相反している。

回 方法

本稿における『老子』諸本の NGSM は、以下の環境において行っている^[4]。

Windows XP Professional Service Pack 1 /
Cygwin v 2.125.2.10 / Active Perl v5.8.0
/ Unicode::String v2.07 / morogram-
0.7.1.pl(師茂樹氏作) / ngmerge.pl v1.1(近
藤泰弘氏作) / Microsoft Excel 2000

先年、齋藤正高氏は形態素解析を行う（単語・文節区切りを行って用例を集める）TFIDF を利用して王弼本『老子』を分析されたが^[5]、本稿では形態素解析を行わない（単語・文節区切りのための辞書に依存しない）n-gram モデルを利用して本文の分析を行う。これは形態素解析を行わないモデルの方が、表意文字を用い、また単語や文節を機械的に区切りにくい漢文資料に適しているという理由による。

また n-gram は網羅的に用例を抽出するため、検索漏れを回避することができる反面、大量のデータの中から有意なデータを選別する作業が必要になる。いかにして膨大なデータの中から有意なデータを探し出すか、また抽出するデータの恣意的な偏重を避けて客観性を保持できるかといったことなどは、分析結果およびその解釈に対する信頼性を示す上で重要な問題となる。この問題に対して、齋藤氏は TFIDF による「重みづけ」によってこれを行っているが、n-gram モデルをベースにした NGSM では、師茂樹氏が NGSM 出力結果から有意な用例を自動抽出する方法として、「1. 各語における偏差値を求め、2. 平均値からの距離（-50）を求め、それを合計し、3. 合計値の大きな順にソートする」という手順をスクリプト化して行っている^[6]。

本稿ではより多くの者が手軽に利用できる方法として、師氏の手法とは別に Microsoft Excel の「分散分析（繰り返しのない二元配置）」ツールを利用し、分散度が上位にある用例（全ての異本で

共起頻度が同一のものは、分散度は"0"となる)を重点的に抽出・検討してみた。この分散度は、字句の共起頻度の延べ数に対する分散度であり、本稿のように『老子』という同一の文献の異本比較には有効だが、異なる複数の文献を比較の対象とするケースにおいては、師氏の手法をとるか別の方法を取る必要がある^[7]。

回 比較に使用するテキスト

一般に『老子』のテキストは、河上公本・王弼本・玄宗御注本・傅奕本の四系統に大別され、これに馬王堆帛書本・郭店楚簡本などの簡帛資料が加わる。先にも述べたように、『老子』はそれぞれの時代ごとに様々な改変が加えられていることから、以下に本稿で傅奕本との比較に使用する諸本について、それぞれ簡単な予備説明を加えておきたい。(【】内は本稿中での略記。)

■ 湖南省長沙馬王堆3号漢墓出土帛書本【甲/乙】

1973年に湖南省長沙の馬王堆3号漢墓から出土した帛書(絹に書かれた)『老子』のテキスト。甲・乙二本あり、避諱や書体などから、甲本は前漢の高祖(在位前206―前195)在世中ないしはそれ以前、乙本は文帝期(在位前179―前157)ないしは、それより若干前の鈔写と推定されている。大きな特徴としては、徳経が前に、道経が後に置かれ、今本と上下篇の位置が逆転していることが挙げられる。漢代以前のテキストで目下唯一、上下両篇がほぼ完備しているものとして採用した。底本は『馬王堆漢墓帛書【壹】』(文物出版社、1980)を用い、適宜、高明『帛書老子校注』(中華書局、新編諸子集成、1996)を参照した。

■ 江戸明和七年宇佐美[㊦]水校刊王弼注本【王】

魏の王弼(226―249)注本の系統からは、王弼本中最善とされる宇佐美[㊦]水の校定にかかる明和七年(1770)刊本を採用した。わが国の江

戸時代に最も流行した『老子』は南宋・林希逸『老子^{けんさいくぎ}[㊦]義』だが、[㊦]水はいわゆる「物門八子」の一人で荻生徂徠門下の古文学派に属し、宋学色の強い『[㊦]義』とは全く校訂のコンセプトを異にしている。明和本の詳細については、波多野太郎「[㊦]水の老子の校訂と南郭」(『文化』(東北大)10-3、1942。のち国書刊行会刊『老子道德経研究』に再録)を参照されたい^[8]。

■ 唐景龍二年易州龍興觀道德経碑【景】

馬王堆帛書本を除けば、現存する『老子』の完足本中最古のテキストが、この唐の中宗の景龍二年(708)に建てられた碑文である。同じ易州龍興觀の開元二十六年玄宗御注道德経幢と混同されることもあるようだが、景龍碑は河上公本の系統に属する。後述するように、しばしば助字を省く傾向が見られ、北朝系の伝本に基づくとされる。底本には、通行する朱謙之『老子校釈』(中華書局、新編諸子集成)本を採用した。

■ 敦煌出土五千字無注本(S6453/P2589)【敦】

敦煌出土の所謂「五千字本」は、本文が4,999字に圧縮されているのが特徴だが、下巻の末尾に「太極左仙公(葛玄)序 係師(張魯)定 河上真人(河上公)章句」とあるように、これもまた河上公本の系統に属する本とされる。前後完備した完足本こそないものの、S6453(天宝十載(751)鈔写、章別字数明記)が第7章から第81章までを備えており、この本の不備を道経(第1章～第37章)を完備しているP2589(章別字数不記載)で補うことで81章全てをカバーすることができる(第1章～第21章までを備え、S6453と同じく章別字数を明記するP2329を補完テキストとして採用しなかったのは、P2589の方がより本文字句がS6453に近かったため)。底本には大淵忍爾『敦煌道経』(福武書店、1978)の図版を用いたが、書写体は必要に応じて適宜、正字体やしかるべき活字体に改めた^[9]。(※“S”はイギリス大英図書館に所蔵されている、いわゆる「スタイン本」の目録番号。“P”はフランスのパリ

国立図書館に所蔵されている、いわゆる「ペリオ本」の目録番号を意味する。）

■ 拋瀧川君山旧蔵古鈔本（武内義雄校定本）【武】

景龍碑や敦煌本をはじめとして、現存する河上公本の多くが北朝伝本の系統を引いた「語氣急」な本であるのに対して、日本に伝存していた南朝伝本の系統を引いた旧鈔本を底本とし、これを前掲敦煌本や景龍碑を中心とする諸本によって校定しているのが武内義雄氏の校定本である。本稿では『武内義雄全集』第六巻所収本を用いたが、周知の通り、岩波文庫にも収録されている。

■ 道蔵所収唐玄宗開元御注本【玄】

唐の開元二十年（732）に玄宗が手ずから『老子』に注を施して理身・理国を表明した、いわゆる「開元御注本」である。この本は、『老子』のテキストとしては、河上公本や王弼本をはじめとする当時あったテキストや注釈の長所を折衷した形になっているというが、例えば第18章末尾の「吾欲獨異于人而貴食母」句を「我獨異于人而貴求食於母」に作り、妄りに「求」・「於」の両字を加えたとするような批判もある（南宋・晁公武『郡齋讀書志』が指摘。因みに馬王堆帛書本も含めて、ここは他本そろって「貴食母」に作る）。本稿では通行する道蔵所収本を使用した。この御注本もまた敦煌唐鈔本や易州龍興觀の開元二十六年幢などがある、本来、別途専門的に分析を要するテキストである。

以上のほか、河上公本系統のテキストについては特に道蔵本【道】と四庫全書本【四】を比較対象として追加した。これは、異本が比較的豊富な河上公本によって唐以降の改変部分を把握するのが狙いである。また今回は傳奕本の分析が目的であるため、今本『老子』と分量・体裁を大きく異にする湖北省荊門市郭店戦国楚墓出土の竹簡本は比較の対象から除外した。

回 傳奕本の特徴——助字

傳奕本最大の特徴は、「現存老子中最も助字が多く句ののびた本」というところにある。そのことは以下に示す NGSM の結果によっても一目瞭然である。

	傳	王	景	敦	道	武	四	玄	甲	乙
也	62	18	1	0	14	28	11	11	174	178
矣	27	11	2	1	10	22	9	11	25	25
焉	11	9	0	0	9	16	8	3	11	10
乎	14	11	1	0	11	12	6	11	12	12
邪	3	0	0	0	0	1	1	3	0	1
耶	0	3	1	1	2	2	2	1	0	0
歟	2	0	0	0	0	0	0	0	0	0

上記分析結果で目を引くのは、(1) 帛書本・傳奕本の句末助字の用例が突出して多いこと、(2) 景龍碑・敦煌本の用例がほとんど皆無であることだろう。ことに顕著なのは(2)だが、景龍碑・敦煌本（唐代の河上公本）は他本に比べて明らかに助字の使用頻度が少なく、中でもそれが句末の助字において最も顕著に表れるという特徴がある。この現象は南北朝から唐代にかけての本文改変に起因するものと考えられている。武内義雄氏は、南北朝時代の顔之推の説を引いて、これを北朝系伝本固有の特徴であると論じている。

「『也』は是れ語已及び助句の辞にして、文籍備へて之れを有するなり。河北の経伝、悉く此の字を略し、其れ間々字の無きこと得べからざる者あり。」

（北齊・顔之推『顔氏家訓』書証篇）

このように、「也」などの句末の助字をことごとく省くのは、当時の北学（北朝系の学問）特有の傾向だったことが判っている。武内本の底本に使用されているわが国伝存の古鈔本は、武内氏によれば、こうした句末の助字が不用意に削られていない南朝系の伝本であり、河上公本には北朝系・南朝系の伝本に由来する詳・略二系統の本があっ

たことの証左としている^[10]。
その南朝系の古鈔本にも増して句末の助字——ことに「也」字が多いのが傳奕本である。帛書本が出土するまでは、「現存最古の『老子』のテキスト」と言えば、景龍二年に建てられた景龍碑を指したが、景龍碑を校勘した清の嚴可均が、

蓋し『道德経』は、御注より後、学宮に頒列して、久しく相伝習せり。故に余の見る所の『道蔵』七十余本も略ぼ同じくして、河上・王弼の二家を以てすと雖も、校者亦た頗る改めて御注に就く。而して傳奕の古本は、字句較々繁にして、亦た尽くは従り難ければ、則ち世間の真の旧本は、必ずや景龍碑を以て最と為さん。（『鉄橋金石跋』）

と評しているように、傳奕本は字句が繁多でその全てには依拠しがたいことから、景龍碑を『老子』の真の旧本に最も近いものとすべきだというのが、概ね帛書本が出土するまでの定評となっていた。この定評が帛書本の出現によって覆されたことは、帛書本の助字の使用頻度を見ても判る通りで、驚くべきことに、それまで助字が多すぎるとされてきた傳奕本の三倍近い頻度で「也」字が使用されている。なお、ここで嚴可均が『『老子』は玄宗御注が出されて以来、学宮に頒布されて久しく伝習されてきたため、『道蔵』に収録されている70余種の『老子』は、河上公本や王弼本にいたるまで、ことごとく御注本によって字句が改められている」と指摘しているが、それを裏付けるかのよう

回 「項羽妾本」 介在の可能性

以下に掲げたのは、秦漢期（劉向以前）の避諱字である。秦の莊襄王の諱が子楚、同じく始皇帝は政、前漢の高祖は邦、以下、恵帝が盈、文帝が恒、

景帝が啓、武帝が徹、昭帝が弗、宣帝が詢、元帝が奭、成帝が驚、哀帝が欣、平帝が衍である。『老子』の場合、書中で使用されている用例から考慮すれば、注目すべきは主に昭帝以前の諱字である。傳奕本が「項羽妾本」なる本に本当に依拠したかを検証するには、このような前漢の避諱字を検してみるのが最も手軽な方法である。つまり、もし前漢の皇帝の諱が避けられていなければ、それが前漢以前の本を元としている可能性が生じるわけである。

	傳	王	景	敦	道	武	四	玄	甲	乙
楚	0	0	0	0	0	0	0	0	1	0
荊	1	1	1	1	1	1	1	1	0	1
正	8	7	8	9	7	8	8	7	8	9
政	4	2	4	9	4	3	2	4	3	3
邦	4	0	0	0	0	0	0	0	6	0
國	24	28	28	28	28	28	28	28	5	29
盈	6	8	8	6	7	8	8	8	10	9
滿	3	1	1	3	1	1	1	1	0	0
恒	0	0	0	0	0	0	0	0	30	29
常	31	30	29	30	30	30	30	29	8	8
啓	0	0	0	0	0	0	0	0	3	3
開	3	3	3	3	3	3	3	3	0	0
徹	2	1	1	2	1	1	1	1	0	0
弗	1	2	0	0	1	2	2	0	39	41
毋	0	0	0	0	0	0	0	0	15	17
不	243	244	242	237	245	246	246	248	192	189

ただし、仮に前漢期の皇帝の諱が漢代に避けられていたとしても、その後の改変によって一旦避諱された字が復活してしまえば、前漢皇帝の避諱字検査の証拠能力はなくなってしまうという懸念がある。ところが上の結果を見ると、意外にも帛書本以外の全ての伝本が、文帝・景帝・昭帝の三帝の諱を避けていることが判る。これは少なくとも現存する今本がことごとく昭帝以後に鈔写された伝本に由来していることを示唆する結果であり、更に言えば、傳奕本が「項羽妾本」に依ったとするのに一見否定的な結果とも見てとれる。

しかし、帛書本以外の諸本の中で唯一傳奕本のみが高祖の「邦」字を避けきっていないのは興味深い事象である。先述した通り、もし本当に傳奕本が北齊の武平五年に項羽の妾の墓から出土した本に基づいているとすれば、鈔写年代は馬王堆の

帛書甲本とほとんど同じ時期となるが、上の表でも明らかのように、帛書甲本は高祖の諱「邦」字を避けていない。そこで、傳奕本の「國」・「邦」字の分布を調査してみると、一層興味深い結果が得られる。

- 國：10（章）, 18, 57, 59, 60, 61, 65, 78, 80
- 邦：36, 54

傳奕本中で「邦」字が使用されているのは第36章と第54章だが、上の結果から、ちょうどそれを挟むような形で「國」字が分布していることが判る。そもそも「項羽妾本」の墓葬時期とそれが出土した北齊の武平五年との間には800年近い時を経ていることを考えれば、竹書なり帛書なりが腐敗して断片化していない方が不自然である。あるいは傳奕が使用した「古本」は、第18章から第57章中の数十章分の断片で、それを元にその欠損部を傳奕が他本によって補ったとは考えられないだろうか。勿論、この「邦」字の分布だけでは、まだ「項羽妾本」演出の偽装工作の可能性も否定できない。先述のような、文帝・景帝・昭帝の避諱から今本が昭帝以降の伝本に基づくことを示す結果も出ているので、もう少し分析を加えてみる必要がある。

■同訓異字ならびに通仮字

ある書物の中に複数の個性が介在している場合——つまり複数の人間がその書物の著述・編纂ないし鈔写（刊刻）に関わっていることが想定される場合、各々の個性を明確に腑分けするのに、その書物がまとまった分量を有するならクラスタ分析を利用するところだが、ここでは、ただでさえ5,000字と分量が少ない『老子』を更に章単位で分析する必要から、単純に同訓異字や通仮字の検出行（≒章）の調査を行うことにした。しばしば、「X篇とY篇では共にA字に作るから、ここもA字に作るべきである」といったような校記を目にするが、この種の校訂は自戒をもって慎重に行わないと、いたずらに書物の内部に介在する個性を

永久に葬り去ってしまうことにもなりかねないとは、こうした作業を行う時に痛感させられることである。

まず以下に掲げたのは、NGSMのデータから、傳奕本中で併用されている同訓異字（語）や通仮字のうち、傳奕本と他本とで分散傾向が異なる用例を抽出したものである。既に分散分析を行っているので、こうした用例も分散度上位の例から容易に抽出できる。

	傳	王	景	敦	道	武	四	玄	甲	乙
居	7	10	10	8	9	7	10	9	25	23
處	20	16	15	18	17	19	16	17	2	4
喪	2	3	1	3	3	2	3	3	3	2
亡	5	4	4	5	4	4	4	4	4	4
吾	27	21	19	21	22	22	22	22	18	19
我	16	19	21	21	20	20	19	19	18	15
至	9	8	8	8	8	9	8	8	9	12
致	5	5	4	5	5	4	5	5	4	1
不道	2	4	2	0	4	4	4	4	4	4
非道	3	0	3	5	0	0	0	0	0	0

次にこれらの各例について、その分布状況を見ていくと次のようになる。

- 居：8（章）, 31, 34, 72, 77
- 處：2, 24-26, 30, 31, 38, 66, 76
- 喪：31
- 亡：33, 41, 44, 69
- 吾：4, 13, 16, 20, 21, 25, 29, 37, 42, 43, 49, 54, 57, 67, 69, 70, 71, 74
- 我：17, 20, 42, 53, 57, 70
- 至：43, 48, 55, 65, 72, 80
- 致：10, 14, 16, 39
- 非道：30, 53
- 不道：55

「居」・「處」や「吾」・「我」は混在していて、特に区別されている様子は窺い難いものの、そのほかの通仮・同訓異字の分布を見ていると、概ねA群：1～30近辺 / B群：30近辺～55近辺 / C

群: 55 近辺～81 の3つのグループに分けられそうである。「近辺」と曖昧な表現を使っているが、竹簡にせよ、帛書や紙の卷子本にせよ、朽ちて断片化してしまった媒体というのは、前後の失われた部分との間に明確な境界を定めがたいのが普通であり、必ずしもここで明確な章数の線引きをしてしまうのは妥当ではない。先刻の「邦」字は、第36章と第54章に見えるのでB群に相当することになるが、仮に「項羽妾本」なるものが実際にあったとすれば、このB群がそれに相当することになる。

ここでB群が「項羽妾本」に相当することを後押しする材料がある。南宋の晁説之^{ちやうまつし}の王弼注跋文と、同じく南宋の晁公武『郡齋讀書志』に、傅奕の『老子注』の佚文とおぼしき一文が引用されている。

傅奕謂ふ、「『常善救人、故無棄人。常善救物、故無棄物』の四句、古本有ること無し。独だ〔河上〕公より得るのみ」と。

(『郡齋讀書志』子部・道家類)

ここで引かれている四句は今本『老子』の第27章に当たるが、この四句が「古本」には見えなかったと傅奕は言っており(現行の傅奕本にはこの四句が既に挿入されている)、第27章が「項羽妾本」の残存部分に含まれていたことを示唆しているのである。「古本に有ること無し」というのが、出土原本の汚損によるものか否かはここでは明言されていないが、少なくとも帛書本、及び『淮南子』道応篇に「老子曰く」として引かれている文には、この句は見えている。ただ景龍碑や敦煌本も含めて諸本「(而 / 故) 無棄人」と作っているのを傅奕本が唯一「故人無棄人」に作っている点は『淮南子』と共通し、またその『淮南子』と帛書本に「常善救物」句が見えない点などは従来指摘されている通りである。いずれにしても、筆者はさきにB群は第30章近辺より始まると述べたが、ここで第27章までは残存していたことが具体的に明示されるわけである。

更に傅奕は「古本」に見えない句を河上公本

によって補ったと続けて言っているが、これは傅奕本の校訂姿勢を窺うことができる重要な記述である。つまり傅奕本は、異本照合によって互いの欠字を補っていくという方針で校定されたテキストだったことになる。謝守灝『聖紀』の「三家の本(項羽妾本・安丘望之本・河上丈人本)、五千七百二十二字有り」という記述は、3種類とも5,722字であったと見なすのはもとより無理があり、むしろこの3種類の本の欠落を互いに補ってはじき出した字数と見た方が妥当である。これは傅奕本が他の今本に比して、際立って字数が多いことに対する説明にもなるが、逆に言えば「項羽妾本」を底本にしていたとしても、必ずしもその古体の保存に十分な注意が払われなかったことをも意味する。或いは傅奕が「項羽妾本」を見た時には、既にこの本が紙媒体に転写された後で、必ずしも十分に原形を窺い知ることができない状態だったのかも知れないが、それを明らかにする手段は今のところない。

回 安丘望之本・河上丈人本

傅奕本『老子』を構成している3群のうち、中央のB群が「項羽妾本」を底本にした可能性については前項で指摘した通りだが、その前後に補填された本についても興味深い結果が出ている。既に示した同訓異字の傾向のほか、句末の助字にまで、「古本」を補填した部分であるはずのA群・C群の間に明らかな使用傾向の差違が見られるのである。以下に提示するのは、A・B・C各群における句末の助字の分布である。傅奕本における分布が他の諸本と異なることを示すため、武内本・道蔵河上公本・王弼本・御注本・帛書甲・乙本の分布も併せて示しておく。(以下、【傳】+【武】・【道】・【王】・【玄】・【甲】・【乙】の順でA・B・C各群に相当する範囲の共起分布を示す。単位は例。)

・也

A 8 (8, 4, 5, 4, 44, 49)

B 23 (8, 3, 8, 1, 58, 54)

C 31 (12, 7, 5, 6, 72, 75)

• 矣

A 6 (4, 0, 1, 1, 5, 6)

B 8 (6, 3, 2, 1, 5, 5)

C 13 (12, 7, 8, 9, 15, 14)

• 焉

A 6 (11, 7, 7, 2, 3, 3)

B 4 (3, 1, 1, 0, 6, 6)

C 1 (2, 1, 1, 1, 2, 1)

• 乎

A 9 (10, 8, 9, 8, 8, 8)

B 0 (1, 2, 2, 2, 0, 1)

C 5 (1, 1, 0, 1, 4, 3)

このように A 群と C 群は、各々別の本によって補填されている可能性が高い。ここで想起されるのは『聖紀』の「安丘望之本」と「河上丈人本」である。『聖紀』では、傳奕本がこれに「項羽妾本」をあわせた「三家の本」を底本としたことが記されているが、奇しくもこの記述と符節を合わせるかのような結果が出た訳である。ただ惜しむらくは、安丘望之本は現存せず、河上丈人本もまた河上公本との相違が判然としないことから、これ以上の追究は難しい。河上丈人は『史記』樂毅伝賛に「樂臣公、黄帝・老子を学び、其の本師、号して河上丈人と曰ふ」とあり、また『隋書』経籍志に「梁に、戦国の時の河上丈人注二卷有り。亡」と著録されている。しかし、本としての出自は定かでなく、西晋の皇甫謐『高士伝』（『太平御覧』巻 507・逸民所引）に「河上丈人…『老子章句』を著はす」と見えて以降にその存在が確認しうることから、河上公本同様、魏晋南北朝以降に定まったテキストである可能性がある。清の姚振宗は『隋書経籍志考証』（開明書店刊、二十五史補編所収）で「河上丈人」と「河上公」、さらには「安丘先生（『後漢書』耿弇伝。すなわち安丘望之）」と「安期生（河上丈人の門弟。樂毅伝賛に見えるほか、『高士伝』では蒯通とともに項羽に謁見したことも見

えている）」が混同されていた可能性を指摘している。だとすれば、傳奕の狙いは前漢より前の伝本の収集にあった可能性がある。

回 おわりに

本稿では、謝守灝『混元聖紀（老君実録）』中の傳奕本の来源に関する記述の妥当性を検証すべく、傳奕本をはじめとする『老子』諸本の NGSM データを起点に傳奕本の内部構造を掘り下げ、謝守灝『聖紀』の傳奕本に関する記述が必ずしも後人の創作にかかるとは言いつれないこと、また帛書本との近似性から意識されてきた「項羽妾本」とおぼしき「古本」が、実は傳奕本の中央部の限られた範囲にのみ認められること、更には中央部の「古本」を挟んで「安丘望之本」・「河上丈人本」に相当する二種類の『老子』の本が前後の補填に利用されていることなどを明らかにすることができた。紙幅の関係で必ずしも十全な議論を尽くしたとは言いがたいが、ともあれここまでを一応の成果としておきたい。

本稿で取り上げた『老子』の分析と評価は、傳奕本の腑分けを別にすれば、ほとんど既に明らかにされていることの再検証と言うべきで、実際、手作業でも出来る内容が大半を占めている。ただ、本質的な中国史や中国思想の研究は、こうした腑分けの作業の後にあるものであって、腑分けの作業自体にあまり多くの時間を費やしてしまうのは避けたいところである。筆者の場合、この腑分け作業にコンピュータを利用したテキスト処理の需要の一つがあると考えている。本稿で示した腑分けの方法は、一時一手に成らずとされる先秦古文獻そのものの腑分けにも一定の応用が利くものと思うが、それはひとまず後日の課題としておく。

注

- [1] 例えば、「座談長沙馬王堆漢墓帛書」（『文物』1974-9）の張政娘氏のコメントなど。
- [2] 傳奕本を底本にして他本と校訂したものに清・畢沅の経訓堂叢書本があり、また傳奕本の異文考証を行った

- ものに近人馬叙倫『老子校詁』（中華書局）がある。
- [3] N-gram 及び 石井公成氏提唱の NGSM については、本誌第 2 号（2001）の特集「N-gram が開く世界」を参照されたい。
- [4] 各ツールの入手先は以下の通り。
- Cygwin（Windows 上で UNIX API 環境を構築。）
<http://cygwin.com/>
 (Unicode::String モジュールの組込みは Cygwin 同梱の CPAN を利用。)
 - Active Perl
<http://aspn.activestate.com/ASPN/Downloads/ActivePerl/>
 - 師茂樹氏作 "morogram" (n-gram 解析用スクリプト)
<http://www.ya.sakura.ne.jp/~moro/resources/ngram/>
 - 近藤泰弘氏作 "ngmerge" (マージスクリプト)
<http://klab.ri.aoyama.ac.jp/tool/>
- なお、NGSM 環境の導入と利用方法については、山田崇仁「初めての N-Gram —— Cygwin もしくは Perl を用いて」（本誌第 2 号、2001）及び、山田氏の Web サイト (<http://www.ritsumei.ac.jp/kic/~tyv07679/>) を参照されたい。
- [5] 齋藤正高「TFIDF の漢文テキストへの応用 ——『老子』王弼本を例に」（本誌第 3 号、2002）。
- [6] 師茂樹「N グラムによる比較結果からの用例自動抽出 —— 実例に即して」（『東洋学へのコンピュータ利用 第 14 回研究セミナー』、2003）。
- [7] 本稿で使用した NGSM データは紙幅の都合上、掲載できなかったが、代わりに筆者の Web サイト上にデータを公開して読者諸賢のご批正を乞うこととしたい。
http://www.karitsu.org/studies/laozi_ngsm.htm
- [8] 王弼本のテキストデータは、筆者の Web サイトに公開しているものをういている。
http://www.karitsu.org/sosho/laozi_etext.htm
- [9] 敦煌本『老子』については、一本にとどまらないため、幾つか代表的な参考書を挙げておく。
- 王重民『敦煌古籍叙録』（中華書局、1979）
 - 姜亮夫『巴黎所藏敦煌写本道德經殘卷叙録』（『中国哲学』1980-2）
 - 姜亮夫『巴黎所藏敦煌写本道德經殘卷叙録』（『雲南社会科学』1981-2 / 3）
 - 程南洲『倫敦所藏敦煌老子写本殘卷研究』（文津出版社、1985）
 - 大淵忍爾『敦煌道經目録』（法蔵館、1960）
 - 大淵忍爾『道教典籍目録・索引』（国書刊行会、1988 / 1999 増改）
 - 楠山春樹等『敦煌と中国道教』大東出版社、講座敦煌、1983）
- また敦煌学の分野は、本誌第 3 号で岩本篤志氏が紹介している『『俄藏敦煌文献』収載文献データベース』も含め、既に優良なサイトが幾つかある。本稿執筆に際して活用したものとしては、以下のサイトがある。
- 漢学研究中心『敦煌学研究論著目録資料庫』
<http://ccs.ncl.edu.tw/>
 - 大英図書館 International Dunhuang Project（国際敦煌学項目）
<http://idp.bl.uk/>
- [10] 武内義雄『老子原始』及び『老子の研究』（ともに『武内義雄全集』第 5 巻所収）。

特集 1

人文学 情報処理教育

本誌前号で組まれた特集「人文学情報処理教育」は、幸い少なくない読者の関心を引いたようである。人文系における情報教育の重要性は、本会がこれまで『電脳中国学』等を通じて主張してきたことであるが、前号の特集においてそれがひとつのまとまった形で問題提起をされたことは、議論の範囲や方向性が明確化されたという点で大きな一歩であったのではないかと思う。今後はここで提案された課題に対してひとつひとつ議論していき、ある程度のディシプリンと呼べるものにまでもっていくことがひとつの目標になるのではないかと思う。

今号では、学科・コース単位で取り組んでいる事例報告を2点掲載するほか、前号で提起された教員のスキルに関する問題を再びオンライン討論において取り上げる。前者は学生の評価、後者は教員の評価というそれぞれ重いテーマについて論ぜられており、今後の議論に極めて有益であろう。

Contents

文学部における情報処理教育

- 大谷大学人文情報学科における試み——……………松川 節… 22
- 花園大学「情報歴史学コース」の現状と課題……………佐藤 誠・田中 正流… 28
- オンライン討論 人文学情報処理教育における教える側の問題……………32

文学部における情報処理教育

——大谷大学人文情報学科における試み——

松川 節（まつかわ たかし）

◇ はじめに

2000年4月、大谷大学文学部^[1]に人文情報学科^[2]が設置された。文学部における情報学科としては日本で最初の試みであり、文字通り「人文学情報処理教育・研究」を行ないつつ、今年度で完成年度を迎えている。ここでは設置時から関わる一教員として、この4年間の試みを回顧しつつ、今後の展望を述べてみたい。

◇ 設置の意図

大谷大学文学部が既存の7学科^[3]に加えて人文情報学科を設置した意図は、情報化社会が高度に発展していく中で、どんな情報を発信するべきか、その中で人間はどうあるべきかを問い直すために、コンピュータを創造的に利用する技術の修得をめざすというものである。社会に氾濫する様々な情報を取捨選択し、対象化した上で再発信するためには、単に情報処理技術を身につけるだけでは不十分であり、長年にわたって蓄積されてきた人文学の伝統を生かしてこそ意味をなすのではないか。そこには、人文学における情報処理とは、いかなる形態で何を目指すべきものかという根源的問いがあった。

◇ カリキュラム

学生は、情報処理技術を修得するとともに、人文学的課題を発見し、修得した技術をもとにデジタルコンテンツ作りに取り組むことになる。そのために次のカリキュラムが用意された。

1年次の「基礎演習」（週2コマ）では、前期にWindowsベースのいわゆる「情報リテラシー」（タッチタイピング、ワープロ、表計算、メール、Webを利用した情報検索、データベースなど）を徹底的にマスターし、パソコンを「文房具」として自在に操れるようになる。しかしGUIだけではOSの何たるかを理解しがたいので、後期はUNIXベースでコマンドラインに習熟し、HTMLやPerlなどのプログラミング言語の初歩を体得することによって、ネットワーク管理を踏まえた情報発信を実習する。この方法によって、1年次修了時には複数OSにおけるデータ共有が可能になり、ネットワーク上の様々なリソースを利用するための基礎が築かれる。端的に言えば、ネット上にあるCSV形式の統計データを加工してHTML形式に変換し、UNIXサーバにFTPで転送し、CGIでアクセス制限をかけたWebページとして提示するといった、情報の検索→加工→再発信の方法を身につけるのである。同時に、通年の「ゼミ演習」ではリレー講義形式で他学科の教員よりそれぞれの学問分野の概要・研究方法の講義を受け、人文学の全体像をイメージした上でその

情報化を考えるきっかけとしている。

2年次においては、情報化社会に内在する様々な問題を自分の力で探し出し、テーマとして設定し、評価・解決してゆく力を養成している。具体的には、4～5人の小グループ単位でテーマを設定し、その解決に向けて協同作業を行ない、PowerPoint を使って成果を発表するという演習である。昨年度は「仏教」という統一テーマを課したところ、「仏教文献デジタル化のメリット」、「声明の人文情報学的解析」など、そのまま卒論の研究テーマにつながりそうなものが見られた。また、2年次ではSQLの理解と修得を目的とした「データベース演習」が必修単位となっており、人文学情報をデータベースによって階層化しつつ蓄積してゆく技法の修得を義務づけている。

3・4年次においては、学生各自の関心に合わせて専門のゼミクラスに分かれる。開講されているゼミは表1のとおり。このなかで、筆者が担当する「マルチメディア民族誌と仮想博物館」ゼミの今年度3回生前期シラバスを表2に掲げる。

表1 3～4回生ゼミクラス・テーマ一覧

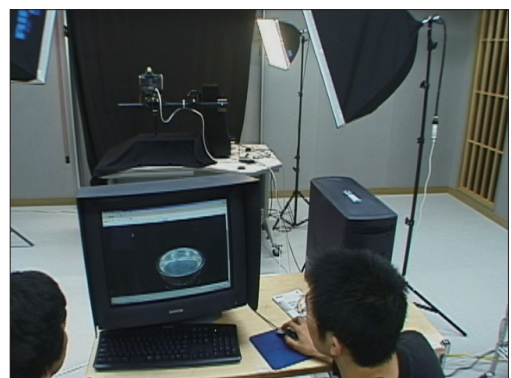
仏教学とコンピュータ
マルチメディアを利用した表現技法
情報化社会とコミュニケーション
デジタルアーカイブと電子図書館
マルチメディア民族誌と仮想博物館
情報倫理
人文学のためのデジタル・ツールの製作

表2 「マルチメディア民族誌と仮想博物館」ゼミ
2003年度前期シラバス

第1回 (4月14日)	講義概要の説明・グループ別テーマの実例提示・ネットワーク設定確認
第2回 (4月21日)	「葵祭」プロジェクト発動／ネットワークインフラの立ち上げ
サブゼミ1 (4月22日)	「葵祭」プロジェクト（ビデオ撮影入門）
サブゼミ2 (5月6日)	「葵祭」プロジェクト（企画書プレビュー）
第3回 (5月12日)	「葵祭」プロジェクト（取材準備1）
サブゼミ3 (5月13日)	「葵祭」プロジェクト（取材準備2）

——5月15日: 葵祭当日(取材)	——
第4回 (5月19日)	「葵祭」プロジェクト（動画処理）
サブゼミ4 (5月20日)	『魏志倭人伝』電子テキストの収集
第5回 (5月26日)	「私の仮想博物館」プロジェクト発動／国立民族学博物館見学会の準備
サブゼミ5 (5月27日)	『魏志倭人伝』電子テキストの文字鏡フォント埋め込み
第6回 (6月2日)	民博見学会の結果報告 & HTML化の指針
サブゼミ6 (6月3日)	『魏志倭人伝』テキスト講読1
第7回 (6月9日)	QTVRを利用した擬似動画撮影実習1（パノラマ・ムービー）
サブゼミ7 (6月10日)	『魏志倭人伝』テキスト講読2
第8回 (6月16日)	QTVRを利用した擬似動画撮影実習2（オブジェクト・ムービー）
第9回 (6月23日)	VR-WorxによるQTVRオーサリング実習
第10回 (6月30日)	葵祭プロジェクト成果報告会1
サブゼミ8 (7月1日)	葵祭プロジェクト成果報告会2
第11回 (7月7日)	「私の仮想博物館」プロジェクト（中間報告）
サブゼミ9 (7月8日)	『魏志倭人伝』テキスト講読3
第12回 (7月15日)	「私の仮想博物館」プロジェクト（成果達成状況の確認）
サブゼミ10 (7月16日)	『魏志倭人伝』テキスト講読4

写真1 QTVRオブジェクト・ムービー撮像装置



以上が学科の基幹となる4年間の演習形態であるが、このほか学生は関心に応じて他学科の開講課目を履修することもできる。さらに表3のように、斯界を代表する専門家に人文学情報処理研究及び情報メディア論の最前線を教授していただいているのが本学科のもうひとつの特色である。

表3 非常勤講師による情報処理教育課目例

- ・東洋学とコンピュータ（アジアの暦法とコンピュータ）：矢野道雄（京都産業大学）
- ・印刷文化とDTP：中西秀彦（中西印刷）
- ・ハイパーメディアデザイン入門（Macromedia Director を利用したオーサリング（制作・統合）：齋藤朋彦（IAMAS）
- ・Photoshop を利用した画像処理とビジュアルライゼーション：伴宙（BAN ファクトリ）
- ・情報文化史：山田奨治（国立日本文化センター）
- ・多言語情報処理論：安岡孝一（京都大学人文科学研究所）

◇ 成果と評価

今年度で完成年度を迎える本学科の成果は、明年1月に提出される卒業論文^[4]をもって問われるべきであろうが、中間報告として以下の点が指摘できよう。

写真2 Director による仮想博物館コンテンツ



まず、文学部における規定により、卒業論文は文字媒体（2万字程度）での提出が義務づけられ、デジタルコンテンツなどの作品は「附録」と位置づけられる。デジタル作品が評価される以前に「学術論文」として成立していなければならないわけで、情報処理能力とともに、日本語そのもののリテラシーや論文執筆のための論理的構想力・表現力が要求される。これは文学部として当然のことと思うが、かえって学生にはプレッシャーになっているようで、3年次の後期になって卒論を意識するようになると、ネットを流れる情報をコピーするだけではとても論をたてられないことに気づき、あわてて情報処理学会研究報告『人文科学とコンピュータ』誌バックナンバーなどを漁ってテーマ探しをはじめるとというのが現状である。

その一方、いわゆるマルチメディアを利用した情報収集・コンテンツ作成・再発信については、期待した以上の成果が見られている。例えば、大谷大学博物館^[5]に所蔵される重要文化財を紹介する仮想博物館立ち上げを企画し、標本の提示には QuickTime VR^[6] 擬似動画像を活用して 360 度全周映像を実現し、コンテンツを Web で配信するとともに Director でオーサリングして CD-ROM で配布する、もちろん解説文も自分で書く——といったことを独力で達成できる学生が育っている。こうした表現能力は、この社会における大学への期待、つまり「社会に貢献する人材の育成」を踏まえれば、文学部における「人文学+情報処理」教育の最も現実的な成果といえるかもしれない。

さて、本誌が扱う「漢字文献情報処理」に関しては、第1に、筆者の3年次サブゼミで『魏志倭人伝』を講読した試みを挙げてみたい。学生の半数は高校くらい漢文など一文字も読んでいないというので、台湾中央研究院漢籍電子文献データベースから GB コードのテキストをダウンロードして整形し、同じくネット上に公開されている訓読文、現代日本語訳、訳注を段落ごとに切りだして対応させ、訓読文を参照しつつ Word で返り点を振らせ、漢文訓読のしくみを理解させてみた。また、図書館で百衲本テキストのコピーを取り、

電子テキストで欠けている文字を照合して探しだし、文字鏡フォントを埋めこませた上で本文の読解に入った。結果として漢文の読解力がついたわけではないが、作業の過程で漢字の異体字や文字コードの問題を考えるきっかけが与えられたこと、そして何よりも漢文の読解に日本語のリテラシーが要求されることを理解させられたことは収穫であった。本来の史料講読とはまったく逆順であるが、文学部にいながら漢文史料とまったく接点を持たない学生には有効な方法といえよう。

第2に、史学科の学生に対して本研究会(編)『電腦東洋学Ⅱ』をテキストにして「漢字文献情報処理入門」という演習(表4)を行なったところ、日頃から漢文に親しんでいる学生だけあって反応もよく、「台湾中央研究院漢籍電子文献を利用して、中国史上「太宗」という廟号をもつ皇帝を古い順に全員挙げよ(王朝名と諱で答えよ)」という課題に対しては、二十五史データベースを活用しつつ、検索結果を『東洋史辞典』で二重にチェックして回答するものも見られた。しかし一方で、文字コード規格の変遷に関する説明に対して関心を示したのは一部の学生だけだった。

この例から、歴史学のような文献学的メソッドが確立されている学問分野ほど情報化に対する親和性が高いことがわかった。またこれと関連して、非常勤先の大学史学科における漢文講読演習

でモンゴル史関係の漢文史料『蒙鞮備録』を講読し、史料の白文テキスト(約4,000字)のWordファイルを配布して、訓読文・現代語訳・訳注をつけて提出するという期末レポートを課したところ、内容の理解度に差はあるものの、ほぼ全員が完成させて提出した。漢字文献情報処理技術は、OSの進化とも相俟って、よい意味でのブラックボックス化が進んでおり、学生が人文学情報処理の醍醐味を感じることでできる最適な教育手段であるといっていよいだろう。

◇ 研究プロジェクト

次に、教育の延長上にある人文学情報処理研究に関して、本学における取り組みを2つ紹介したい。

■ 企業との合同によるパッド型携帯端末コンテンツの開発プロジェクト

(株)三洋電機からパッド型携帯端末「スクリオ」の提供を受け、そのコンテンツを開発する演習授業の中から学生による特別チームを発足させ、全国学会での発表を目指すプロジェクトである。昨年度は、小学校の「総合的学習の時間」に焦点を合わせ、小学生がリバーウォーク(川歩き)をしながら得られる動植物や環境汚染といった実体験(=「体験的知識」)を、携帯端末のデータベースを検索して参照することにより、「抽象的知識」

表4 「漢字文献情報処理入門」授業のシラバス

第1回	電腦東洋学の地平(導論・理解度テスト)
第2回	Windows 2000の中国語設定
第3回	中国語入力(IME)
第4回	中国語メール
第5回	漢文データベースの利用(台湾中央研究院)
第6回	漢文データベースの利用(四庫全書)
第7回	文字コードとユニコード
第8回	マルチリンガル・ホームページの作成
第9回	Wordで訓点レイアウト
第10回	まとめと理解度テスト
第11回	今昔文字鏡の利用
第12回	マルチリンガル・プレゼンテーション
実技テスト	(テスト期間中に実施)

写真3 「スクリオ・リバーウォーク」のコンテンツ説明画面



図.1 パッド型携帯端末 - スクリオ -

として獲得させることを目的としたコンテンツを作成し、日本教育工学会第 18 回全国大会で発表した^[7]。今年度も同学会で、京都における「シティーワーク」をテーマとしたコンテンツを発表する。

■ 大谷大学所蔵『北京版チベット大蔵経』の高精細デジタル画像化

大谷大学に所蔵される貴重書『北京版チベット大蔵経』は、1720 年に北京で木版出版された 358 巻、20 万ページに及ぶチベット語の一切経である。この 1720 年の版本の所在は本学以外には世界中どこにも知られていない。すでに影印出版されている^[8]が、その印刷は部分的に不鮮明で読みにくい部分も多い。また、現物は朱墨刷りで退色が進行しているため、高精細デジタル画像化によって現状を保存することが急務である。同時に、ダウンサイジングした画像を公開して研究者間で共有する目的で、2001 年度からこのプロジェクトはスタートした。

デジタル画像化は、35 ミリ・リバーサルフィルムによる撮影→高精度フラットベッドスキャナ(5,300dpi 相当)によるスキャンという二段階を経てなされるが、それぞれの段階における精度追求の理論とノウハウ^[9]は、アルバイトとして参加する学生たちに伝授され、文化財の保存と公開を目的とした高度な画像処理技術という研究テーマになっている。

◇ 今後の課題

本稿冒頭で提起された「根源的問い」に対して、筆者は未だに回答を用意できていないが、ここでは 4 年間の試みを通して明らかになった 2 つの課題を述べておきたい。

ひとつは、教える側が現代社会における「知」のあり方の変遷にいかに対応してゆくかという問題である。第 2 次ゲーテンベルク革命といわれるデジタルデータ氾濫の時代にあって、教える側と教わる側の情報量の差は限りなく近づきつつある。例えば、漢文講読演習で誰もが苦勞した「用例探

し」や「出典探し」が、『四庫全書』や『四部叢刊』といった大規模データベースの出現によってパソコン上でキーをたたけば容易に実行できるようになった。こうした傾向がさらに進み、あらゆる人文情報がネットワークで結ばれ、いつでも取り出し可能な分散型「知」の総体として捉えられた時、教える側の責務は情報の「切り売り」から、情報をいかに「処理」するかへと移行せざるを得ない。人文系教員に FD が求められる所以である。

もうひとつは、専門に根ざした大学院レベルの人文学情報処理研究を支援することによって、学部学生の研究志向を喚起することである。人文系の学界において真に求められるものは、情報処理ツールを制作する技法ではなく、そのツールを使って生み出される「人文学的成果」そのもののはずであるが、現状は「本末転倒」といっても過言ではなく、その原因の一端は大学院における高度情報処理教育の未整備にある。

今後、文学部における情報処理教育は、人文学的成果という到達点を、教える側と教わる側がいかに共通の目標として認識できるかが鍵になるであろう。

注

- [1] 京都市北区小山上総町。<http://www.otani.ac.jp/>
- [2] <http://www.otani-humane.info/index.php>
- [3] 真宗学、仏教学、哲学、社会学、史学、文学、国際文学の各学科。
- [4] 以下は筆者のゼミ生の卒論テーマの一例である。

デジタルアーカイブを利用した二条城障壁画の商業利用について／博物館におけるデジタルコンテンツの先端的提示方法について／現代社会における活字メディアの可能性／バーチャルリアリティ技術の限界——VR 難壇は実現可能か——／伝統的日本文化のマルチメディアを利用した表現方法について——雅楽の笙をとりあげた電子媒体の作成——／マルチメディアを利用した古典文学のデジタル化——FLASH による「サウンドノベル『万葉集』」——／仮想博物館による無形文化財の提示方法について——大文字五山送り火を例に——／美意識と眉化粧の変遷——画像処理による

分析——／仮想博物館の構築における多層イメージの
追究／伝統文化復元の可能性についての研究——祇園
祭の鷹山を例に——

- [5] 2003 年 10 月 14 日、大谷大学構内に開館予定。
- [6] 米アップル社が開発した仮想的に三次元空間を再現で
きる技術。
- [7] 宮脇裕・和田圭二・櫛村暢彦・水田庄一・森山馨一朗・
平澤泰文・川田隆雄・松川節・目黒雅昭（共著）「パ

ッド型携帯端末を利用した分散協調教育システムの開
発」

- [8] 鈴木学術財団（編）『影印北京版西藏大蔵経』
1955-61, 東京・京都。
- [9] 片岡裕・柴田みゆき（共著）「一次資料として使用可
能な精度保証付き高再現性デジタル画像の作成と提
示」(http://www.dl.uilis.ac.jp/DLjournal/No_16/9-kataoka/9-kataoka.html)

花園大学「情報歴史学コース」の 現状と課題

佐藤 誠（さとう まこと）、田中 正流（たなか まさる）

◇ 0. はじめに

花園大学^[1]「情報歴史学コース」は、昨年度より文学部史学科の1コースとして開設された。本稿では、開設2年目に入った本コースの概況と、問題点、展望について、講義の担当者の立場から報告していきたい。本コース開講以前の課題および理想論的展望については前号において、来年度より本コースのゼミを受け持つ師の論文^[2]を参考していただきたい。

花園大学では2002年度の学科内編成により学生は、2回生時よりゼミの前進である講義を受講することとなった。昨年度の本コースの講義は「コンピュータ歴史学」のみであった。対象の学生は一回生のみで、講義要項には「ある程度ITスキルが要求される」としていた。実際の内容としては、前期はテキストデータの収集・活用についての基礎的な実習を、後期は考古学を主題にして、GISを中心とした遺跡情報のデジタル化についての実習が行われた。後期からは、授業履修者全員にノートPCを貸与し、それぞれがある程度自由に扱える実習室での講義が行えるようになった。本年度から2回生を対象とした、「情報歴史学研究」「研究入門演習」が開講されるようになり、予想以上に履修学生が増えたことによって、残念ながらPC貸与という形は取れなくなった。そこで、不自由ではあるが、実習室にノートPCを固定して、ソフトを自由にインストールし、「使い倒す」ことの出来る環境を維持した。

◇ 1. 講義の概況

本コースで現在行われている講義のうち、特に史学科に関連する授業は、前述した「コンピュータ歴史学」、「情報歴史学研究」、「研究入門演習」の3つである。昨年度は、「コンピュータ歴史学」のみの開講であった。本年度は、2回生を対象とした「情報歴史学研究」、「研究入門演習」の2科目が開講した。以下、それぞれの授業内容について述べていくこととする^[3]。

一、コンピュータ歴史学

「コンピュータ歴史学（前期）」は、これから「情報歴史学コース」へと進む学生を対象としている。情報歴史学と従来の歴史学との違いやデジタルデータとアナログデータについてメリット、デメ

表 1 開講科目表 史学科（情報歴史学コース）一部抜粋

区分	科目名	開講期間	履修年次
必修科目（普通講義）	コンピュータ歴史学Ⅰ	前期	1～
	コンピュータ歴史学Ⅱ	後期	1～
必修科目（特殊講義）	情報歴史学研究Ⅰ	前期	2～
	情報歴史学研究Ⅱ	後期	2～
必修科目（コース必修科目）	データベースシステムA	前期	1～
	データベースシステムB	後期	2～
	情報検索演習1	後期	2～
	情報検索演習2	前期集中	2～
	情報メディア	前期	1～
	情報メディア実習	後期	1～

必修科目	研究入門演習Ⅰ 7	前期	2～
	研究入門演習Ⅱ 7	後期	2～
	日本史学演習A	通年	3
	日本史学演習B	通年	4
	卒業論文	通年	4
選択科目	情報と社会	前期	1～
	コンピュータ基礎A－Ⅰ	前期	1～
	コンピュータ基礎A－Ⅱ	後期	1～
	コンピュータ基礎B－Ⅰ	前期	1～
	コンピュータ基礎B－Ⅱ	後期	1～
	アルゴリズムとプログラミングA	前期	1～
	アルゴリズムとプログラミングB	後期	1～
	デジタル制御	後期	1～
	情報システム管理技法A	前期	1～
	情報システム管理技法B	後期	2～
	情報通信技法A	前期	1～
	情報通信技法B	後期	2～
	ネットワーク論	後期	1～
	マルチメディアの活用A	前期	1～
	マルチメディアの活用B	後期	2～
	情報と職業	後期	1～

リットなど、情報歴史学の入門的な講義を行っていった。

まずディレクトリ型やロボット型など検索エンジンの違いを理解させ、ネット上の膨大な情報の中から研究に必要な正確な情報だけ抽出する情報検索のスキルを身につけさせた。

課題として、平安時代の年中行事を題材に、情報検索のスキルを駆使して、Webcat^[4]などの文献目録や東京大学史料編纂所^[5]などの大規模データベースを活用し、検索を行い、そのテーマについてレポートをまとめる作業を行った。その際、同時に紙史料を使い実証的にまとめた。

次に、テキスト処理に入ったが、まず文字の歴史、文字コードの変遷について講義した。ビットとバイトの話から始まり、アスキーコード、JISコードからUnicodeまで、加えて今昔文字鏡^[6]のことにもふれた。

この際の課題として次のようなことをさせた。まず日本霊異記をシフトJISにて入力させ、無い

漢字はゲタをはかせる。その上で、印字させ、無い漢字に関しては諸橋大漢和を引かせ、大漢和番号を書き込ませる。これにより、シフトJISには収録漢字が少ないこと、漢字が無い場合の対応策も含めて学べたと思われる。

この授業では、前期と後期で担当が変わるので、後期は考古学の資料を用いたコンピュータ処理を行う予定になっている。

二、情報歴史学研究

「情報歴史学研究」では、コンピュータ歴史学の単位取得者である2回生を対象とし、より高度で専門的な講義となった。

前期は、文字コード問題をふまえた、全文検索を中心としたテキスト処理の実習を行った。テキスト検索は、1回生時にコンピュータ歴史学を受講した学生には若干なじみがあるようであったが、1回生時に履修した学生が少なかったこともあり、全体的にはゼロからの出発となった。

具体的には、歴史史料を題材としたテキストデータを用いて、簡易なGREPソフトの使用から始まり、QGREP^[7]を用いた正規表現、続いて今昔文字鏡を通じた文字コード問題、他言語データ（特にBig5）をシフトJIS環境へ変換するためのソフトの利用などを行っていった。

この講義もコンピュータ歴史学と同様に前期後期で担当が変わる。しかし後期の内容としては、前期の内容を踏まえてデータベースの構築と活用及び、マークアップを中心としたテキスト処理について講義する予定をしている。

三、研究入門演習

「研究入門演習」では、2回生を対象としている。この演習は、他コースの場合、その分野の論文を精読し、使用されている論文・史料を集めどのように論旨の組み立てが行われたかを学ぶものである。

しかし、情報コースでは、最終的にコンピュータの特性を活かした歴史学の論文を書くことを目的としているため、上記の内容の他に、研究を行っていく上で必要となるであろうサイトの紹介、

ソフトの使用の実習というような技術面の授業を行っていく必要があった。

前期ではまず研究のための環境を整える技能として前述のような研究に活用可能なソフトの導入と実践を行っていった。タブブラウザソフトや今昔文字鏡、年号変換などのソフトを中心に各自サイトからダウンロード、インストールを行わせていった。サイトの紹介については他のコースでも使用するような、Webcat や国立国会図書館^[8]、また花園大学内の図書館の検索ページ^[9]を皮切りに、データベースなどの研究機関のサイトを紹介していった。

Webcat や学内の図書館のサイト (hops) の活用は、初歩的ではあるがコンピュータを利用して研究していく上で初歩の初歩ともなるため、特に時間を割き課題を与えた。内容は、国際日本文化研究センターの怪異・妖怪伝承データベース^[10]から任意の地域の事例を 10 件選び、その典拠となっている雑誌・論文などを hops 及び Webcat で検索し図書館にあるもの 5 件の記載ページと奥付をコピーする。学内に無いものは所在確認をさせる。検索した事例 10 件はリスト化して別に提出するというものであった。同時に、従来の論文の作成方法や史料論などの講義も行っていた。

後期は、論文精読と、論旨の要約を含めたレジュメの作成 (勿論、PC を活用したレジュメの作成が前提となる) やデータベースの活用と作成について行っていく予定である。

◆ 2. 課題

四月から授業を始めて感じたことは以下の二点である。コンピュータの技能習熟度の格差と歴史的思考(知識)の欠如である。これらの問題点は、本コースの全ての授業に対して言えることであり、現時点においては相乗効果を起こして問題を大きくしている。あるデータベースを活用した場合、試しに何か検索をかけさせると、歴史学の問題意識が薄いため何を入力したら良いか迷ってしまい、検索文字の入力画面で止まってしまう学生が多々見られた。五山文学のテキストデータを用

いた GREP においても同様の事例が発生した。この場合、学生になじみのある近代文学^[11]のデータを用いて検索をするという方法に切り替えるなど対応策が採られた。しかし、今度は検索ソフトの使用方法や正規表現という技術面で立ち止まるというように、技術的な格差も傷口を大きくしている。各自がコンピュータを所有しているのであればもう少し技術面の問題は埋められると思うが、学内で自由に利用可能なコンピュータはソフトのインストールも不可能なほど制約が多いため難しいのが現状である。

2003 年から高等学校の普通科では「情報 A」「情報 B」「情報 C」が新設され、そのうち最低一科目が必修となった。従って、技術面の問題は、3 年後にはコンピュータの基礎をそれなりのレベルで学んだ学生が入って来るような段階になれば解消されるのではないかという甘い期待をしているのであるが、甘い幻想で終わりそうな気もしないでもない。そうなるとこの格差は、やはり本コースに入る早い段階で均一化する (つまり足切りを行う) ことで解消していくしかないのであろうか。

本コースにおける習熟度のラインを何処へ持ってくるか、これも常に考え続けたいといけない問題である。現段階では、卒論案として、「電子化された史料を使った歴史研究」を目標としているが、それが無理な場合として「史料の電子化に関する諸問題」をあつかった場合も可とするゴール設定がなされている。

同時に社会におけるコンピュータのハード・ソフト両面での急速な進歩から考えると、進歩に対し柔軟に対応できるような人材の確保・育成、加えてアップトゥーデイトな PC 環境の整備も必要となってくるのではないかと考えられる。

もっと切実な問題点としては、セメスター制というカリキュラムの制約上、半期一区切りとなってしまう点である。これは特に基礎を学ぶための「コンピュータ歴史学」において深刻で、短期間であつたり、担当教員の入れ替わりがあることなどから、継続性や課題設定などの点が難しくなっている。

◆ 3. まとめ

情報歴史学という学問を進めていく上で、基礎的歴史用語・歴史学認識に対する欠如という問題は、1・2回生時の講義において今後も問題となりうるものといえる。第3号の師論文にも「従来の方法論について知識を深めておくことは肝要であろう（以下略）」とあるが、肝要どころか前提条件としてあるのではないとも言える。教員側は、専門としての分野を持った上でその研究の向上のためにコンピュータを使っている。そのため現在、本コースの学生が置かれた状況とは、コンピュータの利用に対する出発点が異なるのである。その点を考慮すると、本コースに在籍しつつ歴史学の他分野の講義を積極的に受けていくように指導していくことによって、仮にコンピュータが無くても研究できるような歴史研究者を養成するような方向へゴール設定する方法が考えられる。反面、史資料をコンピュータ上で扱うことを前提とした「歴史情報」の研究者を育成するというゴール設定も選択肢として重要である。

以上、講義内容と課題について述べてきたが、これら3つの授業で繰り返し基礎的な実習を行った結果、情報歴史学におけるテキスト処理や情報検索の理解とPCの基本的操作が身に付いたと思われる。教科書も無く、マニュアルも無いという

手探り状態で、まだまだ課題は山積みではあるが本報告では以上とさせてもらう。

注

- [1] 京都府京都市中京区西ノ京壺ノ内町 8-1、<http://www.hanazono.ac.jp/>
- [2] 師茂樹「情報歴史学の課題 花園大学・情報歴史学コースのための主観的航海図」(『漢字文献情報処理研究』第3号、2002年)
- [3] それ以外に、CDC (Career Development Center) の情報ブロックの科目が史学科の科目として組み込まれており、情報歴史学コースの学生はそれを受講している。詳しくは表1参照のこと。
- [4] NACSIS Webcat <http://webcat.nii.ac.jp/>
- [5] 東京大学史料編纂所 <http://www.hi.u-tokyo.ac.jp/index-j.html>
- [6] 今昔文字鏡 <http://www.mojikyo.org/>
- [7] QGREP <http://Oban.com/araken/>
- [8] 国立国会図書館 <http://www.ndl.go.jp/>
- [9] hops (花園大学 OPAC サービス) 学外からアクセスするには、hunet (花園大学学術情報ネットワーク) のアカウントが必要。
- [10] 国際日本文化研究センター「怪異・妖怪伝承データベース」 <http://www.nichibun.ac.jp/>
- [11] 講義中では、青空文庫 <http://www.aozora.gr.jp/> 所収のテキストデータを活用した。

オンライン討論

人文学情報処理教育における
教える側の問題

◇ はじめに

もろしげき

2003/06/16 23:16:01

昨年、『漢字文献情報処理研究』第3号における特集「人文学情報処理教育」において、「人文学情報処理教育はいかにあるべきか」というテーマでのオンライン討論の様子が掲載されました。この討論記事は読者の興味を大いに引き、反響も少なからずあったと聞いていますが、これはすなわち人文学情報処理教育の必要性、あるいはそれがなかなかできないことへの危機感が広く認識されつつある、ということだろうと思います。

先のオンライン討論のなかでは様々な問題が指摘され、またその解決のための提案もいくつかなされましたが、中でも教える側の問題を危惧する声が大きかったように思います。昨年の二階堂氏の言葉を借りれば「人文情報処理教育は、各分野の授業に埋没すべき」、つまり中国文学なら中国文学の、日本史学なら日本史学の、と言った具合に、人文系のそれぞれの分野で固有の問題を解決するためにこそコンピュータは使われなければならないのに、現実には教員側に能力がない、能力がある人材を確保しない（できない）などの問題から、情報処理教育が「各分野の授業」とリンクすることはあまり見られないように思います。また、幸いに人材が確保され、それがなされているところでも、従来の講座の体制が変わることなく

あり、それに情報処理教育がおまけのようについついていような場合が多いのではないのでしょうか。

そこでこの討論においては、議論を始めるにあたっての糸口として (1) 学科・研究室で情報面を担当する教員に求められるスキル、および (2) 学科・研究室全体で持っているべきコンセンサスやスキル、の二点について、皆様のご意見を頂いた上で議論を始めたいと思います。理想論を言えば、教員全員がそれなりのレベルのコンピュータ・スキルを身につけていることが望ましいのでしょうか（そのような立場からの発言も歓迎します）、昨年の討論における千田さんの発言にもあったような「最低限の状況の確立」をとりあえず達成するにはどうすればいいだろうか、というところから議論を始めたほうがよからうと思い、あえてこのような問いかけをしてみました。

では、お願いします。

◇ 人文学情報処理教員に求められるスキル

二階堂よしひろ

2003/06/17 19:59:25

「単に使える人」と「情報技術を教えられる人」の違いは何でしょう。

自分もあまり明確な答えがあるわけではないのですが、一つはもちろん、新しい技術や操作方法に対して、一定のスキルを持っていることが「教

えられる人」の条件であると思っています。

ここで「新しい」と限定したのは理由があります。というのは、人文系情報処理の授業では、かなり「固定化」の問題があるからです。

昨今では、人文情報の授業を課すところも、ある程度は増えてきました。しかし、その授業の中身については、時に MS-DOS 時代のファイル操作であったり、一太郎のかなり前のバージョンでの印刷テクニックであったりと、ずいぶんと古びたものであったりすることがあります。

一般に、人文系の授業というのはかなり固定化される傾向があります。それは一種の職人芸のようなもので、いったん確立されてしまうとなかなか変更されにくいものです。毎年新しい材料と意識でやっている、そういった授業はむしろ少数派に属するのではないのでしょうか。

問題は、これが人文系の情報授業にまで適用されてしまうことです。それはもちろんその教員が苦勞して確立した方法ではありましよう。ただ、情報技術がものすごい速さで進展する今日では、ある手法はあっという間に時代遅れになってしまうのです。

むろんいつも最新であることを求めるわけではありません。しかし、ある程度、常に新しい技術を吸収し、応用することは求められてよいと思います。ある意味で、これは大変なことだと思いますが、このように求められることは、人文系情報教員の「宿命」と言ってよいかもしれません。

さらに、もう一つ「単に使える人」との違いは、「仕組みについて理解している」ことであると思います。

これもどこまでと言われると、線引きが難しいのですが、つまり OS やアプリケーションの基本的な働き、文字コードの仕組みやその実装、さらに、インターネットなどの通信技術についての知識を有し、理解していること。また、自分が研究対象としている言語の処理について、それなりの技能を持っている。こういったことが必要になってくると思います。

実は、「仕組みを理解」している人であればこそ、新しい技術に対応することが容易なのです。原理

を理解している人であれば、あとは表面上のテクニックを身につければ、応用が効くのです。

ですから、「仕組みを理解する」ということと、「常に新しい技術に対応できる」ということは、表裏一体のものであるのかもしれませんが、仕組みが分かっていないからこそ、いったん確立した手法に拘泥して、次の段階に進めないとも言えます。

もっとも、現状では「単に使いこなせる人材」や「古い手法でそれなりに使える人材」ですら不足気味であると思いますが。

だいたい、自分もどこまで「実装」できているかは甚だ心もとないです。「ばりばりスクリプトが書ける」とか、「多言語処理をばんばんこなす」ということまでとなると、かなり苦しいですね。

山崎直樹

2003/06/23 14:05:26

二階堂さんは、「使える人」とちがって、「教える人」は、仕組みを理解していないといけない、と述べておられました。その理由として、「仕組みを理解していれば新しい技術に、容易に対応できる」を挙げておられました。

わたしも、「教える人は、仕組みを理解している必要がある」には、賛成です。そして、その理由として、もうひとつ、以下を付け加えたいと思います。

すなわち：

教える人は、ある目的を達するために、複数の手段があるとき、どうして、ある手段が望ましく、別のある手段が望ましくないのか、説明する責任がある

です。

情報技術の教え方は、descriptive（描写的）ではなく、ある程度は、prescriptive（規範的）であるべきだと思います。簡単に言うと、ああいう手もある、こういう道もある、という網羅的な知識を伝えるのではなく、こうすべきだ、ああすべきではない、という指示も含むべきだということです。これは当たり前ですね。

で、そのとき、説明をする必要があると思うの

です。イメージとしては……

風邪気味のとき、「あったかくして早く寝なさい」とよく言われますね。これを、おばあちゃんが言うのだったら、おばあちゃんは、なぜそうしないといけないかという仕組みを説明する責任はありません。

しかし、医学部の授業で、教師が、「暖かくして早く寝るべきだ」と主張するのでしたら、説明責任を伴うと思うのです。

おばあちゃんは、「使える人」であればいいのですが、医学部の教師は「教える人」ですから。

もろしげき

2003/07/07 01:12:25

「仕組み」を教えられる、ということがひとつのポイントになるという二階堂さんや山崎さんのご意見に、基本的には賛成です。

しかし、二階堂さんが最後のほうで仰った、「仕組みを理解する」ということと、「常に新しい技術に対応できる」ということは、表裏一体のものであるのかもしれない、というご意見には、少し違った印象を持っています。

私の乏しい経験では、そういう主張をする人は、情報系を除く理系出身の先生に多いようが気がします。それで、その先生が新しい技術に対応できているかという点と必ずしもそうではない場合が多い——人文系における「固定化」と同じようなことがよくおきる——ような印象です。

情報系出身の先生は、「仕組み」についても当然知識を身につけておられますが、ヒューマン・インターフェースな部分への理解、あるいは概念レベルの話と実装レベルの話を分けたりする現実性などがあって（という理屈づけは私の想像ですが）か、日進月歩で変わっていく現実への対応能力があるような印象です。どちらも、あくまで印象論なので、その程度の話として聞いていただきたいのですが。

また、山崎さんが仰られた「説明責任」についてのご意見はまったくもっともだと思いますが、人文系においてコンピュータなどの「仕組み」を教えると言うことの重要性は、それ以外のところにもあるのではないだろうかと思っています。

つまり、普段考えない「考え方」を考えるきっかけになる、といひましようか、例えば文字コードについて学ぶことを通じて文字について再考したり、パターンマッチの仕組みを学ぶことを通じて、テキストを読むときに我々が無意識のうちに分節しているあり方について反省したり、といった具合に、使い方の裏にある「仕組み」を説明することをネタにして、人文系に必要な批判的なあり方を身に着ける訓練ができると思うのです。

つまり、教員が持つべき「仕組み」についての知識は、二階堂さんや山崎さんが述べられた、技術を教える前提、背景というような観点において重要であることはその通りだと思うのですが、それだけでなくもっと大きな人文系的な広がりをもたらすためにも重要なのではないか、と思うのですが、いかがでしょうか。

山崎直樹

2003/07/07 18:06:44

もろさんがおっしゃっていた「普段考えない『考え方』を考えるきっかけになる」ということについてですが…

ぱっと、わたしの頭に浮かぶのは、

何のために数学を学ぶか？——思考能力の訓練のため

という類いの考え方ですが、もろさんのおっしゃっているのは、こういうことではないですよね。

情報処理技術といってもいろいろありますが、確かに、ある種の技術を教えるときには、以下の問題に触れざるを得ないと思います。

- 我々は、ふだん、何気なく、何となく似た字を同一視しているが、これは、離散的なものしか区別できない機械にとって、たいへん面倒な作業である。
- シソーラスについても同様。
- 文字コードの問題を考えると、社会言語学でいうところの「言語計画」「言語政策」といった問題を考えざるを得ない。

特に、日本では、言語政策が意識されることが少ない（大学生の、この問題に対する意識は、驚くほど低い）ので、文字コードを切り口にした東アジアの言語政策の講義などは、面白いかな、と思ったことがあります。

逆に、言語政策の研究者は、情報処理の問題に疎いことが多いですね。

確かに、ある種の情報処理技術は、われわれに、「人間の認知能力」とか、「社会と言語のかかわり」とか、の問題を考え直させてくれますね。でも、後者は、とくに、文系だけに必要なことではないかもしれません。「社会とのかかわり」を考えずに、知的好奇心だけで、新しいテクノロジー（例えば、核兵器）を振り回されても困りますからね（むちゃくちゃな飛躍だ）。

二階堂よしひろ 2003/07/08 18:31:30

もろさんの言う教員が持つ知識は「もっと大きな人文系的な広がりをもたらすためにも重要なのではないか」というご指摘については全くその通りだと思います。

但し前提としては、われわれはそれを視野に入れつつ、まずは目先の「技術」を教える必要があります。「人文系の授業は職人芸」といったことを前に書いたような気もしますが、紙のテキストや辞書索引を使いこなす職人芸を教えるように、電子テキストについてもやはり職人芸的なものが必要となるわけでしょうから。むろん、そこでは一定のスキルが要求されます。

とはいえ、だいたいの授業がこの時点で終わってしまうような予想もあります。これは大学のどの段階で教授されているかにもよりますが、そもそも多くの学部生のレベルでは、なかなか応用面にまで辿り着きません。

それでも、その授業の過程においていろいろと新しい可能性を含む「ネタ」を振りまいておくのも、人文系情報教員の楽しみかもしれません。

二階堂よしひろ 2003/07/08 19:01:50

人文情報教員に必要な「スキル」とは具体的に何を指すのでしょうか。まず、求められる最低限

のレベルは、次のようなものでしょうか。

- ワープロ・表計算ソフトなどの使いこなし
- ネットワークに関する基礎知識
- ハードウェアへの一般的理解

これは一般の情報処理授業でも変わらないと思います。ただ、意外にハードウェアに関しては追究されないですね。これはよくない傾向だと思います。

さて、次に求められるのは人文学的な使いこなしに関してでしょうか。

- Unicodeなどの文字コードと多言語処理に関する知識
- 特定分野の主たるデータベースの用法
- 論文や関連資料の検索技術

自分などはこの段階までしか教えられないですね。さらに欲を言えばこうなりますか。

- 電子テキストやコーパスを使用した統計的な解析
- PerlなどのスクリプトやCなどのプログラムによるテキスト処理

しかし、このようなスキルをそれぞれの分野で会得するというのは、現段階ではやはり大変なことではないでしょうか。そもそも教える人がいないような気がしますし…。

◇ 学科・研究室のコンセンサスと人材育成

千田 大介

2003/09/09 13:07:22

人文学情報処理授業を設置する大学は、おそろしく教員免許の絡みもあって、ここ一・二年でかなり増加したようですが、しかし聞こえてくるのは担当教員の慨嘆ばかりで、それらの授業が有効に機能しているという話はなかなか伝わってきませ

ん。この点、師さんの前振りの通りだと思いますし、「学科・研究室全体で持っているべきコンセンサスやスキル」について議論する必要は確かにあるでしょう。

前回の討論で、情報処理の人文学授業への埋没が話題になりました。その点について私は、人文学情報処理は文献学（版本学・書誌学）と同様に扱われるべきものだと考えています。文献学は、東洋学の基礎となる領域ですが、その理論を完全に理解している人文学研究者は、そう多くないのではないのでしょうか。文献学の知識は文献の読解や理論の授業にも必要になりますが、基本的なところを押さえておけばことさらに取り上げなくても済むケースが大半です。しかし、図書館目録の作成といった作業になると、概略的知識では歯が立ちませんので、文献学プロパーの関与が必要になります。

人文学情報処理でも、データベース作成などの情報発信、あるいは N-gram などの高度なテキスト分析の場合には、「仕組みを理解」した人文学情報処理プロパーが関与したり教授したりする必要があります。一方、多くの人文学授業では、教員は必要なデータの引き出し方を修得していて、その「手順の説明」ができれば、必ずしも「仕組みを理解」していなくてもいいと思います。

例えば、唐詩の授業であれば、ごく自然に北京大学『全唐詩』データベースの使い方と、その読解・研究への応用方法について、あわせて内容の信頼性の評価や検索 Tips などについて教えるべきですが、文字コード問題やデータベースの構造などについて踏み込む必要はないでしょう。

現状では、人文学情報処理に対応できる、つまり「仕組みを理解」している人材が非常に不足していますし、パソコンを使いこなせる教員すらもそう多くありません。理想的には、あらゆる教員が「仕組みを理解」するレベルに達して、情報処理を人文学授業に埋没させられればいいのですが、現実的には、一人二人の「仕組みを理解」した教員を確保することすらままならないでしょう。

その前提に立つと、人文系「学科・研究室全体」の情報化の形態としては、

- 使いこなしレベルの多数の教員による、授業の中で「手順の説明」
- 「仕組みを理解」した少数の教員による、人文学情報処理講座

というのが現実的であると考えられます。

ところが、このような形態すらも実現が困難な現実があります。特に問題なのは、使いこなしレベルにすら到達することができない、あるいは到達しようとししない教員の存在です。そのような人たちは、とすると、少数のできる人に情報関係の全てを押しつけ、自らは電腦素人を自認して旧来の枠に閉じこもろうとします。甚だしきに至っては、自分ができないことが他人にできるのが我慢ならないとばかりに、情報化否定論に凝り固まってしまう。

過度の期待とその裏返しとしてのパソコンやインターネットへの憎悪、それを排除して、道具としてパソコンとインターネットをとらえる意識を定着させる、このスタート地点にも立てていないのが、大半の大学の現状だと思います。

漢情研が長らく活動を積み重ねながらも、結局実現できていないのが、この意識改革です。結局、コンセンサスの確立こそが、適正な人文学情報処理教育を普及・定着させる上での、最難関なのでしょうね。

もろしげき

2003/09/09 22:54:12

情報化否定論者には、昔ずいぶんいじめられました（苦笑）。昨今の IT ブームの中で一時より大分減ったと思いますが（むしろ妙な IT 礼賛論者の増加が目立つような気がします）、まだまだいるところにはいるんでしょうか。インド学・仏教学界は情報化に理解があって、めぐまれているような気がします。

それはともかく、千田さんの仰るような研究室の体制は、非常に現実的だと思いますし、個人的にも、従来のゼミに対して人文情報処理のサブゼミが存在する、みたいな体制は一つの解だと考えています（本誌前号「情報歴史学の課題」参照）。

しかしこのような体制だと、「人文学情報処理

プロパー」はどこで育成されるのだろうか？という疑問が残ります。現在の人材不足は、裏を返せばそのような人材の育成がなされてこなかったことに原因があるのは言うまでもありません。今後の人材育成を考えると、従来の授業に埋没しない人文学情報処理をメインにした学科、ゼミが文学部の中にできていいような気がします。

この点について、千田さん、みなさんのご意見はいかがでしょうか

二階堂よしひろ 2003/09/10 13:56:26

いまやコンピュータやインターネットそのものについては、それほど特別視されることは無くなりました。会社の風景などを見ても、もっぱらパソコンに向かって仕事をしている感じで、ワープロやメールは、完全に一般の生活にも取り込まれています。

では、教育の現場で「埋没」していくほど浸透しているかという、これがまたお寒い状況です。本当は、教育や研究の場においてこそ、一層の発展が望まれているにもかかわらず、です。

悩ましいのは、これがまた情報処理の「一般化」が進んだことに起因するものなのかもしれないということです。何故なら、情報化が進んだことで、なるほど反対したり妨害したりする「負の圧力」は減じました。しかし、情報化を推し進めなければならぬという「正の圧力」もまた減少してしまったのではないと思われるからです。

いまたぶん、人文学における情報化について、現場の反応は「他人任せ」か「無関心」といったところでしょう。

これはこれで、困った現象です。というのは、情報化が一般化した今だからこそ、学問研究や教育の現場においては、「通り一遍ではない」真の情報化のステップに向けて進んでいくべきなのに、「これでいいや」というレベルに留まってしまう危険性があるからです。

むしろいまこそが、学部や研究室での対応、また学会における対応が求められていると思うのです。

そのカギとなるのは、やはりなんといっても人

材です。それは一般の人文学教育の場でも、人文学情報処理に特化した組織で行われても構いません。しかし、いずれにせよ「単につかひこなす」だけでは不十分で、今後の人文学における情報処理の方向を見極めながら進む必要があるでしょう。

学部や研究室、また学会などの組織には、もう少し長期的なスパンを考えて情報化の戦略を持つてほしいものです。特に人材育成ということを考えて。

千田大介 2003/09/12 12:25:23

人文学情報処理対応教員の育成は確かに難しい問題ですが、その際、最も重視されなくてはならないのは、大学院卒業生の進路の確保でしょう。

人文学情報処理学科は、その趣旨からして、一般的な情報処理人材に求められるプログラミングやサーバ管理などのスキルを即戦力レベルにまで高める教育は行いできませんので、そのような社会的ニーズとは合致しません。従って一般の人文系大学院と同様に、教育や出版・マスコミ業界が主な進路になるでしょう。

しかし、それらの業種では、あくまで人文学的な知識・能力が第一に評価されるので、人文学情報処理スキルは、いかに重視されても、プラスα的な扱いにならざるを得ません。

となれば、人文学情報処理は、やはり師さんのおっしゃるように、人文学大学院を構成する一つのゼミ、もしくはサブゼミとして提供されるしかないでしょう。もちろん、人文学情報処理に特化した大学院でもいいのですが、人文学の専門を確立した上で電腦もできる人材が育つように、十分配慮する必要があるでしょう。

もっとも、本誌に紹介されている人文学情報処理系コースはきわめて先進的な事例であって、大多数の大学では未だに学科はおろか、人文学情報処理のゼミを設置する機運すらも見られないのが実情です。アンチ電腦派は以前ほど目立たなくなってきたとはいえ、二階堂さんの書かれたような「他人任せ」か「無関心」という雰囲気は私も折に触れて感じています。

残念なことに、中国学全般の情報化は、未だに

啓蒙の段階にとどまっていると言わざるを得ません。となればむしろ、中国学を専門とする大学教員・院生を対象とする講習会を開いて、意識改革と基本スキルの普及につとめることを考えなくてはならないのかもしれませんが。

以前、『電腦中国学』を出版したとき、このマニュアルを読んで電腦を使う人が増えて、教員の意識改革が進み、人文学情報処理のできる若手も育ってくるようになる、というバラ色の21世紀を夢見ていましたが、絵に描いた餅でした。人文学で特に顕著な「固定化」の問題は、教員はおろか若手院生の中にもしっかりと根を張っているのでしょうか。

「革命未だならず」というのが正直な感想です。

二階堂よしひろ 2003/09/13 02:05:04

確かに『電腦中国学』を出した時には、漢籍データやツールの動向、それに人文学情報学の進む方向性を示しておけば、あとは自然に人材が育ってくるものだという期待感がありましたね。

しかし、当時と比べてデータベースや電子テキストは飛躍的に増加したにも関わらず、応用面での進歩はそれほどでもない感じがします。

そこで自分などは「待っているだけではダメ、人材は育成せねば」という考え方に宗旨替えすることにしました。そのためのノウハウの確立については、それなりに試行錯誤を積み重ねてきたつもりです。

とはいえ、これがまた可笑しいのですが、例えば自分の現在いる部署では、ほとんど人文学情報学など必要ではありません。せいぜい語学教育への応用ですが、それですら、設備や環境などの面から、限定的なものに過ぎません。

一方で、他の部署では、「コンピュータを応用しろ」と厳命されている教員も存在するようで、なんだかミスマッチを感じたりもします。必要なのに必要な人材が配されていないということは、実はこの方面では致命的な問題なのかもしれません。

それは何といっても、組織の上の側が、人文学

情報についての考え方が甘い、ということがあると思います。「現有の教員に、ちょっと付け焼き刃的にやらせれば何とかなる」くらいに考えているのでしょうか。残念ながら、人文学情報については、その程度の認識ではどうにもなりません。

おそらく、このような状況が続くと、それなりに形成されつつある人文学情報学の芽も潰れていくものと思われます。人材の重要性については、これまで何度も繰り返し主張していますが、どうも人文学情報については、組織の側に徹底して「人材を評価する」能力が欠けていることが大きな要因であるようです。

◇ おわりに

討論を通じて、現在の危機的な状況が再確認されるとともに、前号の討論で提示された「授業に埋没すべき」という方向性がより明確になった。すなわち、今後の人文学情報処理教育においては、

(1) 人文学に対して情報処理が有機的に連携するような「仕組みを理解した」人材の育成、および(2) 学科・研究室全体での「使える」能力の底上げが必須だということである。「組織の側に徹底して「人材を評価する」能力が欠けている」という現状認識は、まさにこの二つが達成されていないことを意味している。

討論中に見られたこれらの主張は、奇しくも本特集に投稿された二つの事例報告の主張と、軌を一にするものである。松川氏の言葉を借りれば、上記(1)は「真に求められるものは…ツールを制作する技法ではなく…「人文学的成果」そのもの」という主張に、(2)は「人文系教員にFDが求められる」という指摘に対応しよう。

危機感を共有する同志の存在を心強く思う反面、学界を覆う「無関心」の濃霧が晴れる日は来るのだろうかと考え、正直気持ちが悪くなる。人文学情報処理が色物や表層的なものに終わらないためにも、より積極的な働きかけが必要なのだろう。いずれにせよ、「待っているだけではダメ」なのだ。

東洋学情報化と 著作権問題

報告

インターネットは一個の巨大なコピー機である。電子メールが届くのも、ホームページが見られるのも、遙か遠くのサーバからその都度ファイルの複製が送られてくるからにはほかならない。したがって本来複製する権利（コピーライト）であるところの著作権が、インターネット普及当初から絶えの間ない議論を惹起してきたのも、むしろ当然の成り行きといってよいだろう。そして近年、東洋学におけるコンピュータ利用の一般化にともない、日々膨大なデジタル情報をやりとりする我々研究者も、もはや著作権問題に全く無関心ではいられなくなった。今回漢情研の公開講座では、こうした状況をふまえて、法学者の石岡克俊氏を講師に招き、著作権に関する講演及び参加者との質疑応答をお願いした。会場は法学研究者と東洋学研究者双方にとって大変刺激的な「異文化交流」の場となったが、その具体的な様子は以下に掲載される諸編から十分にうかがい知ることが出来るだろう。

※「漢情研 2003 年夏期公開講座 東洋学情報化と著作権問題」は、以下の要領にて開催された。

- ◆日時：2003 年 8 月 2 日（土） 13 時 30 分より
- ◆会場：慶應義塾大学 日吉キャンパス 来往舎中会議室
- ◆講師：石岡克俊（慶應義塾大学産業研究所助教授）

Contents

「東洋学情報化と著作権問題」参加レポート……………	師 茂樹 ……40
東洋学情報化と法律問題——第 1 回	
所有権の行使と無体財産の法的保護 判例の分析と解説 ……	石岡 克俊 ……43
著作権についての知識を深めよう	
東洋学のための著作権サイト・ページ指南……………	小島 浩之 ……53

「東洋学情報化と著作権問題」 参加レポート

師 茂樹（もろ しげき）

◇ 著作権講座、盛会に終わる

去る 8 月 2 日に慶応大学日吉キャンパスで開催された漢字文献情報処理研究会・2003 年夏期公開講座「東洋学情報化と著作権問題」は、30 人以上の参加者を集め、予定された 4 時間があったという間に過ぎてしまうほどの盛会で終わった。

前半は一時間半の予定で、講師の石岡克俊氏（慶應義塾大学産業研究所助教授）に著作権法の概論をしていただき、後半は参加者があらかじめ用意した具体的な事例に対して石岡氏が法律的な解説・解釈を提示するとともに、より詳細な議論を全員で行う、というものであった。実際には、前半の予定時間があっさりと超過されてしまい、後半の議論もはしょり気味という司会者泣かせの進行であったが、それが全体の充実感を損なうことはなかったように思う。

それは、単なる教える、教えられるという関係ではなく、石岡氏が我々の業界の問題を法学的に興味深い研究対象として捉えておられたこともあって、講演は非常におもしろく、それを踏まえた議論がさらに白熱（その勢いで懇親会も議論が尽きず）したためである。加えて、参加者に図書館やデジタルアーカイブ構築の実務者が多く問題意識が高かったこと、また石岡氏自身が漢籍や日本の古典を楽しんでおられた点なども理由としてあげられるだろう。有益な研究会になったことを講師の石岡氏、主催者、参加者各位に感謝申し上げたい。

以下、講演会の内容をメモ的にレポートしたい。まさに生齧りの知識にもとづくものであるため、誤謬、誤解についてはご寛恕の上、ご指摘いただければ幸いである。

◇ 法律用語の基礎知識

講演会の前半は「著作権法のパースペクティブ」という題で、法律の条文の読み方を（それぞれ 1 条 1 項 1 号などという数え方から）講義していただいた。強く印象に残ったものとして、次のようなものがあげられる：

- ・大陸法と英米法
- ・自然人
- ・有体物と無体物
- ・「利用」と「使用」
- ・「みなす」と「推定する」
- ・任意規定と強行規定
- ・私的自治の原則

例えば、ついついさらりと読み流してしまう著作権法の第一条にも多くの背景がある：

この法律は、著作物並びに実演、レコード、放送及び有線放送に関し著作者の権利及びこれに隣接する権利を定め、これらの文化的所産の公正な利用に留意しつつ、著作人等の権利の保護を図り、もつて文化の発展に寄与することを目的とする。

これを正確に読むためには、次のような法学的「常識」が必要とされるのである：

- 「著作＜物＞」という、現実に触れることができる＜物＞を想像してしまうが、それは法律的には「有体物」と言って民法などの範囲に属し、著作権法が保護しようとする「無体物」とは区別される。
- 「隣接する権利」は、法律の二大潮流である大陸法（独仏など）と英米法では前者にのみ認められるものである。ソフトウェアなどの著作権に関する議論では、コンピュータ大国アメリカの法律家の意見がしばしば参照されるが、日本は大陸法の系統に属するため、注意が必要である。
- 著作権法はもとよりアーティスト保護的な側面があったため「文化」という用語が見えるが、ソフトウェアなど、狭い意味ではアートには属さないものも著作権法の対象となっている。したがって現在、この「文化」は「国民の活動の総体」というような意味で用いられ、アートとしての質などは考慮されない（したがって賢えが悪いが、立ち話、便所の落書きも著作物になる）。
- 「利用」とは、著作財産権（22条以下）で規定されている許諾が必要なコピーなどの行為を指し、それ以外の「使用」とは区別される。

このような背景知識を持たないまま議論してきたことを恥じ入る一方、我々が問題としている古典籍やデジタル・データについては、日本の著作権法だけでは語れないことを痛切に知ることになった。

例えば、ある図書館が所蔵している写本について考えてみよう。著作権法が保護の対象としているのは、写本に書かれた内容（無体物）であって、物理的な実体をもつ写本そのもの（有体物）ではない。したがって、写本そのものをどう扱うかについては所有者である図書館に権利（民法で規定）があり、またその閲覧などについては図書館の所

有する土地に進入するわけであるからその制限に従うしかない、という具合である^[1]。

同様に著作権だけでは語れない例として、例えば漢和辞典などの文字番号やページ番号を使ったデータベースのことがあげられる。このような番号に著作権を認めることは基本的には困難であり（アメリカではインデックスに権利を認める判例があるらしい）、それを使ったデータベースを作ることも著作権法的には問題がないと考えられる。しかし、もし漢和辞典の出版社がデータベースによって売れ行きが損なわれると（主観的に）判断した場合には、不正競争防止法などによってデータベースの公開を差し止めることも可能であろう、という。

普段、論文などを書いたりしていることから著作権に対する親しみはあったものの、まさか東洋学に関連して民法や不正競争防止法、独占禁止法などが話題に上るとは思っても見なかったので、正直驚きを禁じ得なかった。

◆ 校訂は「創作」か、「真理の発見」か

次に、第2条第1項第1号を見てみよう。著作物とは何か、という規定である。

著作物 思想又は感情を創作的に表現したものであつて、文芸、学術、美術又は音楽の範囲に属するものをいう。

ここに見える日常的な用語も、法学的なフィルターを通してみると次のようになる：

- 「思想又は感情」は、自然人である人間のみが持つことができるものである。企業などの法人には著作物を作ることができない。ちなみに人工知能については、英米法では著作者であるとのこと。
- 「創作的」とはオリジナリティがあるという意味であって、クリエイティビティがあるという意味ではない。
- 前述の通り、「文化」＝国民の活動の総体

なので、ここでいう文芸云々の規定は現在ではあまり意味を持たない。

ここで興味深いのは「創作的」という言葉の内容であろう。オリジナリティというのは、大雑把に言えば他の著作物のまねをしていないということである。ただ、ここで「まねをしていない」というのは実際にまねをしたかどうかでは判断されず、比較対象が「周知」であるかどうかによる。例えば、A氏が書いた論文が、たまたま図書館に並んでいたB氏の本の一部と一緒だった場合、A氏がその本を見たことも聞いたこともなくても、その論文は「創作」とは見なされない。図書館に並んでいれば、その本は「周知」されているからである。

ところでこの「周知」の範囲であるが、学術書にありがちな、きわめて少数のもの、かつ著者がほとんど買い上げたりしている場合でも、それは「周知」であるという。つまり、不特定多数を意味しないのである。この点は注意が必要であろう。

また、数学における証明や公式の発見などは「真理の発見」であって「創作」に当たらない、という点も重要である。我々東洋学研究者が仕事として行っているテキストの校訂作業は、原テキストを復原するという意味では「真理の発見」に近く創作性が認められないかもしれない、という解釈が石岡氏によって提示された^[2]。ちなみにドイツでは「校訂権」という権利が、非常にハイレベルな場合に限るが認められているという。

これは大変興味深い議論である。文献研究の方法論を語る際、客観主義・実証主義であれとナイーブに主張する人もいれば、「作者はいない」という立場をとる人も少数ではあるがいる。上記の考え方を乱暴にあてはめれば、前者は「真理の発見」主義者、後者は「創作」主義者ということになるのか。そこまで極端ではなくても、校訂作業に「質」や「価値」、「個性」を見出す我々の感覚との齟齬については議論を深めなければならないだろう。

◇ さらなる議論に向けて

石岡氏も、校訂作業をはじめとする東洋学の研究がどのような性格を持ち、それをどのように法的に解釈できるのかについて、今後、共同研究をしていきたいと提案された。その際、前述の通り著作権法だけにとどまることなく、国内の関連法、中国・台湾・ドイツ・アメリカなどの法律など、多角的な視点が求められることは言うまでもない。

また、強調しておきたいのは、民法や著作権法には、公序良俗などに反しない限りコミュニティの合意や契約を優先する「私的自治」の原則が存在し、法律の内容も多くがそれを優先する「任意規定」であって、絶対に法律通りでなければならぬ「強行規定」（刑法など）とは異なる、という点である。つまりこれは裏を返せば、研究・教育情報をめぐる東洋学コミュニティの合意形成がこれから強く求められるということであろう。

今後、研究・教育活動はますますの国際化を求められ、またデジタル技術の重要性は決定的となってくる。国会のような場で議論・制定される法律が、現代の高度情報化社会に追いついていないことは人口に膾炙されているところであるが、我々研究者も、天から降ってくる法律を待つのではなく、国際的な議論、合意形成への積極的な参加が求められるのである。議論に参加することなく、旧来のドメスティックな慣習や方法に拘泥するのでは、せっかくの研究業績も、予算をかけた便利なデジタルアーカイブも、死蔵される可能性があるのではなかろうか。

注

[1] 本誌所収の石岡克俊氏の論文を参照。

[2] これに関連しそうな判例として、訓読文に著作性を認める「将門記訓読事件」がある（岡島昭浩氏のご教示による。『著作権判例百選第三版』、有斐閣、138～9頁など）。

東洋学情報化と法律問題——第 1 回

所有権の行使と無体財産の 法的保護^[1]

判例の分析と解説

石岡 克俊（いしおか かつとし）

【目次】

- 序——対話を終えて
- 一 はじめに
- 二 著作権と所有権の峻別
- 三 判例に見る法的保護の展開
 - (一)「広告用ガス気球」事件
 - (二)「長尾鶏」事件
 - (三)「クルーザー」事件
 - (四)「かえでの木」事件
- 四 おわりに

◇ 序——対話を終えて

筆者は、本年 8 月 2 日に催された漢字文献情報処理研究会主催の 2003 年夏期公開講座「東洋学情報化と著作権問題」に講師として招かれた。この講座では、まず、著作権法の基本的な考え方や枠組みについて筆者が説明及び解説を試み、講座参加者に共通の前提が出来上がったところで、次に、東洋学の調査・研究活動の中で諸氏が抱える現下の著作権問題^[2]についての報告が行われ、

続いてこれに基づく、議論・意見交換が行われた。筆者にとって、ここでのやりとりはとても興味深く有意義な経験であったことはいうまでもない。いや、そればかりではない。法学者として深い関心と問題意識を駆り立てられる大きな契機となったことを正直に告白し、今回の講座参加諸氏の熱心な議論への寄与に感謝する次第である。

今回のように、学術的な関心事項につき法律上の問題を検討する場合、法学者は、通常、自らにとって最も身近な学問領域である法学を念頭に置いて考えていく。他の学問領域における調査・研究の仕方やその作法を詳しく耳にする機会がないからである。しかし、このような態度のまま問題を掘り下げていくと、法学の領域において妥当する考え方であっても、他の領域においては全く見当はずれな、その学問領域に身を置く人々に全く省みられない結論を導くことになりかねない。法やルールというものは実効性がなければそもそも意味をもたず、実効性というものは、その社会やコミュニティを構成する人々の合意の上に成り立っていないければ、覚束ないものになってしまう。

講座に先立ち、あらかじめ提出されていた質問

の内容を検討する過程で、その中にある混同や誤解を発見し、当日、その一応の回答を試みた。これによって、疑問や問題の中には解決を見たものもあるかもしれない。しかし、筆者は、ここでの議論を通じ、東洋学の調査・研究活動の実態に即した、より詳細な議論と検討とが必要とされていること、そしてそれぞれの論点について筆者なりの一応の結論に到達しておく必要を痛感した。

そこで、この場を借りて、今回の対話によって見出されたいくつかのポイントを何回かに分け、議論の整理をしていくことにしたい。法律の規定だけではなく、関連する判例や学説、場合によっては諸外国の動向などにも配慮しながら執筆をしていくことになるだろう。当面、今回採り上げた論点を含め、(1)著作権と所有権をめぐって、(2)所蔵権をめぐって、(3)校訂権をめぐって、(4)タイプフェイスないしフォントの法的保護、等々が検討されるべきテーマとして考えられる。研究会は終わってしまったが、常時、法学者との意見交流の場として、この紙面を利用させてもらえれば幸いである。

なお、本稿における注等その他の記述は法律学の慣例に従うこととし、取り上げた判例等における当事者の取扱いなども同様とすることをあらかじめお断りしておく^[3]。

◇ 一 はじめに

著作権法を議論する際に、まず出会う躰きは有体物と無体物の区分であるといわれる^[4]。民法では、所有権の客体として「物」を定義し、ここでいう「物」とは専ら有体物を指すとしている(民法85条)^[5]。有体物とは、無体物に対する概念で、空間の一部を占め、有形的に存在するものをいい、液体、気体、固体がこれにあたる。他方、無体物とは、権利、自然力(電気・熱・光)など姿のないものを指す。また、人間の知的・精神的活動の成果たる著作物や、発明・考案などの技術的思想も無体物である。

民法がわざわざ「物」の定義をした趣旨は、物に対する全面的な支配権能である所有権の客体を

限定するためである。所有権の有無が問題となった場合、その客体が適切に確定されていなければならないからである^[6]。なお、所有権とは、その所有物を使用、収益及び処分する権利であり(民法206条)^[7]、ある特定の物を全面的に支配する権利である。

しかし、書籍・雑誌、レコード・CD、DVDやゲームソフトなど、著作物をさまざまな媒体に固定し内蔵している物、さらには、絵画、写真や彫刻など著作物と媒体とが一体化している物については、所有権を有する者であっても、その物を全面的にかつ排他的に支配することができないことがある。それは、既に述べたとおり、所有権の客体が有体物であるのに対し、著作権のそれは無形の思想・感情の表現といった観念的な無体物だからであり、有体物の支配は可能であっても、その有体物に固定され内蔵されている無体物をも支配することは不可能だからである。このような両権利における客体の相違は、これらの併存を可能とし、その限りにおいて、物の所有者は、自らの物でありながら自身で自由に支配できない場面に遭遇することになる。後述の「顔真卿自書建中告身帖」事件上告審判決においても指摘されているように、所有権は著作物を直接、排他的に支配する権能を持ち合わせてはいないのである。

他方、著作権法の中には、絵画、写真及び彫刻などの美術の著作物につき、著作権と所有権との間の調整を企図した規定が存在することに注意を喚起しておきたい。

著作権法は、美術の著作物及び写真の著作物について展示権を定め、これらの著作物につき、その主要な利用態様である原作品の展示に関し、著作者が排他的権利を有する旨を定めている(著作権法25条)^[8]。しかし、法は、屋外の場合に恒常的に設置する場合を除き、美術の著作物をその原作品の所有者(及びその同意を得た者)がその所有権に基づき公に展示することを認めている(著作権法45条)^[9]。この規定は、美術又は写真の原作品が譲渡された場合における著作権者の展示権と原作品所有者の所有権との間の調整を図るために設けられたものである。美術や写真の原

作品は、通常、有体物としてその美的価値を享受することになる。従来から原作品を購入した者がその所有権に基づき展示が可能であるということは認められているところであり、また、このような場合に著作権者の許諾をいちいち要するとなると、原作品の商品としての流通を阻害する結果を招来しかねないとの配慮に基づくものである。これは、所有者がその著作権の行使として展示することができるのではなく、著作権が制限された結果、展示可能となったものであり、その所有権の行使としての展示行為が著作権に抵触しないということを意味するに過ぎないとされている^[10]。

これまで見てきたように、有体物をその客体とする所有権と無体物をその客体とする著作権との間には、絵画や写真、彫刻など具体的な著作物の態様において、法は一方でそれらの併存を認めつつ、他方でその調整を必要としてきた。近年、著作権のように特定の法律によって必ずしも権利が明定されているわけではないが、相応の経済的価値を有しうる無体物（無体財産）^[11]に関する訴訟が増えており、それらに対して一定の法的保護の必要を認める判決が現れてきている。

これらの判例は、所有権の行使と無体財産の関係を整理・検討する上で重要な示唆に富む。

本稿では、著作物のように具体的な法律によって権利として定められ保護されている無体物のみならず、何らかのかたちで法的保護の必要性が裁判所によって認識されるに至った無体物を俎上にのせ、これらが所有権と如何なる関係に立つものなのか、また、これらの法的保護が判例において、如何なる理論的根拠の下、認められてきたのかを、判例の展開とともに検討していく。

◇ 二 著作権と所有権の峻別 ——「顔真卿自書建中告身帖」事件^[12]

【事案の概要】

上告人X（原告・控訴人）は、書の原作品を多数所蔵している博物館であり、本件の「顔真卿自書建中告身帖」も、その所蔵品の一つとして、現在Xの所有に属している。

本件は、書道関係の図書を出版する出版社並びに編集発行者である被上告人Yら（被告・被控訴人）が、X所有の「顔真卿自書建中告身帖」（原作品）の複製物を製作^[13]、販売したことに對し、XがYらの行為自体をXの「顔真卿自書建中告身帖」に対する所有権の侵害であるとし、Yらに對し本件出版物の販売の差止及びその廃棄を請求した事案である。

第一審、控訴審とも、Yらの本件出版物の製作販売は、Xの「顔真卿自書建中告身帖」に対する所有権を侵害するものではない旨判示され、Xの請求を棄却した。これを不服としてXが上告したのが本件である。

【判旨】上告棄却

「美術の著作物の原作品は、それ自体有体物であるが、同時に無体物である美術の著作物を体现しているものというべきところ、所有権は有体物をその客体とする権利であるから、美術の著作物の原作品に対する所有権は、その有体物の面に対する排他的支配権能であるにとどまり、無体物である美術の著作物自体を直接排他的に支配する権能ではないと解するのが相当である」^[14]。

「したがって、著作権が消滅しても、そのことにより、所有権が、無体物としての面に対する排他的支配権能までも手中に収め、所有権の一内容として著作権と同様の保護を与えられることになると解することはできないのであって、著作権の消滅後に第三者が有体物としての美術の著作物の原作品に対する排他的支配権能をおかすことなく原作品の著作物の面を利用したとしても、右行為は、原作品の所有権を侵害するものではないというべきである」^[15]とした。その上で、著作権が現存しない本件の場合に、Yらの行為（本件出版物の製作販売）は、Xの所有する「顔真卿自書建中告身帖」の排他的支配をおかすものではなく、所有権をなんら侵害するものではないと、原審同様、Xの請求を棄却した。

【解説とコメント】

本件は、X所有の「顔真卿自書建中告身帖」（原作品）複製物の製作・販売を、本件原作品の所有権に基づき出版物の販売差止並びにその廃棄を請

求することが可能か否かが問題となった事案である。

判決は、「所有権は有体物をその客体とする権利であるから、美術の著作物の原作品に対する所有権は、その有体物の面に対する排他的支配権能であるにとどまり、無体物である美術の著作物自体を直接排他的に支配する権能ではない」とし、著作権と所有権とを明確に峻別した。他方、「第三者が…原作品に対する排他的支配権能をおかすことなく原作品の著作物の面を利用したとしても、…原作品の所有権を侵害するものではない」と判示し、第三者が原作品の排他的支配権能をおかした場合には所有権侵害を構成することに含みを残し、而後の判決の展開を示唆している。

◇ 三 判例に見る法的保護の展開

(一)「広告用ガス気球」事件^[16]

【事案の概要】

広告宣伝会社である控訴人X(原告)は、西ドイツの訴外Aからガス気球を輸入し、これをあるイベントの会場シンボルとしてその主催者訴外Bに賃貸した。Bは、このイベントの宣伝の一環として本件気球を公開した。その際、職業写真家である訴外Cが営利を図る目的をもって、本件気球を撮影し、その後撮影した写真を第三者に賃貸するため写真代理店業を営む訴外Dに本件写真の賃貸につき仲介を委託した。

数ヵ月後、被控訴人Y₁(被告)は自らが製造・販売する自動車の宣伝のため、また被控訴人Y₂(被告)が製造するカー・エアコンの宣伝のため、広告宣伝を営業目的とする訴外Eに広告ポスターの製作を共同で注文した。Eは、広告ポスターの素材として使用するのに適当な気球を含む数枚の写真をDから借り受け、これを使ってポスターを製作し、Y₁に引き渡した。Y₁はこの広告ポスターを営業所・代理店宛てに直接発送し掲示をした。

Xは、Yらに対して、Yらの本件ポスターの使用に因り、広告媒体としての新規性・独占性を喪

失したとし、結果、本件気球を他に賃貸することが不可能となったと主張し、賃料相当額800万円にのぼる逸失利益の損害賠償を請求した。

第一審では、「本件において、Yらは本件気球の「影像」を利用したにすぎないものではあるけれども、その手段の故を以って、かかる利用が許される、と判断することはでき」^[17]ず、「Yらが、本件ポスターの…使用により、本件気球(の影像)を利用したことは違法である、と言わざるを得ない」^[18]と判示した。しかし、「Eは広告宣伝を営業目的とする株式会社であり、広告ポスター製作についての所謂「専門企業」である。このような双方の立場および広告ポスター製作に要請される技術的専門性・能力的特殊性に鑑みれば、Yらは、本件ポスターが…Eにおいてその企画の設定から素材の利用(本件写真の利用を含む)に至るまで適法に処理した上で製作されたものである、と信じていた事実を推認することができる。…されば、Yらには、本件ポスターの…利用に際して、Xが本件気球について有する権利を妨害もしくは侵害するであろうことを予見し得る余地はなかった、と言うべきである」^[19]とし、Xの請求を棄却した。Xは控訴したが、東京高裁は次のように判示して、原審同様、Xの請求を棄却した。

【判旨】請求棄却

「一般に、物の所有者は、その所有権の範囲を逸脱し若しくは他人の権利・利益を侵害する結果となるような場合を除き、その所有物を如何なる手段・方法によっても使用収益することができ、第三者は、所有者から使用収益を承認されている場合を除いては、直接にせよ間接にせよ、他人の所有物を利用することによって所有者の使用収益を阻害してはならない法的関係にあるものといわなければならない。これを本件気球についていえば、右気球を…広告媒体として使用することにより利益をあげることが広告宣伝業者たるXの所有目的と認められ…、第三者がXの右目的の実現に先立って、右気球を特定…の宣伝に利用し、それにより、…右気球に特定の…イメージを密着させてしまつて、所有者であるXが使用収益の目的を達成することを不可能にしたとすれば、X

が右気球の所有者として有する利益を侵害したものであるというべく、かかるXの所有目的及びこれを阻害する結果の発生を予見しうべき地位にある第三者が、あえて前記のような挙に出たときは、Xに対し損害賠償の責に任ずべき場合の生じうることは、これを否定することができない。そして、本件におけるように、気球を撮影した写真を素材とするポスターによって宣伝行為をすることも、ここにいう右気球の利用に含まれると解すべきである」^[20]。

「しかしながら、…本件ポスターは、この種ポスター作成の専門業者であるEが、Y₁を注文者とする製作物供給契約に基づき、独自の判断と手法により作成して納入したものであって、注文者たるY₁は、作成されるべきポスターの内容につき広告効果の見地から意見を述べたにとどまり、ポスターの素材となる気球の影像を如何にして用意するかというような、仕事を完成させるための手段・方法等は、すべてEがその責任において選択決定のうえ作業を進めた」。「したがって、本件ポスターに使用された気球の影像が、特定企業ないし特定商品の広告媒体として使用することを予定している宣伝業者の所有する気球を撮影した写真によったもので、これを本件ポスターに利用することにより、所有者の右使用目的を阻害する結果をも招きかねないものであることまでを予見し、所有者の使用許諾の有無等を確認すべき注意義務を注文者たるY₁が負い、右注意義務を尽くすことなく本件ポスターを作成納付させて使用した点においてY₁に過失があったとすることはできない」^[21]。

【解説とコメント】

本判決は、原審同様、Y₁による本件ポスターの利用がXの使用目的を阻害する結果の招来を予見することはできず、Y₁による過失は認められないとし、Xの請求を棄却したが、本件気球を撮影した写真をポスターとして宣伝行為に利用した行為は、本件気球の所有者Xが使用収益の目的を達成することを不可能にするような行為として評価され、所有者の利益を侵害し所有者に対し損害賠償の責を負うとされた。その根拠として、

所有権の使用収益機能が指摘されている。

（二）「長尾鶏」事件^[22]

【事案の概要】

長尾鶏（以下「本件長尾鶏」）を飼育している被告Yは、原告Xが被告の承諾を得ずに、本件長尾鶏を写真にとり、これを複製し観光写真を作成して高知県下の観光業者等に販売したことに對し、著作権侵害を理由とする損害賠償請求訴訟を提起したが、この訴えが後になって請求放棄により終了したことを受け、この訴訟を提起したことそれ自体が、不法行為に当たるとして原告Xが被告Yに対して損害賠償を請求した事案である。

【判旨】請求棄却

まず、本件長尾鶏の著作物該当性についての検討がなされた。すなわち「本件長尾鶏が…著作物と認められるかを検討してみるに、…本件長尾鶏が思想または感情を表現した美しいものであるとは考えられるものの、このことから更に本件長尾鶏が著作権法2条1項1号に定める思想等を創作的に表現した美術の範囲に属するものとみるのはいささか無理であると思われる」^[23]とし、その該当性につき否定的ないし消極的な立場を示した。その上で、「本件長尾鶏には、…独特な美しさがあり、その管理、飼育にもそれなりの工夫と人知れぬ苦勞があり、永年の努力の積み重ねの結果、ようやくにしてこれが育て上げられたものであることを考えると、本件長尾鶏を写真にとったうえ絵葉書等に複製し、他に販売することは、右長尾鶏所有者の権利の範囲内に属するものというべく、その所有者の承諾を得ることなくして右写真を複製して絵葉書にして他に販売する所為は、右所有者の権利を侵害するものとして不法行為の要件を備えるものとみられ、右権利を侵害した者はその損害を賠償する義務がある」^[24]とし、…そう考えると、先に取り下げられた訴えは「その主張する権利が立証不能な違法不当なものであるとまではいえないこととな」^[25]り、「Yが提起した訴えがX主張の如き不法行為に当たるとは認め難いので、更にその余の点について判断するまでもなくXの請求は失当としてこれを棄却」^[26]

するとした。

【解説とコメント】

本件では、本件長尾鶏が著作物として認められなくとも、それ自体独特な美しさがあり、その管理、飼育にもそれなりの工夫と人知れぬ苦労があり、永年の努力の積み重ねの結果、育て上げられたものであることを認定し、本件長尾鶏を写真に撮り、絵葉書等に複製・販売することは、長尾鶏所有者の権利の範囲内に属し、かつ無許諾で本件長尾鶏の写真を複製・販売することは、所有権者の権利を侵害するものとして不法行為を構成するものとした。

(三)「クルーザー」事件^[27]

【事案の概要】

原告Xは、大型サロンクルーザー「プリンス号」（以下「本件クルーザー」）を所有し、自らが経営するマリンホテルのシンボルとしてホテル利用客の観光用等の用途に使用していた。本件クルーザーは、被告Yが輸入し、訴外Aに販売し、その後原告Xに転売したものであった。Yは、Aに対し、本件クルーザーの写真を自らの宣伝広告用に雑誌に掲載することなどの販売協力を依頼しその承諾を得たが、それとほぼ同じ時期に本件クルーザーの所有者がXに移転したことを知っていながら、Xに対する所要の手続をとらず、月刊雑誌「オーシャンライフ」に本件クルーザーの写真をユーザーの使用例として掲載した。

Xは、写真掲載によって、「本件クルーザーが売りに出されている」ばかりではなく、自らが経営するホテルも「経営悪化により売りに出されている」等の噂が広がり、その名誉が毀損されたばかりではなく、その噂によって従業員が辞めていき、売上が減少し、廃業やむなさに至ったと主張し、損害賠償の請求と謝罪広告の掲載を求めた。

【判旨】一部認容、一部棄却

「Xは、本件クルーザーの所有者として、同艇の写真等が第三者によって無断でその宣伝広告等に使用されることがない権利を有していること明らかである」^[28]。

「Yは、本年クルーザーの前所有者の時にその

承諾を得て、平成元年2月上旬頃、同艇の写真を自己の宣伝広告用に「オーシャンライフ」に掲載することの依頼をしたが、同月末頃には本件クルーザーの所有権が右前所有者からXへ移転したことを知ったのであるから、Yは、その時点で、新所有者であるXに対し、改めて右写真の雑誌掲載についてその承諾を得るべき義務があったといふべきである。しかし、Yは、Xに対してその承諾を求めることもせず、また右雑誌掲載中止の手続をとることもせずに、そのまま掲載したものであるから、Yは、Xに無断で本件クルーザーの写真を自己の宣伝広告用に右雑誌に掲載したこととなり、これによってXが蒙った損害を賠償すべき責任があるといわざるを得ない」^[29]。

【解説とコメント】

本件では、本件クルーザーの所有者であるXが、同艇の写真等を第三者によって無断で宣伝広告等に使用されない権利を有し、Yの写真掲載行為によってXが蒙った損害を賠償すべき責任があるとされた。

(四)「かえでの木」事件^[30]

【事案の概要】

本件は、高さ15メートルにも及ぶ大きなかえでの木（以下「本件かえで」）を所有する原告Xが、本件かえでの写真を掲載した書籍を出版、販売等した被告Y₁（出版社）に対して、本件かえでの所有権に基づき上記書籍の出版等の差止めを、被告Y₁及び上記書籍に掲載された写真を撮影した被告Y₂（フリーカメラマン）に対して、本件かえでの所有権侵害による不法行為に基づき損害賠償の支払いを求めた事案である。

本件において、原告Xは、被告Y₁に対して、本件かえでの所有権に基づいて、本件書籍の出版等の差止め及び本件書籍の回収を請求することができるか。また、原告Xは本件かえでの所有権侵害の不法行為に基づいて、被告Y₂が本件かえでの写真を撮影し、被告らが本件書籍を出版したことについて損害賠償を請求することができるかが争点となった。

【判旨】請求棄却

「所有権は有体物をその客体とする権利であるから、本件かえでに対する所有権の内容は、有体物としての本件かえでを排他的に支配する権能にとどまるのであって、本件かえでを撮影した写真を複製したり、複製物を掲載した書籍を出版したりする排他的権能を包含するものではない。そして、第三者が本件かえでを撮影した写真を複製したり、複製物を掲載した書籍を出版、販売したとしても、有体物としての本件かえでを排他的に支配する権能を侵害したということとはできない。したがって、本件書籍を出版、販売等したことにより、Xの本件かえでに対する所有権が侵害されたということとはできない」^[31]。

「本件かえでを撮影し、その写真を掲載した本件書籍を出版、販売等したことにより、Xの本件かえでに対する所有権が侵害されたということとはできない。また、本件全証拠によっても、被告Y₂が、本件かえでの枝を折るなど、本件かえでの所有権を侵害する行為を行ったと認めることはできない」^[32]。

【解説とコメント】

本件では、所有権はあくまでも有体物をその客体とする権利であるから、本件かえでに対する所有権の内容は、もっぱら有体物としての排他的支配権能にとどまり、本件かえでの写真の複製や複製物を掲載した書籍の出版についての排他的権能を含まないこと、また、第三者による本件かえでの写真の複製や、複製物を掲載した書籍の出版・販売によって本件かえでの排他的支配権能は侵害されないことから、本件書籍の出版・販売等により本件かえでの所有権は侵害されなかったとした。

◇ 四 おわりに

冒頭において触れた「顔真卿自書建中告身帖」事件では、美術の著作物の原作品による複製物の製作・販売行為に対し、原作品の所有権を根拠として、出版物の販売差止・廃棄を命ずることが可能か否かが争われた。判決は、美術の著作物の原作品に対する所有権と美術の著作物自体を明確に

峻別した上で、第三者が原作品に対する排他的支配権能をおかすことなく原作品の著作物の面を利用したとしても、原作品の所有権を侵害するものではないとした。

他方で、判決は、別の箇所において、著作権が消滅した後、第三者が美術の著作物の複製物を出版すると、原作品の所有権者への許諾を求める者が減少し、結果、収益をあげる機会が奪われ、経済上の不利益を受けることを認めている。しかし、この第三者による複製物の出版が、原作品の排他的支配をおかすことなく行われたときには、この複製物の出版は既に公有に帰した著作物を利用するに過ぎないのだから、この行為が原作品の所有者の経済上の不利益をもたらしたとしても、これは第三者が著作物を自由に利用することができる事実上の結果であり、この行為を捉えて所有権者の原作品に対する使用収益機能を違法におかしたとは認められない^[33]とし、第三者が原作品の排他的支配権能をおかした場合には所有権侵害を構成することもありうることを示し、而後の判決の展開を示唆している。

そこで問題となるのが、如何なる場合に原作品の排他的支配権能をおかすこととなり、所有権侵害を構成することになるのかという点である。この点、従来判例によれば、①気球が広告媒体として使用することで利益をあげることを目的としているところ、第三者がこの目的に先立ち、気球を撮影した写真を素材とするポスターを別の広告宣伝に利用したために所期の目的が達成困難となった場合（「広告用ガス気球」事件）、②それ自体独特な美しさを備え、その管理・飼育にも工夫と苦勞を要し育てられた長尾鶏にあって、その写真を絵葉書等に複製・販売する場合（「長尾鶏」事件）に、それぞれの行為は所有権を侵害し、不法行為を構成するとされている。①においては、排他的支配権能、すなわち使用収益機能の侵害の態様が、具体的目的との関係で詳細な認定を行っている一方で、判決後半部分では、それと同じウェイトで予見可能性ないし過失論の中で損害賠償責任を排除できるロジックが用意されている。また、本件においては特有の事情がかなり考慮されたよ

う見受けられ、本判決の射程はかなり限定的であるといつてよい。また②につき、管理・飼育に対する一定の努力を所有権との関係で法的に保護すべきものと捉えているようであるが、本件は、当該複製販売行為それ自体が不法行為を構成するか否かが争われた例ではないため、この先例的意義はかなり乏しい。

以上からも明かなように、所有権の排他的支配機能の侵害といっても未だ不確定のところが多分にあることは否めないように思われる。

これらに対し、「クルーザー」事件では、所有権の内容たる排他的支配権能について何ら言及することなく、「本件クルーザーの所有者として、同艇の写真等が第三者によって無断にその宣伝広告等に使用されることがない権利」を有するとして、不法行為の成立を認めている。ここでは、所有権との関係が必ずしも明確ではなく、かかる権利が認められる背景的事実に関する認定も十分とは思われないが、所有権の機能とは異なる見地から不法行為の構成を指摘している点、他の判例と比べてユニークなものであると評価できよう。

ところで、最近の判決である「かえでの木」事件は、「顔真卿自書建中告身帖」事件同様の考え方に立ち、本件かえでのに対する所有権の内容は、もっぱら有体物としての排他的支配権能にとどまり、本件かえでのの写真の複製や複製物を掲載した書籍の出版についての排他的権能を含まないこと、また、第三者による本件かえでのの写真の複製や、複製物を掲載した書籍の出版・販売によって本件かえでのの排他的支配権能が侵害されない限り、本件かえでのの所有権は侵害されない旨判示した。

排他的支配権能の侵害の有無（損害賠償請求の根拠）につき、判旨は、現に被告Y₂が本件かえでのの枝を折るなど、本件かえでのの所有権を侵害する行為を行ったとは認められない^[34]と認定していることから明かなように、専ら本件かえでのの所有権侵害、しかも直接的・物理的侵害を指摘し、それらとの関係で評価がなされている点、注意を要する。

なお、「かえでの木」事件判決では、さらに加えて「念のために」本件かえでのの所有権以外の法

的利益侵害についての検討が行われている^[35]。

判決では、①原告は平成12年7月頃までは本件かえでのを鑑賞するために平穏な態様で原告所有の土地へ立ち入ることを一般に容認していたものと認められること。②本件かえでのの保全を図るため、看板を立て、根を踏む等の本件かえでのの生育に悪影響を及ぼす行為や、営利目的での土地への立ち入り及び撮影に制限を設けるなどの一定の条件を設定したが、観光客による本件かえでのの鑑賞や、私的な目的での撮影のために土地に立ち入ることを何ら禁止していなかったこと。③被告が当該看板設置以降撮影をしていないこと、また、本件かえでのの生育に悪影響を及ぼす可能性のある行為をしていないこと、を認定し、原告の「本件土地」の所有権侵害の可能性、さらにはその他の法的保護に値する利益侵害の可能性をも検討している^[36]。

本判決に見るように、「本件かえで」の所有権ではなく、「本件土地」の所有権に目を向けることは、博物館・美術館等における収藏品へのアクセスの法的性格ないし料金徴収の根拠を考える際に格好の素材を提供してくれる。

冒頭において取り上げた「顔真卿自書建中告身帖」事件において、最高裁判所は、博物館・美術館における料金徴収の根拠を次のように述べる。すなわち、「博物館や美術館において、著作権が現存しない著作物の原作品の鑑賞や写真撮影について料金を徴収し、あるいは写真撮影するのに許可を要するとしているのは、原作品の有体物の面に対する所有権に縁由するものと解すべきであるから、右の料金の徴収等の事実、所有権が無体物の面を支配する権能までも含むものとする根拠とはなりえない。料金の徴収等の事実、一見所有権者が無体物である著作物の複製等を許諾する権利を専有することを示しているかのようにみえるとしても、それは、所有権者が無体物である著作物を体現している有体物としての原作品を所有していることから生じる反射的效果にすぎないのである」^[37]と。

最高裁は、被上告人の主張に応えるかたちで上のように述べたわけだが、「かえでの木」事件判

決を前提とすれば、これが必ずしも十分に正確ではないことが明らかとなる。すなわち美術館や博物館が所蔵品につき一定の条件を付したり、料金などを徴収するのは、所蔵品のそれ自体の所有権に基づいているのではなく、美術館や博物館の土地の所有権（管理権）に由来しているにすぎないということなのである。

「かえでの木」事件判決は、最後に付言して次のように言う。「原告は、本件かえでの所有権に基づき…各行為を阻止できない限り、本件かえでを保全することができない旨述べる。しかし、原告が、本件土地に所在する本件かえでの生育環境の悪化を憂慮して、本件かえでの生育等に悪影響を及ぼすような第三者の行為を阻止するためであれば、本件土地の所有権の作用により、本件かえでを保全する目的を達成することができる。…現に、原告は、本件土地への立ち入りに際しては、本件かえでの生育等に悪影響を及ぼす可能性のある行為をしてはならないこと、許可なく本件かえでを営利目的で撮影してはならないことを公示しているのであるから、第三者が上記の趣旨に反して本件土地へ立ち入る場合には、原告は当該立ち入り行為を排除することもできるし、上記第三者には不法行為も成立する。また、本件土地内に、美観を損ねないような柵を設けること等によって、より確実に上記目的を達成することもできるというべきである」^[38]と。なお、「かえでの木」事件は、「本件土地」の所有権侵害可能性以外のその他の法的保護に値する利益侵害可能性の認定は行っていない。

注

- [1] 本稿の表題に掲げた「無体財産」なる語は、近年、「知的財産」と呼ばれるようになったが、ここではかかる財貨の無形の側面に着目し、論旨を展開していくことの関係であえてこの語を使うこととした。
- [2] 「著作権問題」というよりも、より広く「法律問題」といった方が、ここで交わされた議論の内容を示すには適当かもしれない。
- [3] 原告をX、被告をYとし、その他訴外A・B・C…と呼ぶ。

また、それぞれの事案は、独立のものとして取り扱う。

- [4] 斉藤 博『著作権法』（有斐閣、2000 年）1 頁。
- [5] 民法 85 条は「本法ニ於テ物トハ有体物ヲ謂フ」と規定する。
- [6] しかし、今日、無体物の上に所有権などの物権が成立することも認められ、いまやこの規定は所有権の客体を限定する意味を有していない旨がしばしば指摘される。ここでは、民法における「物」概念の確定や現代的な把握が主題ではないので、これ以上踏み込むことはしないでおくが、ここでは一応、この規定を所有権の対象ないし客体として典型的なものを示しており、「物」の範囲を類推・拡張によって捉える際の根拠ないし指針として把握しておきたい。
- [7] 民法 206 条は「所有者ハ法令ノ制限内ニ於テ自由ニ其所有物ノ使用、収益及ヒ処分ヲ為ス権利ヲ有ス」と規定する。
- [8] 著作権法 25 条は「著作者は、その美術の著作物又はまだ発行されていない写真の著作物をこれらの原作品により公に展示する権利を専有する」と規定する。
- [9] 著作権法 45 条 1 項は「美術の著作物若しくは写真の著作物の原作品の所有者又はその同意を得た者は、これらの著作物をその原作品により公に展示することができる」と定め、同条 2 項において「前項の規定は、美術の著作物の原作品を街路、公園その他一般公衆に開放されている屋外の場所又は建造物の外壁その他一般公衆の見やすい屋外の場所に恒常的に設置する場合には、適用しない」と定めている。
- [10] 加戸 守行『著作権法逐条講義』（著作権情報センター、四訂新版、2003 年）（296 頁）。
- [11] しばしば、「物のパブリシティー権」と呼ばれているものである（伊藤真「物のパブリシティー権」『田倉整先生古稀記念・知的財産をめぐる諸問題』（発明協会、1996 年）507 頁）等。
- [12] 書籍所有権侵害禁止請求事件、最高裁昭和 58 年（オ）171 号、昭和 59 年 1 月 20 日第 2 小法廷判決、上告棄却（民集 38 卷 1 号 1 頁、判例時報 1107 号 127 頁）。
- [13] なお、事実認定によれば、Y らの本件出版物は、原作品から複製したものではなく、原作品の元の所有者 A から許諾を受けて写真撮影した人 B の承継人 C が持っていたものを譲り受け、この原作品の写真乾板を用い複製物である本件出版物を製作したという事実が認定

されている。

[14]「顔真卿自書建中告身帖」事件・書籍所有権侵害禁止請求事件、最高裁昭和58年（オ）171号、昭和59年1月20日第2小法廷判決・判例時報1107号127頁〔128頁〕。

[15] 同上。

[16] 控訴審：損害賠償等請求控訴事件、東京高裁昭和52年（ネ）792号、昭和53年9月28日民事8部判決（著作権判例集Ⅲ846頁、肖像権関係判例集859頁）、原審：損害賠償等請求事件、東京地裁昭和48年（ワ）7540号、昭和52年3月17日民事15部判決、棄却（控訴）（判例時報863号64頁）。

[17]「広告用ガス気球」事件第一審判決・損害賠償等請求事件、東京地裁昭和48年（ワ）7540号、昭和52年3月17日民事15部判決・判例時報863号64頁〔69頁〕。

[18] 同上〔70頁〕。

[19] 同上。

[20]「広告用ガス気球」事件控訴審判決・損害賠償等請求控訴事件、東京高裁昭和52年（ネ）792号、昭和53年9月28日民事8部判決（著作権判例集Ⅲ846頁、肖像権関係判例集859頁）

[21] 同上。

[22] 損害賠償請求事件、高知地裁昭和57年（ワ）31号、昭和59年10月29日判決、棄却（控訴）（判例タイムス559号291頁）。

[23]「長尾鶏」事件・損害賠償請求事件、高知地裁昭和57年（ワ）31号、昭和59年10月29日判決・判例タイムス559号291頁〔293-294頁〕。

[24] 同上〔294頁〕。

[25] 同上。

[26] 同上〔295頁〕。

[27] 損害賠償等請求事件、神戸地裁平成元年（ワ）265号、平成3年11月28日判決、一部認容一部棄却（確定）（判例時報1412号136頁）。

[28]「クルーザー」事件・損害賠償等請求事件、神戸地裁平成元年（ワ）265号、平成3年11月28日判決・判例時報1412号136頁〔138頁〕。

[29] 同上。

[30] 出版差止等請求事件、東京地裁平成14年（ワ）1157号、平成14年7月3日民事29部判決、棄却（確定）（判例時報1793号128頁）。

[31]「かえでの木」事件・出版差止等請求事件、東京地裁平成14年（ワ）1157号、平成14年7月3日民事29部判決、棄却（確定）判例時報1793号128頁〔132頁〕。

[32] 同上。

[33]「顔真卿自書建中告身帖」事件・書籍所有権侵害禁止請求事件、最高裁昭和58年（オ）171号、昭和59年1月20日第2小法廷判決・判例時報1107号127頁〔128-129頁〕。

[34]「かえでの木」事件・出版差止等請求事件、東京地裁平成14年（ワ）1157号、平成14年7月3日民事29部判決、棄却（確定）判例時報1793号128頁〔132頁〕。

[35] 同上。

[36] 同上。

[37]「顔真卿自書建中告身帖」事件・書籍所有権侵害禁止請求事件、最高裁昭和58年（オ）171号、昭和59年1月20日第2小法廷判決・判例時報1107号127頁〔129頁〕。

[38]「かえでの木」事件・出版差止等請求事件、東京地裁平成14年（ワ）1157号、平成14年7月3日民事29部判決、棄却（確定）判例時報1793号128頁〔133頁〕。

※ 石岡氏には、今後数回にわたり、東洋学情報化と法律問題について連載して頂けることになっている。（編集局）

著作権についての知識を深めよう

東洋学のための著作権サイト・ページ指南

小島 浩之（こじま ひろゆき）

本稿では主として東洋学を専攻する研究者や学生が、著作権の知識を得るのに役立つ Web サイト、ページを紹介する。なお中国のサイトについては千田大介氏、台湾のサイトについては野村英登氏よりご助言をいただいた。

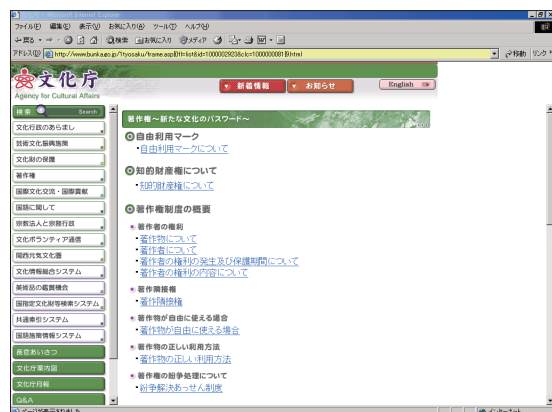
◇ 1. 著作権法一般（国内を中心として）

（1）立法府および行政府からの提供情報

●国会会議録検索システム

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

国立国会図書館が提供するシステム。1 回国会（1947 年 5 月開会）以降の本会議、全ての委員会等の会議録全文が検索できる。著作権関係のキーワードを入力することで、立法府での著作権関係の審議内容や議論の内容を窺うことが可能で



ある。

●文化庁著作権課

[http://www.bunka.go.jp/1tyosaku/frame.
asp?Of1=list&id=1000002923&clc=100000
0081{9.html](http://www.bunka.go.jp/1tyosaku/frame.asp?Of1=list&id=1000002923&clc=1000000081{9.html)

行政府として著作権関係の事務処理を統轄しているのが文化庁著作権課である。著作権法の概要など基本的な知識に加え、常に最新の情報を入手することができる。ただしここに載せられている見解は、あくまで行政府としてのものであり、実際の法的可否については司法府たる裁判所の判断に留意しなければならない。

●文部科学省審議会情報

[http://www.mext.go.jp/b_menu/shingi/
index.htm](http://www.mext.go.jp/b_menu/shingi/index.htm)

著作権法では、補償金の額を定める場合、文化庁長官が審議会に諮問することを定めている（71 条）。また著作権法の改正など著作権制度に関する重要事項についても、原則審議会を経て文化庁長官に答申される。平成 12 年までは著作権審議会がその任に当たっていたが、中央省庁の再編により平成 13 年以降は文化審議会の著作権分科会がこれを担当している。このページでは主として 1998 年前後からの両審議会の答申、議事録、名簿などを読むことができる。

●法令データ提供システム

<http://law.e-gov.go.jp/cgi-bin/idxsearch.cgi>

総務省行政管理局が官報を元に作成したデータベース。法令（憲法・法律・政令・勅令・府令・

省令・規則)の内容を検索できる。検索には“法令索引検索”と“法令用語検索”の二種がある。“法令索引検索”が指定した法令の全文を閲覧できるのに対し、“法令用語検索”では、全法令の中から指定した用語を含む条項だけをピックアップしてくれる。

(2) 判例情報

●裁判所

<http://www.courts.go.jp/>

著作権に関する司法判断は上記裁判所 Web サイト中の次のページで確認することができる。

1) 知的財産権裁判例集

[http://courtdomino2.courts.go.jp/chizai.nsf/\\$Help](http://courtdomino2.courts.go.jp/chizai.nsf/$Help)

2) 最高裁判例集

<http://courtdomino2.courts.go.jp/schanrei.nsf>

1) では高等裁判所以下の下級審における昭和44(1969)年以降の判例を検索できる。著作権をも含む知的財産権のみで、一つの判例検索システムが構築されているのは非常に興味深い。2)は最高裁判所の判例に関する検索システム。収録は『最高裁判所判例集』の範囲内ということなので、民事は昭和22(1947)年9月、刑事は同年11月ということになる。なお上記データベースに収録されていない最新の判決については、“知的財産権判決速報”([http://courtdomino2.courts.go.jp/chizai.nsf/\\$About](http://courtdomino2.courts.go.jp/chizai.nsf/$About))および“最近の最高裁判決”(<http://courtdomino2.courts.go.jp/judge.nsf/View1?OpenView>)を参照のこと。

(3) その他の団体によるサイト

●社団法人著作権情報センター

<http://www.cric.or.jp/>

著作権思想の普及事業や、著作権関連の調査研究を行っている社団法人のサイト。著作権に関する多くの情報を入手することができ、電話による無料の著作権相談も行っている。“著作権文献・資料データベース”では、著作権関連の知的所有

権についての図書、論文、記事等のデータ(収録範囲:1993-2001)を検索でき、“著作権審議会報告”では、先に紹介した“文部科学省審議会情報”には収録されていない過去の審議会報告書を閲覧できる。“外国著作権法令集”ではアメリカ、フランス、イタリアと並んで中華人民共和国著作権法の邦訳があるのも嬉しい。タイプ別に初心者向けの冊子も発行しており、全て Web 上で閲覧できる。我々のように法学の専門家でない者が、著作権について学びたい場合、最初に覗いてみるべきサイトだと言える。

●CI-NIME 教育メディア著作権関連情報

<http://deneb.nime.ac.jp/>

大学共同利用機関であるメディア教育開発センター(http://www.nime.ac.jp/index_ie.html)が提供するサイト。高等教育においてマルチメディア資源を活用する際に、必要となる著作権上の知識等を提供する。教育と著作権についての情報が満載なので、高等教育に携わる方には必見のサイトである。

情報の提供だけでなく、権利者との権利関係調整のためのシステム開発も行っている。著作物の利用許諾契約を迅速に得るためのシステム“著作権等流通促進システム”は興味深い試みである(詳細はhttp://www.nime.ac.jp/nnl/no30/4_23_00p.htmlを参照)。このシステムを利用して許諾を得られる著作物はまだ多くないが、著作権関係の判例が公開対象に含まれており利用価値がある。ただし利用者登録(無料)が必要で、かつ高等教育機関の教員等だけに利用を制限している。また同システムはサイト内におけるコンテンツとしての位置が解りにくい(コンテンツの一つである“著作権に関する知識”のページ最下部にリンクがある)。利用を促進したいのならばコンテンツ内の位置づけを再考すべきであろう。

(4) 学会

●著作権法学会

<http://www2.odn.ne.jp/~aaf77690/>

昭和37(1962)年設立の学会。サイトの主要内容は活動案内であり、一般に得べき情報は

少ない。

●情報ネットワーク法学会

<http://www.in-law.jp/>

2002年に設立された新しい学会。情報ネットワーク社会における法を広く取り扱うので、参加者も研究者、政府機関、企業など幅広い。著作権法専門の学会ではないが、情報、ネットワークといったことに関する法律は、著作権とは切り離せないもので、常に注意を払う必要がある。メールマガジン“IN-LAW Newsletter”を月に2回発行しているので、興味のある方は購読してみてもいいかであろうか。

（5）個人サイト

●Net Law

<http://www.netlaw.co.jp/>

弁護士荒竹純一氏のサイト。著作権の情報が上手くまとめられており安心して見ることができる。氏の著書『知らないと危ない・インターネットと著作権』（中央経済社、1997）の完全改訂版、および『インターネット著作権：知っておきたいITビジネスの法知識』（ソフトバンク・パブリッシング社）のWeb版を提供しており有用である。

●サイバースペースの法律

<http://www.law.co.jp/okamura/index.html>

弁護士であり、近畿大学で「情報法」の講義も担当する岡村久道氏のサイト。氏は先に紹介した“情報ネットワーク法学会”の理事も務めている。ここでは大学での講義ノートや論文等が公開されている。コンテンツの一つであるリンク集“著作権法入門”は重宝する。

●駒田亭文庫

<http://www1.odn.ne.jp/~aac13570/>

群馬大学駒田泰土氏のサイト。著者の論文や私的な書評メモなどが公開されており参考になる。

●リンク問題資料集

http://www-h.yamagata-u.ac.jp/~matumoto/link/link_index.htm

山形大学松本邦彦氏のページのコンテンツの一つ。

インターネットと著作権ということで、必ず話

題に上げられるものにリンク問題がある。すなわちリンクを自由に貼ることが著作権侵害に該当するのかという問題である。現在ではフレーム内に取り込むようなリンクを除き、原則リンクは著作権法上問題ないと思えるむきが主流のようだが、反論・異論もないわけではない。標記のサイトはこの問題についての論著や実例をまとめたもので、インターネットで情報を発信する立場の方は一読されることをお勧めする。

●知的財産用語辞典

<http://www.furutani.co.jp/index.html>

弁理士である古谷栄男、松下正、眞島宏明の3氏が運営する、特許、著作権、商標など知的財産権に関するサイト。用語辞典がメインのサイトではあるが、3氏の論文や、初心者用の解説文書も用意されている。

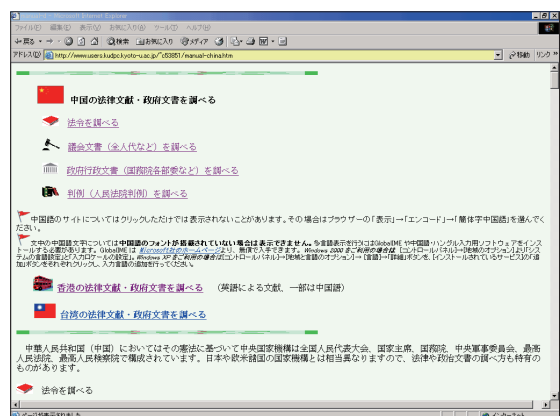
◇ 2. 中国・台湾の著作権法

（1）日本語による情報提供

●京都大学大学院法学研究科附属国際法政文献史料センター

<http://www.users.kudpc.kyoto-u.ac.jp/~c53851/>

インターネット上で、中国、台湾、韓国、東南アジア諸国に関する法律文献、政府文書を入手するためガイドページ。単なるリンク集ではなく、それぞれに丁寧かつ親切な解説が付されている。



1) JETRO 北京センター知的財産室

<http://www.jetro-pkip.org/>

2) 貿易・経済事業 (財団法人 交流協会)

http://www.koryu.or.jp/trade/1_03base.html

1) は大陸、2) は台湾に関する情報を提供する。著作権というより産業分野における知的財産権についての情報という趣が強い。1) では中華人民共和国の関連法規の邦訳 (著作権法も含む)、判例の情報を公開しており、2) では台湾著作権法の邦訳 (台湾知的財産権マニュアル (H10 年度) 内) や、台湾における知的財産権の最新動向を随時紹介している。

(2) 中国

●中華人民共和国国家版權局

<http://www.ncac.gov.cn/>

中国における著作権法の実務機構。著作権関連法規はもちろんのこと、国内外の著作権をめぐるニュース・事件についても掲載されている。

●中華人民共和国国家文物局

<http://www.sach.gov.cn/>

中国の文物の写真を撮影したり、それを研究上利用する際に、著作権法以外に文物保護関係の法律に抵触する可能性がある。これら文化財保護関連の法令は国家文物局の“政策法規”にまとめられている。

●北大法律信網

<http://211.100.18.62/>

北京大学法制信息中心が運営するサイト。中



国における法律のポータルサイトの役割を果たしており、リンクも充実している。法令、判例のデータベースもあるが、原則会員制 (有料) である。なお無料のデータベースとしては法律法規庫 (<http://search.law.com.cn/>) などがある。

●中国知識産権司法保護

<http://www.chinaiprlaw.com/default.htm>

最高人民法院判事蔣志培氏のサイト。中国の知的財産権関連の法令、判例、情報に加え、氏の論考も掲載する。情報の更新もこまめで早い。著作権法に特化しているわけではないが、中国の知的財産権に関する情報を入手するには非常に有効なサイトである。

●冬涛專欄

<http://www.ccw.com.cn/news2/zl/dongtao.asp>

計算機世界網>中国 IT 新聞>專欄天地の中にあるコンテンツ。冬涛とは、北京市海淀区人民法院知識産權庭 (知的財産権法廷) の判事、李東濤氏のことで、本誌第 3 号のレビューで千田大介氏が紹介した中国電子図書館事件の裁判長を務めた人物である。ここでは主に李氏が扱った知的財産権裁判に関する解説が公開されている。

(3) 台湾

●文化建設相關法規彙編 (行政院文化建設委員會)

<http://www.cca.gov.tw/law/index.html>

日本の文化庁に該当する文化建設委員会サイト内のページ。台湾の著作権、文化財関連の法令を掲載している。

●全国法規資料庫

<http://law.moj.gov.tw/>

法務部全国法規資料庫工作小組が管理するサイト。台湾における法令、判例を検索できる。専門外の一般利用者が検索に困らないように、通俗用語 (法律用語でない言葉) を法律用語に変換するシステムまで用意されている。日本のお役所もこういった心遣いを見習って欲しいものである。

特集2

漢字処理技術の 最新動向

かつて、コンピュータにおける漢字の問題は、収録字数の壁をいかに乗り越えるかという点にあった。しかし今、我々は新たな壁にぶつかっている。Unicodeをはじめとする大規模な漢字コードが規格化されたおかげで便利にはなったが、増えた文字の数だけ便利になったわけではない。規格書に文字はあるが何の字かわからない、使おうと思ってもフォントがない、これだけあってもまだ足りない、といった問題が浮上してきたのである。

文字コード規格が文字数の膨張に比例するように混乱を招いている一方で、文字コードの外側に解を摸索する動きもある。とは言えそれらは現在のところ未成熟であり、当面は両方に目配りをしていかねばなるまい。ここに集められた論者は、いずれも批判精神に満ち、建設的な提案をするものであると自負している。今後の漢字処理をよりよいものにするためにも、本特集が議論を喚起するきっかけになればと願っている。

なお、本特集の一部は、昨年12月の漢字文献情報処理研究会第5回大会の発表に基づくものである。

Contents

CHISE Project

……………守岡 知彦、江渡 浩一郎、苔米地 等流、宮崎 泉、師 茂樹… 58

文字コード外フォント処理……………上地 宏一… 70

Embedding Glyph Identifiers in XML Documents…………… Christian Wittern… 74

JIS X 0213 の改正と UCS との関係について……………川幡 太一… 80

Unicode 4.0……………師 茂樹… 86

CHISE Project

守岡 知彦（もりおか ともひこ）、江渡 浩一郎（えと こういちろう）、
苦米地 等流（とまべち とうる）、宮崎 泉（みやざき いずみ）、
師 茂樹（もろ しげき）

■ 文字コード超克への道

前号の師の記事にあるように^[8]、CHISE Project は、巨大な収録文字数を誇る符号化文字集合やエンコーディングを工夫した優れた文字コードを作ろうとするのではなく、符号化文字技術というものの自体に疑いの目を向け、それに代わるものを再構築することによって文字コードにまつわる問題を抜本的に解決することを目指している。

検索や再利用が容易であるという電子テキストの優れた特徴は符号化文字技術に基づいているが、これはもともと英語用ラテン文字のために設計されたものといえ、英語用ラテン文字との差異が大きい用字系（script）ではしばしば技術的な齟齬に悩まされてきた。そうした齟齬は符号化文字技術を発展させる原動力となり、この技術を豊かにしてきたが、それでも齟齬が無くなった訳ではなく、また、本来単純であった符号化文字技術の複雑化も招いている。漢字においても、まず、文字数の多さが問題になり、その後、大規模文字集合の登場により、何を文字と考え何を符号化するかとか、文字の同一性をどう考えるかなど、符号化文字技術を支える抽象文字概念に関する根源的な問題が顕在化した。そして、文字が増えても依然として外字を必要とする局面は残っている。そして、インターネットの発達によって、各々の外字を大域的に利用可能にするために標準化しようとする欲求が高まったといえる。その結果、符号化文字集

合の標準を拡大しようとする動きは際限なく続いている。このような状況により、文字概念に起因する根源的な問題はむしろ拡大しているといえる。

こうした問題意識に基づいて、守岡は符号化文字技術に依存しない文字表現を模索するようになった。その後、田中哲氏やg新部裕氏らとの議論を通じ、この問題は次のように整理された：

1. 脱コード化：画像のような符号化文字列以外のものでも記号を表現する
2. 文字定義：文字は1の記号を基に、性質や関係を定義することによって定義する
3. (カラーパレット的) 局所的文字符号：2で定義した文字オブジェクトに対応する短い整数値（符号値）が必要な場合は、局所的に利用する整数値を適当に作る。この際、文字オブジェクトとの対応関係を管理する

そして、1997年には1と2に焦点を当てたLisp系記号処理モデルである“My Symbolic System”が構想され、1998年には2と3に焦点を当てた文字処理モデルである“UTF-2000”モデルが構想された。^[11]

1998年にはg新部氏によってUTF-2000 based on GNU Emacs 20.2.90が試作された。そして1999年には守岡によるXEmacs UTF-2000（現在のXEmacs CHISE）^{[12][13]}の開発がスタートし、UTF-2000モデルの具体化が本格化した。

CHISE Project

CHISE Project は前述の UTF-2000 モデル（現在は、これを“Chaon”モデルと呼んでいる）を計算機環境全体に適用することを目指すためのものである。ちなみに、CHISE というのは“CHaracter Information Service Environment”の略で、文字に関する知識処理のための計算機環境というようなものを表している。この中核となるのは計算機環境で共有される文字データベース・システムであり、各種アプリケーションがこの文字データベースを参照して動作することにより計算機環境全体で文字に関する知識が共有されることになる。

UTF-2000 プロジェクトにおいてもこうした計算機環境の実現を目指していたものの開発者は事実上守岡だけで、こうした目標を実現することは困難であった。一方、CHISE Project では新たなコアメンバーの参加によって計算機環境の実現に向けて踏み出すことができた。2001 年度には Christian Wittern 氏が参加するとともに、未踏ソフトウェア創造事業のプロジェクトとして採択され、XEmacs CHISE 内蔵文字データベースの外部化や TopicMaps^[4] エンジンの試作、UCS^{[5][6]} に収録された約 7 万字の漢字などを対象とした漢字構造情報データベースの実現など開発や大規模データ整備が進展した。2002 年度には江渡・鈴木泰博氏・藤原義久氏の可視化グループ、上地宏一氏のグリフ・グループ (KAGE), 苫米地・宮崎の組版グループ (Ω), そして、前述の師が参加し、CHISE Project の開発が本格化した。

CHISE Project の概要に関しては前号の師の記事で紹介されているが、それ以降での重要な成果の 1 つは江渡による Ruby/CHISE と師による Perl/CHISE の実現であろう。これらは Chaon モデルのスクリプト言語への適用の試みであるだけでなく、CHISE 文字データベース・システムを実際に複数の言語処理系から共有可能になったことを意味しており、CHISE の計算機環境を実現する上で非常に重要な一歩であるといえる。ま

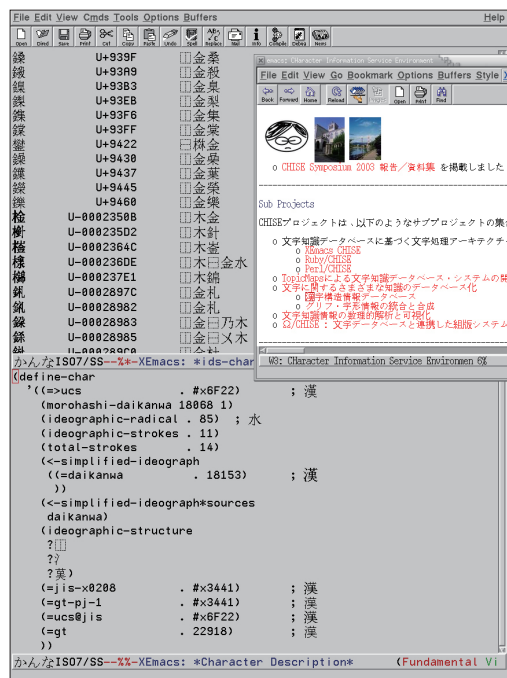
た、多言語組版システム^[7]と漢字字形合成システム KAGE を用いた組版システム Ω /CHISE が苫米地と宮崎により開発され、フォントやグリフデータなどの問題からまだ完全ではないものの、XEmacs CHISE で表示される文字や IDS を使って指定した漢字を TEX で組版・印刷することが可能となった。

XEmacs CHISE

XEmacs CHISE は、師の言う所の『UNIX 界の怪物エディタ (的統合環境)』^[6]のひとつである XEmacs^[11]に基づく Chaon 実装である。XEmacs に内蔵された Emacs Lisp の力と XEmacs CHISE が拡張した文字データベース処理機能を併用することにより、高度な文字知識処理が可能である。このようなことから XEmacs CHISE は文字処理の実験場として優れており、実際、CHISE Project における基本概念の多くはこの開発の中で育ってきたといえる。

このように XEmacs CHISE は実用的な文書編集環境の側面と実験環境の側面を持つ。これは次の

図 1: XEmacs CHISE



4つの側面に分けて考えることができる：

1. XEmacs Mule 互換
2. UCS/Unicode 実装
3. 多字化実装
4. Chaon 実装、CHISE client

1 番目の側面は XEmacs で利用可能なさまざまな Emacs Lisp プログラムが利用可能であることを意味し、実際、Wanderlust や Mew などの mail/news reader, w3 や Emacs-w3m などの WWW browser, Lookup などの辞書検索プログラム、AUC-TeX や野鳥などの TEX 編集環境、PSGML などの SGML/XML 編集環境などさまざまな用途に応じたパッケージが利用可能である。特に、文書編集の点では今も最強の環境のひとつではないかと思う。さまざまなタブ補完やダイナミック・マクロ^[2]、インクリメンタル・サーチ、置換、スペル・チェッカー、フォント・ロック^[3]やキーボード・マクロなどの強力な基本機能がある。拡張パッケージを含めた全機能は1つの環境に統合されており、例えば、キーボード・マクロで各種機能を利用することができる。これが単なるエディタではなく統合環境^[4]といわれる所以である。こうしたことから、Emacs Lisp で書かれたパッケージを使うと、Windows 上の辞書検索プログラムや今昔文字鏡といった単体のユーティリティを使うのとは異次元の感覚が得られるのである。そして、XEmacs CHISE はこうした高度な操作性をそっくりそのまま継承しているのである。

2 番目の側面では、XEmacs CHISE は BMP (基本多言語面) の文字のみならず拡張漢字 B などにも利用可能な高度な UCS/Unicode 実装と見做すことができる。GNU Emacs や XEmacs Mule でも Mule-UCS を使って UTF-8 などの Unicode 系 CES (Emacs 用語では“coding-system”^[1]) を利用可能であるが、これらの内部では JIS X 0208 や GB 2312 や CNS 11643 といった ISO/IEC 2022 系 CCS (Emacs 用語では“charset”^[1]) に写像して文字を表現するため、これらの CCS

で表現可能でない文字は扱うことができない。XEmacs CHISE 自身も UCS/Unicode に依存した実装とはいえないが、4 番目の側面によって UCS/Unicode を含めたさまざまな文字を容易に定義可能である。そして、実際に UCS/Unicode に関するさまざまな文字定義を内蔵している。

3 番目の側面では、XEmacs CHISE は 2³⁰ 個、即ち、約 10 億個の文字オブジェクトを同時に表現可能な広大な文字空間と、さまざまなフォント・エンコーディングを利用可能な表示能力を備える。これにより Unicode フォントで GT の文字を表示したり、台湾中央研究院の CDP 外字や『漢字庫』などの外字用フォントに取り敢えずの文字表現とフォント・エンコーディングを与え、すぐに利用可能にすることができる。これも実は 4 番目の側面に起因した性質である。

そして、この 4 番目の側面により、外字データを段階的に UCS/Unicode などの標準コードに変換することができる。例えば、守岡が開発や運用に関わっている東洋学文献類目データベース^[5]では大型計算機で作成された JEF 漢字+独自の外字 (JEF-CHINA3 外字) のレガシーデータと新たに入力された UTF-8 のデータをシームレスに扱う必要があるが、JEF-CHINA3 外字から UCS/Unicode への完全な写像表は現在存在しない。このため、JEF-CHINA3 外字に形式的な文字表現を与えると同時に、CHISE の文字データベースに JEF-CHINA3 外字に関する情報を追加して行くことで、徐々に UCS/Unicode に変換可能な文字を増やしている。この際、元データを UTF-8 化する必要はなく、1 番目の側面を継承した Mule の coding-system 機能によって、自在に CES を変換することが可能である。また、XEmacs CHISE が拡張した、実体参照自動変換機能によって、UCS/Unicode に変換不可能な文字をファイル中では SGML の実体参照にしておき、変換可能になった時点で XEmacs CHISE に自動的に変換させることも可能である。

4 番目の側面は現在の所、直接一般利用者の目に触れるものではないが、XEmacs CHISE の持つ文字処理の柔軟性を支える基盤となっている。そ

して、この基盤は文字オブジェクトの定義、文字オブジェクトの素性値の参照・設定・消去、文字素性に対するイタレータ、そして、文字データベース管理用の処理といった非常に単純な機能によって実現されている。しかし、これらのプリミティブと Emacs Lisp の提供する強力な機能を組み合わせることにより、さまざまな処理が可能となる。

■ CHISE 文字データベース

Chaon モデルに基づく文字処理システムを活用するためには文字知識のデータベース化が不可欠であり、我々はこうした文字データベース・コンテンツの開発にも力を入れている。現在、CHISE 環境用に提供されている文字データベースには

1. XEmacs CHISE 附属のデータベース
2. CHISE 漢字構造情報データベース (CHISE-IDS)

の 2 つがある。

前者は XEmacs CHISE に附属する文字データベースで、XEmacs CHISE の `define-char` 関数を用いた Emacs Lisp プログラム (`define-char` 形式) で表現されている。

後者は ISO/IEC 10646-1:2000^[5] で定義された IDS (Ideographic Description Sequence) を用いて漢字構造情報を表現したデータベースと XEmacs CHISE 用関連プログラムからなるパッケージである。

この漢字構造情報というのは漢字の部品の組合せ構造に関する情報のことである。ご存知のように、多くの漢字は偏と旁などの部品の組み合わせによって構成されているが、こうした漢字の部品の組合せ構造は形の抽象的表現となるだけでなく、字義や音価にも関係しており、字源に基づく文字構造の分析は「解字」と呼ばれ、そうしたデータは重要な辞書記述の 1 つである。そこで我々は、CHISE 文字データベースに収録されている全ての複合 (会意・形声) 漢字に対し漢字構造情報を付けることを目標に、漢字構造情報データベ

スの開発を計画した。そして、2001 年度には未踏ソフトウェア創造事業の予算を得て、漢字の部品の組合せ構造を表現した CHISE 漢字構造情報データベース (CHISE-IDS) の開発を始めた。現在、ISO/IEC 10646-1:2000^[5] 基本統合漢字 (Unicode の例示字形)、同 統合漢字拡張 A、および ISO/IEC 10646-2^[6] 統合漢字拡張 B に対する入力を一応完了しており、これを元に JIS X 0208:1990 (の例示字形)、大漢和文字などに対する編集作業も行っている。

CHISE-IDS の開発以前から存在する漢字構造情報を含んだデータベースとしては、台湾中央研究院の CDP データベースと台湾の中華電子佛典協會 (CBETA) の外字データベースがある。また、この他、日本でも「今昔文字鏡」がある。CDP のデータベースと CBETA の外字データベースはそれぞれ独自形式を採っており、そのままでは IDS 形式に変換することはできない。また、今昔文字鏡のデータは一般には公開されていない。「今昔文字鏡」は専有的 (proprietary) ソフトウェアであり GPL によって配布されている XEmacs CHISE で利用することはできない。一方、CDP のデータベースと CBETA の外字データベースは GPL で配布されるプログラムで利用可能であるので、可能な限り変換して利用することにした。

XEmacs CHISE 附属のデータベースと CHISE-IDS パッケージは今のところ別々に配布されているが、XEmacs CHISE が利用可能な環境では、CHISE-IDS パッケージ附属のインストーラを使って、CHISE-IDS データベースを XEmacs CHISE 附属のデータベースに統合することができる。また、CHISE-IDS パッケージは漢字構造情報を XEmacs CHISE で利用するためのプログラムを含んでいる。これは、現在の所、

- `ids.el` IDS 形式の構文解析器など
- `ids-read.el` CHISE-IDS データベースを XEmacs CHISE に読み込むプログラム
- `ids-dump.el` XEmacs CHISE 内部で用いている ideographic-structure 形式のデータを CHISE-IDS データベース形式で書き出す

プログラム

- `ids-util.el` 漢字構造情報を Unicode 例示字体風や大漢和辞典風に変換するプログラム
- `ids-find.el` 部品によって漢字を検索するためのプログラム

からなる。

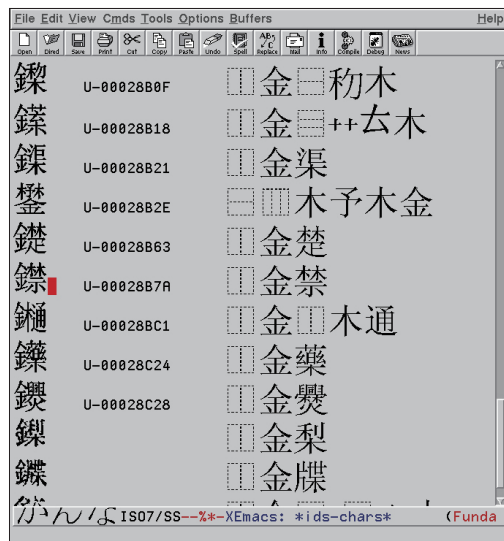
ids-find.el は 現 在 ids-find-chars-including-components (別名 ideographic-structure-search-chars) と ids-find-chars-covered-by-components という 2 つの検索用コマンドを提供している。前者は M-x ideographic-structure-search-chars [CR] 部品列 [CR] で指定された部品列の各部品を少なくとも 1 個は含んでいる漢字の一覧を表示するコマンドである (図 2)。

後者は M-x ids-find-chars-covered-by-components [CR] 部品列 [CR] で指定された部品列中の任意の部品を 0 回以上用いて全体として 2 個以上の部品から構成される漢字の一覧を表示するコマンドである (図 3)。

■ Ruby/CHISE

Ruby/CHISE は江渡によって実現された Ruby^[8]

図 2: M-x ideographic-structure-search-chars [CR] 木金 [CR] の出力結果



における Chaon 実装である。CHISE の文字データベースを利用する機能を中核とし、漢字構造情報を扱うための機能や文字データベースの可視化のため作成されたいくつかの関連ライブラリが存在する。

- Character class

Ruby/CHISE は Chaon モデルにおける文字オブジェクトを Character クラスのオブジェクトとして扱っている。この点では、オブジェクト指向言語ではない Emacs Lisp に基づく XEmacs CHISE における文字オブジェクトよりも Chaon モデル的な言語インターフェースを与えていると考えられる。

以下に、Ruby/CHISE におけるプログラム例を示す：

```
require 'chise'

include CHISE

p "字"# "字"

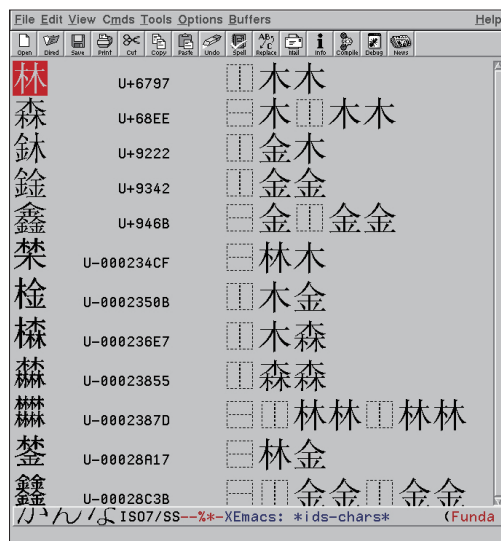
p "字".ucs# 23383

p "字".total_strokes# 6

p "字".inspect_all# いろいろでてくる

p "字".ideograph_daikanwa # 6942 (大漢和番号)
```

図 3: M-x ids-find-chars-covered-by-components
[CR] 木金 [CR] の出力結果



つまり文字そのものが自分の持つ素性を知っていて、そこにメソッドとしてアクセスできる。

上記の例における"字"は実際にはStringオブジェクトであるが、Ruby/CHISEは1文字のみからなる文字列を自動的にCharacterオブジェクトに変換する。なお、

```
p "字字".ucs # エラー
```

のように1文字ではない文字列に対して文字素性を求めようとする場合にはエラーが発生する。

また、Character class のオブジェクトを生成し、それを直接用いることもできる。

```
# Character class のインスタンスを生成
char = "字".char
```

```
# 上記とまったく同じ
char = Character.get("字")
```

Ruby/CHISEではflyweightパターンを用いているので、同じ字の場合は同じインスタンスとなる。

■ 実体参照

```
# 数値参照化する
p "字".to_er # "&#x5b57;"
```

```
# 数値参照を復号
p "&#x5b57;".de_er # "字"
```

```
# JIS X 0208-1990 による実体参照を復号
p "&J90-3B7A;".de_er
```

```
# 大漢和番号による実体参照を復号
p "&M-06942;".de_er
```

このように、実体参照の復号処理が可能となっている。

■ String の拡張

```
String#each_character
```

Stringの一文字毎にイテレータを実行。引数としてCharacterが入る

```
String#map_character
```

同様にmapした結果の文字列を返す

```
String#char_length
```

UTF-8的な文字列の長さを返す。

他にもto_er, de_erなどCharacterと共通の要素を各文字毎に摘要するためのメソッドがいくつか用意されている。

■ 素性の定義

```
"木".mydepth = 1
"林".mydepth = 2
"森".mydepth = 3
```

実行すると自動的にデータベースに保存されるため、プログラム終了後も素性は保存される。

```
p "木".mydepth
p "林".mydepth
p "森".mydepth
```

このようにして、自分独自の必要な素性を作り、その素性に基づいたプログラムを書くことができる。

■ 漢字構造情報関連

Ruby/CHISEには、CHISE文字データベースを扱うための一般的な機能に加え、漢字構造情報(字形分解・合成)を扱うための機能が強化されている。

解字

Stringにはdecompose, decompose_allという

2つのメソッドがある。decompose は一段階だけ分解する。decompose_all はそれを再帰的に行う。

```
p "字".decompose
p "字".decompose_all
p "柵".decompose
p "柵".decompose_all
p "終了".decompose
p "終了".decompose_all
p "鬱".decompose
p "鬱".decompose_all
```

部品による文字探索

IDS 形式で表現された漢字構造情報に対応する文字オブジェクトを生成するには compose メソッドを用いる。

```
p "#x2ff0; 木木".compose
```

find メソッドを用いることである部品が漢字の一部として含まれている漢字群を探すことができる。

```
p "日雲".find
```

ここでは、日と雲という部品が含まれている全漢字を出力する。

■ Ruby/CHISE の今後

これまで述べてきたように Ruby/CHISE は CHISE 文字データベースを扱うための機能を一通り備えており、これらは Ruby が提供するオブジェクト指向の枠組を用いて Chaon モデルを自然に実装しているといえる。

しかしながら、現在の所 CES 関連の機能は不足している。とりあえず、Shift_JIS と EUC-JP の相互変換だけは対応しているものの、XEmacs CHISE が提供するような各種 CES の変換はまだ実装されていない。これは、将来、外部ライブラリーと連携をとり対応する予定となっている。

■ Perl/CHISE

Perl/CHISE は師によって実現された Perl における Chaon 実装である。XEmacs CHISE や Ruby/CHISE と同様に、CHISE 文字データベース（現状では XEmacs CHISE 用の Berkeley DB 版文字データベース）を Perl から参照することが可能である他、文字オブジェクトの集合としての側面に焦点を当てた処理の開発が進んでいる。また現在は、CHISE 文字データベースを用いた、音や意味でもパターンマッチが可能な正規表現の拡張に取り組んでいる。

■ メソッド

Perl/CHISE は現在の所以下のようなメソッドを備える：

```
new( 属性名 => 値 )
```

属性名とその値に対応する文字オブジェクトを生成する。関連する他の属性も自動的に取り込まれる。

```
define_char( 属性名 => 値 [, ...] )
```

属性名とその値 (の集合) を持つ文字オブジェクトを新しく生成する。

```
dump_attr
```

文字オブジェクトが持つ属性の名前と値を出力する。

```
add_attr( 属性名 => 値 )
```

文字オブジェクトに属性名とその値を追加する。

```
del_attr( 属性名 )
```

文字オブジェクトから属性を削除する。

compare(オブジェクト名)

文字オブジェクトの属性どうしを比較して、その重なり具合を返す。

例

```
use CHISE;

my $s1 = CHISE->define_char(strokes
=> 12, radical => 9);
my $s2 = CHISE->define_char(strokes =>
12, radical => 9, ideograph_daikanwa
=> 694);
my $c = $s1->compare($s2);

if ($c == $CHISE::EXCLUSIVE) {
    print "排他的\n";
} elsif ($c == $CHISE::HAVE_INTERSEC
TION) {
    print "共通部分あり\n";
} elsif ($c == $CHISE::PROPER_SUBSET)
{
    print "完全部分集合\n";
} e l s i f ( $ c = =
$CHISE::PROPER_SUPERSET) {
    print "完全上位集合\n";
} elsif ($c == $CHISE::EQSET) {
    print "完全一致\n";
}
```

libchise

XEmacs CHISE や Ruby/CHISE, Perl/CHISE と
いった Chaon 実装は Chaon モデルに基づいて文
字を処理するという点と同一の CHISE 文字デー
タベースにアクセスするという点で極めて似通っ
た部分を持っており、実際に同様な処理を行って
いる。しかしながら、こうした部分の部品化がな
されていないために、各言語処理系はそれぞ
れ別々に実装する必要があった。これは実装効率
とデータ形式の一元管理の点で問題が多いので、

こうした部分のライブラリ化を行うことになった。
そうして誕生したのが libchise である。

libchise の開発は CHISE 環境で共有される機能
をライブラリ化することを目標としている。これ
は言い替えれば XEmacs CHISE の機能を徐々に
libchise 側に移して行くことだといえるかもしれ
ない。この際、問題となるのが複雑なデータの扱
いである。現在の CHISE 文字データベースでは
Lisp を前提とした複雑なデータ構造が使われて
いる部分があるが、こうしたものを Lisp 以外の
環境で扱うのは容易ではない。もちろん、そうし
たものを libchise の中で隠蔽し、libchise を利用
するプログラムに提供することは可能であるが、
記憶管理や実装の手間を考慮すればなるべく複雑
なデータ構造を使わないことが望ましいといえる。
そこで、libchise を実現する上で、

1. 脱 Lisp 化
2. スケーラビリティ

を目標とした。前者は libchise において必要以上
に複雑なデータ構造を扱わないことによって記
憶管理を単純化しようということであり、現在
XEmacs CHISE に附属している CHISE 文字デー
タベースの構造にも関わる問題である。

後者の問題を考えるために表 1 に示すような
階層モデルを考えている。

現状の libchise は第 2 層の基本部分の実装され
ており、当面は第 2 層までの実装を目標として
いる。

表 1: CHISE 階層モデル

第 4 層	オブジェクト層 文字に対する抽象的なサービスの提供 (関数、message 送信式、TopicMaps 等)
第 3 層	構造データ層 複雑なデータ構造の実現 (S 式、XML 等) (記憶管理や型システムの実現)
第 2 層	文字層 素性名による文字データへのアクセスの実現 (C の文字列レベルに限定、記憶管理無し)
第 1 層	データソース層 (Berkeley DB や PostgreSQL などのデー タベース・システム等)

KAGE

XEmacs CHISE を始めとする Chaon 実装は自由に文字を定義可能な環境を提供しているが、画面上に文字を表示するにはなんらかのグリフデータが必要である。現在、XEmacs CHISE では文字表示に通常のフォントを利用しているので、フォントがない限り表示可能な文字を定義することができない。また、このことはCHISE 文字データベースで定義された文字の性質と画面上に表示されるフォントのグリフや字形の間に齟齬が生じ得ることも意味している。この問題を解決するためには、文字に関する知識とグリフ・字形を統一的に扱うシステムの実現と、文字に関する知識に基づいてグリフ・字形データを合成するシステムの実現が不可欠である。

そこで、CHISE Project では上地宏一氏が開発した KAGE システムを採用し、これを CHISE 文字データベースと関係させることによりこの問題を解決することにした。なお、KAGE システムの詳細に関しては、前号および今号の上地氏の記事をご参照頂きたい。

Ω /CHISE

Ω /CHISE は CHISE 環境で定義された各種文字を印刷可能な組版システムを目指したもので、苦米地と宮崎によって開発が進められている。

Ω /CHISE はその名の通り組版エンジンとして Ω^⑦ を採用している。これは Ω が TEX より継承した精密な組版能力に加え、入出力文字コードを変換する強力な言語 OTP (= Ω Translation Process) を内蔵していることから、多言語・多文字種の文書を処理するにあたって現時点で最適なツールと考えられるからである。さらに、OTP は単に文字コードの変換だけではなく、組版規則に従った入出力の変換にも利用可能である。また、外部のプログラム（標準入出力フィルタ）をプラグイン（外部 OTP）として付加し、極めて複雑な処理を行わせることも容易である。このような

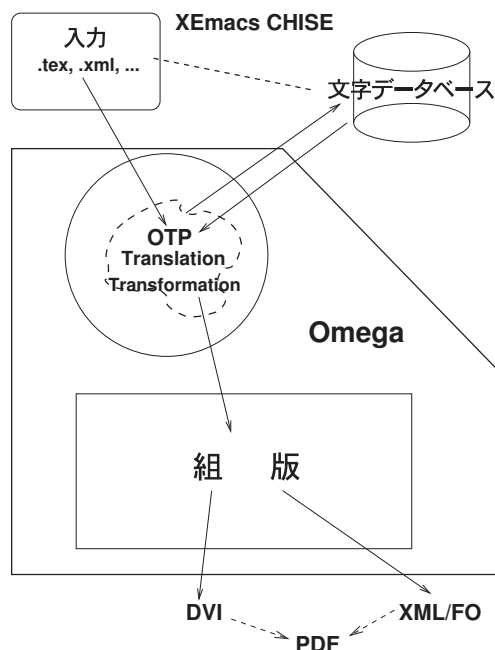


図 4: Ω を利用した高品位組版システム

Ω の拡張性は、文字データベースと連携した組版に不可欠なものであり、符号化文字集合・文字オブジェクト情報の両者を柔軟に組合せた実用的で強力な組版システムの構築を可能にすると考えられる。

また、KAGE サーバーを用いたオンデマンドなフォント生成機能を有している。

現在の所、KAGE サーバーの処理能力やグリフ・字形データの網羅性の点から、全ての漢字字形を KAGE サーバーに頼るのは現実的ではないため、指定した言語環境に対応する基本文字セットに対しては対応するフォントを利用している。

また、それ以外にも GT フォントも利用可能になっている。そうした漢字字形がない場合には KAGE サーバーに対して SVG 形式のグリフ・データを要求し、それを処理することで Postscript 用フォントを生成している。また、グリフ・データはキャッシュされる。ただ、このようなことから、現時点では XEmacs CHISE で表示可能な全文字をそのまま印刷できるという段階にはないといえる。

可視化

文字に関するさまざまな知識を編集するためのユーザー・インターフェースを実現するための第一歩として、江渡と藤原義久氏と鈴木泰博氏は漢字に関する知識の視覚化システムを研究・開発している。

これまでに行ったものとしては、(a) 藤原氏が中心になって行った漢字ネットワーク構造の解析^[2]、(b) 江渡が中心になって行った漢字ネットワーク構造の可視化システム、(c) 藤原氏が行った甲骨文字部品を用いた複合現実感の実験がある。

(a), (b) は漢字間の関係からなるネットワークの構造を解析し、その結果の可視化しようとするものである。現在の所、漢字は大雑把にいて形・音・義の3要素からなるといえるが、網羅的なデータが揃っているのは漢字構造情報（形）であるので、とりあえず共通する漢字部品を持っているという関係に着目したデータ解析を行っている。視覚化には Graphviz^[3] を用いているが、これは接続関係が増えてくると視覚化が非常に難しくなる。

(c) は紙に書かれた甲骨文字を画像認識し、複数の甲骨文字を並べた時に、それらを部品として結合した漢字をコンピューターで表示するものである。これも画像認識の制約から、現在の所、非常に少数の部品にしか対応できない。しかし、人文情報学の未来の姿を垣間見せてくれる実験だったと思われる。

画像マークアップ

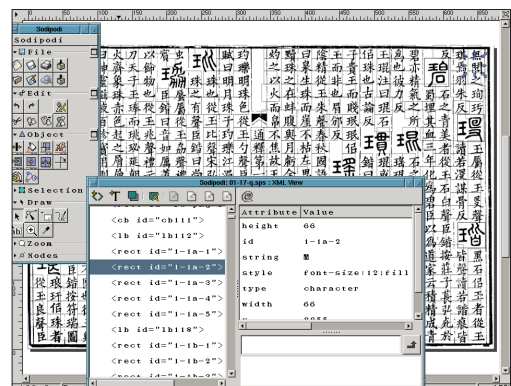
これまで述べてきたように、符号化文字モデルを置き換えるものとして考案した2つのモデルの内、文字の性質に関する知識を処理することに焦点を当てた Chaon モデルに基づく文字処理技術は CHISE Project によって徐々に具現化されてきた。一方、実際に画像などの符号化文字列とはかけ離れたもので記号処理を行うことに焦点を当てた “My Symbolic System” 的アプローチを具体化するものとして、守岡は画像マークアップ技

術を提案している^[5]。これは、XML に基づいて文書の論理構造と視覚的構造の双方を表現可能な電子テキストを実現する試みであり、視覚的構造の記述に SVG (Scalable Vector Graphics)^[10]、文書の論理構造の記述に TEI^[9] を用いることで文書の視覚的構造と論理構造の双方を表現可能な多面的マークアップ・テキストを実現しようとしている。そして、これは “My Symbolic System” 的アプローチにおける文書データ構造の叩き台となるものでもある。

同種のものとして、PDF に基づくフォント埋め込み技術を用いて異体字形を表示する手法^[12]や SVG を用いた検索可能な文字画像^[14]という手法が提案されており、プレゼンテーションや文書公開のための手法としての有効性や可能性が示されている。しかし、文書のデータベース化という観点からいえば文書の持つ視覚的構造を記述するというのがより本質的な問題であると考えられる。

我々が提案する文書の視覚的構造のマークアップ手法はこうした問題を解決するものである。これは、文書内容の論理構造と同様に文書の視覚的構造を想定し、その論理構造をマークアップする方法である。例えば、画像情報のプリミティブとしてピクセル・データやベクター・データを捉え、その上部構造として文字を捉え、文字の上部構造として行を捉え、行の上部構造として欄や柱を捉え、これらの上部構造として頁を捉え、頁の上部構造として本を捉えるといった視覚的構造に関する階層構造が考えられる。そして、文字や行や

図 5: SVG による画像マークアップの例



頁といった視覚的構造上の要素と文字や文や章といった論理構造上の要素との対応関係を記述すれば、多面的な文書記述が可能になると考えられる。

そして、将来的には画像マークアップ技術と CHISE 文字データベースやグリフ・字形関連機能などの CHISE が提供する機能が連携することにより、版本や石刻などを画像的側面も含めて Chaon 的手法で分析したり、さまざまな時代のさまざまな文字・字体・字形を再利用したりと、いろいろな可能性が膨らんでいくのではないかと考えている。

■ おわりに

CHISE Project のこれまでのあゆみと現状、そして、今後の展望について紹介した。CHISE Project は現在の豊富な計算資源を背景に文字に関するさまざまな知識をデータベース化し、それを参照することによって文字やテキストを処理するシステムを作ることによって、文字やテキストを多面的に操作する文書処理環境の実現を目指している。

こうした可能性を現実のものとするためには、プログラム・データ双方のさらなる整備が必要であり、多言語化技術、国際化技術、データベース、分散化技術など各方面のハッカーや、言語学、文字学、文献学など各種言語や字書やテキストなどを扱う人文科学者など、多方面に渡る研究者の幅広い協力が不可欠である。

こうしたことを鑑み、CHISE プロジェクトはオープン・ソースで開発されており、その情報は

- <http://cvs.m17n.org/chise/>
- <http://kanji.zinbun.kyoto-u.ac.jp/projects/chise/>
- <http://mousai.as.wakwak.ne.jp/projects/chise/>

で公開されている。これらの WWW 頁群を含め、各種成果物は CVS で管理されており、最新の開発状況を知ることができる。日本語用と英語

用の 2 つのメーリングリストも用意されており、参加方法は上述の WWW 頁で説明されている。CHISE プロジェクトに興味を持たれた方は、是非、お気軽に御参加願いたい。

参考文献

- [1] bit 別冊「インターネット時代の文字コード」, 第 9 章「文書編集系における文字コード」, 共立出版, 2001.
- [2] Yoshi Fujiwara, Yasuhiro Suzuki, and Tomohiko Morioka. Network of words. In Artificial Life and Robotics 2002, 2002.
- [3] Graphviz – open source graph drawing software. <http://www.research.att.com/sw/tools/graphviz/>.
- [4] International Organization for Standardization (ISO). Information technology — SGML Applications Topic Maps, January 2000. ISO/IEC 13250:2000.
- [5] International Organization for Standardization (ISO). Information technology — Universal Multiple-Octet Coded Character Set (UCS) – Part 1: Architecture and Basic Multilingual Plane (BMP), March 2000. ISO/IEC 10646-1:2000.
- [6] International Organization for Standardization (ISO). Information technology — Universal Multiple-Octet Coded Character Set (UCS) – Part 2: Supplementary Planes, November 2001. ISO/IEC 10646-2:2001.
- [7] The omega typesetting and document processing system. <http://omega.cse.unsw.edu.au:8080/>.
- [8] The object-oriented scripting language Ruby. <http://www.ruby-lang.org/>.
- [9] C. M. Sperberg-McQueen and Lou Burnard, editors. TEI P4: Guidelines for Electronic Text Encoding and Interchange — XML-compatible edition. University of Oxford, 2002.
- [10] The World Wide Web Consortium (W3C). Scalable Vector Graphics (SVG) 1.0 Specification, September 2001. <http://www.w3.org/TR/SVG/>.
- [11] XEmacs. <http://www.xemacs.org/>.

- ⑫ 安岡孝一. 外字と異体字. アジア情報学のフロンティア 全国文献・情報センター人文社会学学術セミナーシリーズ No.10, 全国文献・情報センター人文社会学学術セミナーシリーズ, 第 10 巻, pp. 25-34, 2000.
- ⑬ 守岡知彦. UTF-2000 汎用文字符号に依存しない文字表現系の展望. アジア情報学のフロンティア 全国文献・情報センター人文社会学学術セミナーシリーズ No.10, 全国文献・情報センター人文社会学学術セミナーシリーズ, 第 10 巻, pp. 13-24, 2000.
- ⑭ 守岡知彦. ポスト文字コード時代の文書処理技術に関する展望. データベースの活用と人文社会学 全国文献・情報センター人文社会学学術セミナーシリーズ No.12, 全国文献・情報センター人文社会学学術セミナーシリーズ, 第 12 巻, pp. 59-70, 2002.
- ⑮ 守岡知彦. 文字画像のマークアップの試み. 東洋学へのコンピューター利用第 14 回研究セミナー, 京都大学学術情報メディアセンター第 71 回研究セミナー, pp. 2130, Mar 2003.
- ⑯ 師茂樹. CHISE プロジェクト. 漢字文献情報処理研究, Vol. 3, pp. 143-147, 2002.
- ⑰ 東洋学文献類目 (第 4.6 版) <http://kanji.zinbun.kyoto-u.ac.jp/db/CHINA3/>

注

- [1] どちらも、g新部裕氏の命名である。
- [2] 編集中の文書に含まれる用例を用いた補完
- [3] 文書構造に基づいて自動的に書体や色を変える機能
- [4] しばしば、“kitchen sink”と揶揄される。

文字コード外フォント処理

上地 宏一（かみち こういち）

■ 多漢字環境とフォント処理

ISO/IEC 10646 (Unicode) の普及と拡張によって PC 上における多漢字処理環境が整備されつつあるが、表示や印刷処理に必要となるフォントの整備は今一步といったところである。こうした状況の中で、ここ一年の間に公開された文字コード外のフォント処理技術を紹介する。

■ 漢字構形資料庫

台湾中央研究院 資訊科学研究所 文献処理実験室では、「漢字構形資料庫」と名づけた漢字データベースを 2003 年 3 月より Web 上で公開している^[1]。ファイルサイズは ZIP 圧縮で 50MB 程度と大きく、もともと CD-ROM で配布されていたようである。データベースには、以下の漢字が収録されている。

- ・『漢語大字典』および Big5 コードに収録される 54,711 字
- ・『説文解字詁林』の小篆および重文の 11,100 字形

これらの収録字に対する部首、注音、字典掲載場所、構字式、異体字リンクなどのデータを持ち、専用の閲覧ツール^[2]（図 1）を利用して情報を検索する。

■ データベース機能

閲覧ツールは、漢字または構字式を入力して検索された候補の中から目的の字を選択し、さまざまな情報を閲覧する。

構字式とは、漢字の構造を「部品とその位置関係」または「部品とその筆記順」で表す式のことである。中央研究院独自の構字記号（表 1）を用

図 1：データベース閲覧ツールで「荊」を表示

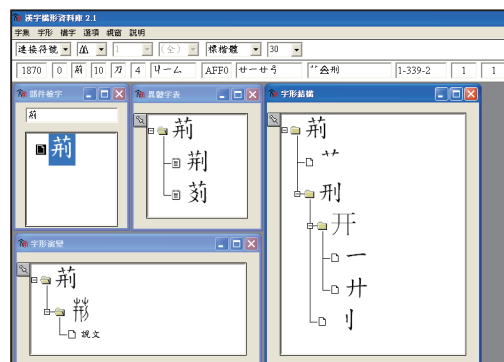


表 1：構字記号の一覧（説明書より抜粋）

類別	符號	説明	構字式範例
連接	△	當部件的連接順序由左至右	順 = 川△頁
	△	當部件的連接順序由上至下	含 = 今△口
	△	當部件的連接順序由外至內	圍 = □△韋
部件序	☐	按部件書寫順序輸入，前後以起始符號（☐）和終止符號（☐）包夾	牖 = ☐片戶甫☐
	☐		
方便（符號）	8	三個相同部件直連	炎 = 8 火
	8	三個相同部件直連	
	∞	二個相同部件橫連	朋 = ∞ 月
	∞	三個相同部件橫連	
	∞	三個相同部件呈三角狀排列	焱 = ∞ 火
	∞	四個相同部件橫連	
	∞	四個相同部件直連	
	∞	四個相同部件呈四角狀排列	𤇓 = ∞ 火

いて、収録されるすべての漢字に対して、それぞれ別々の構字式が与えられている。

■ フォントも収録

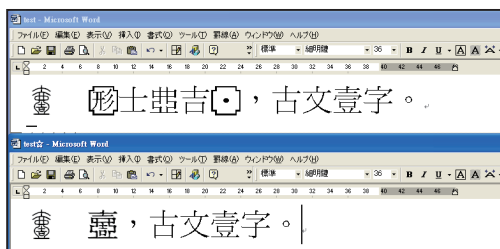
Big5 に収録されない字形は、すべて外字フォント形式で呼び出され、7つの外字フォントを切り替えることで4万字を処理する。このためフォントデータだけで80MB程度の容量となっている。フォントは、明朝体と楷書体、そして『説文解字詁林』の篆書体も用意されている。明朝体と楷書体はそれぞれ「細明體 (MingLiU)」と「標楷體 (KaiU)」をベースとしていて、『漢語大字典』の字形と完全に一致しないグリフもある（たとえば草冠は、明朝体は3画、楷書体は4画に統一されている）。

惜しむらくは、一部の文字は明朝体・楷書体の共にグリフが用意されず、その代替として印刷された字形（『漢語大字典』のものと思われる）をスキャンして輪郭抽出したものが含まれている。おそらくフォント製作が完成していないものと思われる。

■ Microsoft Office 製品での利用

収録される文字はすべて Word などの Microsoft Office 2000/XP 製品で利用できる。具体的には、構字式を外字識別文字列としてドキュメント内に記述し、メニューから変換機能を実行すると、データベースを元に構字式を外字に置き換える機能を持つ（図2）。また、外字ではなく埋め込み画像にも変換できる。フォントの切り替えによってコード外字を表示するため、フォント設定を変更すると別の字に化けてしまうことである。このため実際の利用では、ドキュメント完成

図2: 構字式を含む元データ（上）と変換後（下）



後に、一括で構字式を外字字形に変換する方法が望ましい。

■ 今後への期待

『漢語大字典』収録字に関しては、ユニフィケーションされた字を含めてすべて ISO/IEC 10646-2 (Extension B) までに収録されているため、「Simsun (Founder Extended)」などのフォントを利用することによって、現在は外字処理を用いなくともデータの処理および字形の表示が可能である。データベースには『漢語大字典』の掲載位置情報が含まれているため、UCS コードとの関連付けも可能であろう。そこで、次のバージョンでは外字領域ではなく Ext. B 領域への変換に対応してほしい。また、データベースからフォント部分を独立させ、Ext. B まで収録した繁体字系デザインフォントとして無料で配布してほしいと思うのは欲張りだろうか。

■ 易符動態組字編輯器

台湾の易符智慧科技 (eForth) では、先に紹介した文献処理実験室の研究成果をもとに、文字コード外字をその場でフォント (グリフ) 生成してドキュメント内で利用することが可能な簡易エディタと、作成した文書を Internet Explorer で表示することのできるツールバーを2003年4月より公開している^[3]。

■ efeditor

efeditor はプレインテキストエディタであるが、加えて以下の機能を持っている。

- 漢字検索機能
- グリフデザイン調整機能

漢字検索機能は、Unicode の CJK 統合漢字、Ext. A および Ext. B までの漢字とハングルなどの合わせて83,232コードポイントを、文字に含まれる部品から検索する。またすべてのコードポイントに対してフォント (グリフ) が用意されている。

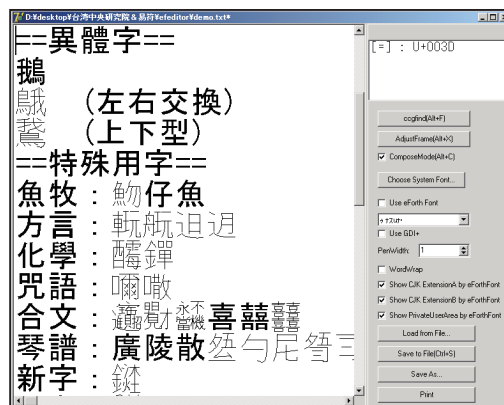


図3: エディタ画面 (細字が生成されたグリフ)

エディタ内部では、アウトライン形式やビットマップ形式ではなく筆画の骨格情報で漢字部品を表現し、後述する智慧型編碼によって各グリフをデータ化している。このため、エディタで表示するグリフはその場で生成される。現在は、骨格情報から明朝体への肉付けのアルゴリズムは無く、単純に骨格線の描画によってグリフを生成する。

■ 智慧型編碼

コード外文字は「智慧型編碼」と名づけられた独自形式で表現されている。

[㊦ : 開始記号]

[部品字・番号] [左上座標] [右下座標]

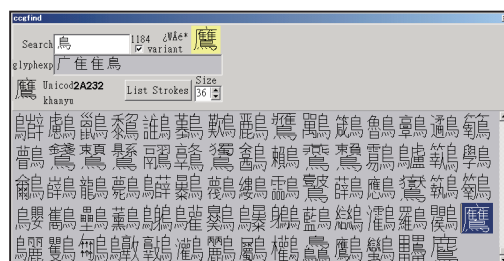
(部品の数だけ繰り返す)

[㊦ : 終了記号]

鵝→㊦我 00061080DE 鳥 00871072DF㊦

つまり、部品の種類と、その位置を与えることで、グリフを表現 (生成) する、パラメータ付与

図4: 文字検索で「鳥」を含む漢字を検索



型グリフ生成法^[4]を利用している。表現内に部品の位置情報まで入れてしまうことには多少違和感があるが、実装が容易なため有効である。

■ 新しい字の利用

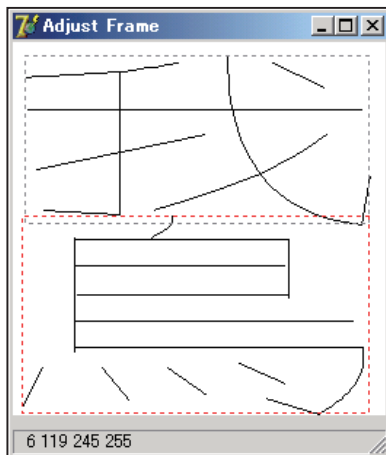
efeditor では、文字コードに収録されない文字をユーザが自由に作ることが可能である。「智慧型編碼」の形式で含まれる部品を記述し、グリフデザイン調整機能で見た目を整える。調整機能はそれぞれの部品の大きさと位置を変更することが可能である。新たに作字した字だけでなく、CJK統合漢字、Ext. A、Ext. Bのコードポイントのグリフデザインも調整が可能である。ただし調整を行った段階で、その字の内部データは Unicode のコードポイントではなく「智慧型編碼」による表現に変換されてしまう。調整は個別の字に別々に適応されるため、同じ文字であっても別々の場所にある文字のデザインを一括で調整することはできない。

このほか、削筆記号や削部品記号が用意されており、漢字部品の任意の筆画や部品を削除することで、漢字の好きな部分だけを取り出して新しい字を作ることが可能となっている。

■ Internet Explorer 用ツールバー

efeditor で作成したドキュメントを他のアプリケーションで表示すると「智慧型編碼」の部分は、

図5: デザイン調整画面



グリフではなく構造表現に分解されてしまう。そこで、専用のツールバー利用すると、efeditor で作成した HTML 文書を、そのままの形で表示することが可能となる。残念ながら、内部のプログラムによってグリフはビットマップ画像に変換されるため、印刷には適していない。

ツールバーは日本語版 Windows でも動作した^[5]。多少動作が不安定なのが残念である。

■ 実際に利用してみても

efeditor はテキストエディタとしてはいろいろと改善の余地があるが、コード外文字を気軽に利用できるという点で、非常に斬新である。今後のバージョンアップに期待したい。

■ KAGE システム

昨年著者が本誌第 3 号にて紹介した KAGE システム^[6]は誠に不本意ながら公開を停止している。これは主にサーバを設置する環境を提供できなくなったことと、サーバプログラムの安定性に問題が残っているためである。そこで現在は、プログラムを一から見直している。具体的には、

- gcc (C 言語) でプログラムを書き直す
- 不安定要素となるサーバ部分を実装しない
- 単体のプログラムとしてユーザが自分の手で利用することもできる

といった改良を行っている。さらにプログラムソースや部品データを GPL ライセンスで公開し、自由に応用してもらうことを考えている。本誌発刊時期には、機能限定版を公開するスケジュールである。

注

- [1] <http://www.sinica.edu.tw/~cdp/>
入手にはライセンスへの同意と、利用者情報の登録が必要である。
- [2] 説明書では動作要件として Microsoft Windows 95/98/Me/2000/XP の繁体字版としているが、Windows 95 繁体字版では動作せず、また Windows 2000/XP の日本語版 (内部ロケールを台湾 / 香港に変更) でも動作しなかった。
- [3] <http://www.eforth.com.tw/efeditor/>
- [4] パラメータを部品の特徴から自動的に計算する「フォント (グリフ) 自動生成」と厳密には異なる。
- [5] インストールするファイルを日本語名 (たとえばデスクトップなど) のフォルダに置いた場合、インストールに失敗する。
- [6] <http://fonts.jp/index.cgi?KAGE>

補記：本稿で紹介している技術の選定は「紅眠の愛の秘密日記」<http://www.din.or.jp/~khoming/nikki/> を参考にした。

Embedding Glyph Identifiers in XML Documents

Christian Wittern (クリスティアン・ウィッテルン)

訳：師 茂樹

■ はじめに—— EGIX とは何か

2002 年 12 月 20 日、川俣晶氏を筆頭とする日本人の技術者グループが「Embedding Glyph Identifiers in XML Documents」(通称 EGIX) と題する技術文書を W3C に提出した。この文書は元々、日本規格協会の標準情報 (TR) である JIS TR X 0047:2001^[1] として発表されていたものであり、その一部はシフト JIS の外字領域を透過的に利用する方針について述べた Windows NT 漢字処理技術協議会 (XKP) の『XKP GAIJI 交換仕様書』等に基づいている。

この文書には、共著者として村田真氏や小町祐史氏などの著名人が名を連ねており、その内容を端的に言えば「<http://www.xml.gr.jp/xmlns/PRE/Reference>」という名前空間と、グリフを参照するときに用いる name 属性を定義したものである。すなわち、要素に内容があれば、それを name 属性で示されるグリフで表される別の字形 (ほとんどの場合正字) として解釈するようにする。name 属性の値は、ISO 10036:1996, Cor.1:2001 *Information technology — Font information interchange — Procedures for registration of font-related identifiers*^[2] に則り、登録機関 (Registration Authority) によって割り当てられ

たグリフ識別子でなければならない。また、この文書では、附録として 3 つの参考例が挙げられている。EGIX がどのような利用を想定して制定されたかを理解するために有用なので、以下にそのすべてを挙げておこう^[3]。

A.1 Example 1

```
<html xmlns="http://www.w3.org/1999/
xhtml">
<body xmlns:glyph="http://
www.xml.gr.jp/PRE/Reference">
<p><span glyph:name="ISO/IEC
10036/RA//Glyphs:10003290"
>吉</span>田茂</p>
</body>
</html>
```

A.2 Example 2

```
<html xmlns="http://www.w3.org/1999/
xhtml">
<body xmlns:glyph="http://
www.xml.gr.jp/PRE/Reference">
<p><span glyph:name="ISO/IEC
10036/RA//Glyphs:10003290">吉
(The version of Short Upper Line)</
span>田茂</p>
```

```
</body>
</html>
```

A.3 Example 3

```
<html xmlns="http://www.w3.org/1999/
xhtml">
<body xmlns:glyph="http://
www.xml.gr.jp/PRE/Reference">
<p>田茂</p>
</body>
</html>
```

以下、本稿では、グリフ参照交換のために提案されたこの規格についてレビューしたいと思う。その際、この規格が問題解決のために適切なかどうか、あるいはその有効性、完全性、総合的な有用性などについて、他の提案との比較しつつ検討したい。

■ 符号化されないグリフ（文字）をどうするか

EGIX は、コンピュータや情報技術の利用が拡大するにともなって、より多くの文字を使いたい、あるいは実際ある文字に対してグリフをきめ細かく選択したいという要求が増え続けている現状を踏まえ、それを解決しようとして提案されたものである。漢字に基づく書記系(writing system)^[*]には長い歴史があり、また時代や地域によっても様々な正書の規範があることから、これまで生み出されたグリフには非常に多彩な字体、字形が存在する。ある複数の文字が、あるコンテキストでは置き換え可能な文字として使われる一方、別のコンテキストではそれが許されないといったことが起こるため、事態は複雑化している。逆に、情報処理用の符号化文字の開発過程に目を向けると、

開発が始まった当初はコンピュータのリソースがまだ貧弱だったこともあり、符号化する文字の数をなるべく少なくする努力がなされてきた。

中国において長い伝統のある文字学の分野では、古典籍電子化の試みや、特に符号化・標準化の試みが、重要かつ新しい刺激となった。漢籍の情報処理のための文字の符号化は 1970 年代に始まるが、それ以来、文字の標準化についての理論や用途もまた目覚ましい発展を遂げてきたと言ってよい。その結果として、情報伝達の抽象的な実体である文字そのものと、それが図形的表現を伴って様々な具体化したもの（インスタンス）を区別する重要性が認識されるに至った。具体化においては、字形や大きさ、画数、部品などが異なってくるにもかかわらずそれらは同一の抽象的な文字を参照しているのである。情報交換自体や様々な情報処理を目的とする場合、視覚的な具体化については意味がないものとして無視することができる。一方、特殊なグリフが用いられる固有名を扱う場合など、視覚的表現が情報として不可欠である場面も存在する。

1980 年代後半になると、現在使われている言語のすべての文字を、単一の文字レパートリーとして符号化してしまおうという取り組みが始まった。そこでは、たとえ意味的に異なった漢字であっても、複数文字として符号化しなくてもよいと判断された場合には、それらを統合することが決定された。漢字統合のためのルールが一通りできあがると、1990 年代はじめには 20,000 字以上の漢字を収録した最初の漢字集合がリリースされた。この時、グリフをより明確に参照できるようにしたいという要求があることも認識されていた。例えば、組版などにおいてフォントの中から適切なグリフを選択したい、というような場合である。この点を考慮して、ISO 9541:1991 (Information technology -- Font information Interchange -- Part 1. Architecture) と前述の ISO 10036 に則ったグリフの個別の登録が始められ、グリフのレジストリ^[*]の管理は Association of Font Information Interchange (AFII) が行うとされた。

しかしながら、もっと多くの文字が欲しいとい

う要求は収まることがなく、さらなる文字の符号化が始められることになった。追加文字の新しいリリースがなされるたびに（本稿執筆時点で、7万字を超える漢字が符号化され、さらに30,000字以上が審議中である）、漢字統合のルールは厳密さを欠くようになり、それにもなってこれまで符号化されなかったグリフまで符号化されるようになってしまった。また、当然のように、文字符号化に対する代替策と見なされていたレジストリへのグリフの登録作業は、ほとんど進まなかった。UCS (Universal Character Set) に非常にたくさんの文字を収録することが認められたため、AFII に対してすべきことはあまり残っていなかったのである。結果的に AFII の代表は、この業務の運営を担当する委員会に対して、規格の取り下げを要請した^[4]。

ここで指摘しておきたいのは、符号化文字集合の領域をさらに拡張しようとしている現在の風潮は、上述のような問題への解決策にはなりえないということである。なぜなら、文字集合とは開いた集合であり、視覚的に異なる字形を情報処理的に区別したいという要求はコンテキストに依存するものであって、両者は明らかに異なったレイヤーに属するからである。

■ これまでに提案された解決策

上に述べたように、抽象的な文字と具体的な形をもつグリフとのあいだには、重要な概念上の差異がある。情報処理においては、この二つを別のレベルにおくようにできることが決定的に重要となってくる。マークアップ言語を使うことができるのであれば、より豊かな表現力を持ち、グリフの微妙な違いを有意に区別することができる、マークアップのレイヤーが追加されることになる。本レビューで紹介している EGIX はまさに、マークアップを用いた解決策を提案しているのであり、重要な成果だと言えよう。これ以外にも同様の解決策を提案しているものがあるが、これらと比較することは有意義ではないかと思う。ここで取り上げようと思うのは、次のものである。

- SVG: altGlyph(<http://www.w3.org/Graphics/SVG>)
- MathML: mglyph (<http://www.w3.org/Math>)
- Rick Jelliffe (1999): Elements for Non-Unicode Characters in XML (http://www.ascc.net/~ricko/xcs/missing_chars.html)
- TEI: Writing System Declaration (<http://www.tei-c.org/Activities/CE>)

これらはすべて、未だ符号化されていないまったく新しいグリフについて参照するためのマークアップ方法と、既存の文字に対して特殊なグリフを指定する方法のどちらか、もしくは両方を備えている。最初の2つがマークアップ要素によって図形的な実体を記述する方法であるのに対して、後の2つは新しい要素を導入するだけでなく、文字に対して追加情報を記述するためのマークアップ構文も提供している。これらの方法と比較すると、EGIX はグリフを参照するために属性をたった1つしか用いないという点で、特異な存在だと言える。SVG や MathML と同様、EGIX には既存の文字やグリフに対して別の情報を追加する機能はない。しかし、EGIX 文書にある使用例に〔吉 (The version of Short Upper Line) のような〕普通の言葉で書かれたコメントが書かれていることからわかるように、追加情報を記述しなければならない場合もあるという意識を EGIX のエディタたちが持っていたことが読み取れる^[5]。また、EGIX が上にあげた方法と比べて大きく異なっている主な点は、グリフへの参照が規定 (normative) とされていることであろう。これは EGIX の肝心なところである。レジストリに事前に登録されたグリフについてのみ参照できるようにすることで、符号化されていない文字の交換が確実に安定したものになるための基盤を提供しているのである。

情報交換というのは、同一の情報項目への参照が理解され、その理解に基づいて処理されるということが前提になっている。もし、1つのグリフをここでとりあげている5つの方法を使って別々に参照したとすると、この5つが同じもの

を参照していることを処理システムが理解することはまったく期待できないだろう。換言すれば、情報交換という目的のために使うのであれば、世界共通の参照システムがない限り実際問題としてうまく機能しないのである。

Rick Jelliffe 氏と（現在作成中ではあるが）TEI が提案する方法では、世界共通のグリフ識別子を前提とすることはない。その代わりに、対象となるグリフと追加情報とを結びつけるためのマークアップ構文を利用する。これによって、これらの方法に従って記述されたファイルをシステムが受け取った場合、要求された処理のコンテキスト（例えば表示用に文字を描画しているのか、検索用のインデックスを作っているのか、など）に応じてグリフを処理することができる。標準化されたグリフとの結びつきを表現する場合、EGIX ではできたとしても暗黙的にならざるをえないが、TEI では明示的に可能である。Unicode などの文字符号化標準では、符号化した文字を一意に同定するためのコードポイントに加えて、多くの文字プロパティを提供している^[6]。適切な処理を確実にものとするためには、これまで符号化されたことがない文字やグリフについても、プロパティが要求されたらそれに答えられるようになってなければならない。私が、グリフ蓄積センターのようなものに本質的な欠陥があると考えているように思われるのは、本意ではない。しかし、処理に必要なプロパティにアクセスすることができるような方法を構築しなければならないのである。まれにしか使われないグリフの場合には、その出典がどこにあるかという情報も強く求められる。

グリフ・レジストリについての議論が EGIX の根幹をなしていることをふまえ、<http://www.glocom.ac.jp/iso10036/> のウェブサイト（最終更新日 2001 年 5 月 1 日）で誰でも見られるようになっているレジストリの現状について、いささか述べておきたいと思う。そこで表になっているグリフの合計は、2003 年 7 月 31 日現在、以下の通りである。

表 1 ISO 10036 の登録機関が割り当てたグリフ^[7]

英数字	7,901
漢字	21,204
ハングル	11,506
上記以外	81,743
合計	122,354

AFII は 2 年半の間にたった 1 つのグリフの登録申請を受けたただけだったが、GLOCOM はほぼ同じ期間で 81,743 というとてもない数のグリフの登録を受け付けており、それは上記の公開ウェブページで確認できる。

これは確かに重要な成果であるが、問題も残されている。まず、このウェブサイトでは、登録されたグリフを探す方法が登録番号を除いて何も提供されていない。登録番号の割り当て方についてどこにも説明がないため、深く考えずに見た者にはでたらめな順番で並んでいるようにしか見えないだろう。また、もし誰かが果敢にもここにあるすべてのページの中から使いたいと思っているグリフを探そうとしても、更なる問題がある。番号 103590 のグリフ（<http://media.glocom.ac.jp/kmmr/10036/glyph-table.html?cjk&103>）と、番号 10003234 のグリフ（http://media.glocom.ac.jp/kmmr/10036/glyph_id_03.html）とは、前者が少し太いぐらいで私の目にはほとんど同じとしか見えない。この 2 つは Unicode の U+53E5（句）にもそっくりである。情報交換という点から考えれば、これは望ましくないことである。すなわち、ある同じ情報単位に対して複数の表現が存在した場合、グリフへの参照情報のみに依存したシステムにおいては、このような状況を知ることが不可能だからだ。さらに言えば、フォント関連オブジェクトの登録機関が、サードパーティーが提供する巨大なグリフの集合を、すでに登録されたグリフのデータベースとの整合性をチェックせずにそのまま収録してしまっていたとすれば^[8]、このような重複は大きな数に上るだろう。これは大いにありうることである。上にあげた例は適当にピックアップしたものであるが、実際には恐らく数千の、場合によっては数万のこうしたグリフの重複が登録簿の中にあるのではないだろうか。

■ まとめ

以上の分析から、EGIX という提案は次の 2 点において貧弱な基盤上にあるのではないと思われる。すなわち、意図する意味の伝達手段としてのマークアップの使い方と、グリフを参照するときの方法についてである。EGIX ではある特定のグリフ・レジストリへの参照しか認めていないが、それは不幸なことではないかと思う。また、参照の際には URI ではなく Formal Public Identifiers (FPI) を用いることになっているが、World Wide Web の技術仕様を考えた場合少し奇異な気がする。もし EGIX が URI による参照を認め、他の W3C 勧告（例えば RDF や HTML 用のモジュールなど）に従ったマークアップ構文を採用したならば、利用価値はかなりあがるのではないだろうか。加えて、この提案が東アジア言語のテキスト処理における異体字問題を特に念頭においたものであることを踏まえると、Unicode がすでに漢字の構造情報を伝達するために所謂 Ideographic Description Sequences (IDS) を規定していることについても指摘しておかなければならないだろう。EGIX であげられている例で言うと、あの文字は□土口と説明されたほうがより利用価値が高くなるのではないだろうか。CHISE プロジェクト (<http://www.kanji.zinbun.kyoto-u.ac.jp/projects/chise/> 参照) では、このような識別子がグリフデータベースへアクセスするための糸口になるのと同時に、その文字がどこで符号化されているのかなど、豊富なデータベースを提供しており、現在ではこのように参照されたグリフをその場で合成してしまう仕組みも用意されている。

■ 免責事項

筆者は、文字符号化に関する TEI のワークグループの議長を務めている。また、筆者は CHISE プロジェクトのメンバーでもある。

■ 訳者謝辞

日本語訳においては、Christian Wittern 氏をはじめ、川幡太一氏、守岡知彦氏のコメントを頂いた。感謝申し上げる。但し、誤訳等はすべて訳者の責任である。

注

- [1] http://www.y-adagio.com/public/standards/tr_lsi_xml/lsi_xml.htm 参照。邦題は「XML による画像参照交換方式」で、英語に直訳すると “Exchange of image references with XML” で、「Picture Reference Exchange by XML」という英語の副題がつく。
- [2] この規格の本文は <http://www.glocom.ac.jp/iso10036/docs/main.htm> [訳注：日本語訳である JIS X 4165:2002 は http://www.y-adagio.com/public/standards/jis_fntobj_reg/toc.htm] からオンラインで入手可能。
- [3] ここに 3 つの例をそのまま引用したが、正直に言えば、これらの例が十分な検討を経て選ばれたものとは思えない。ちなみにこの例に出てくる吉田茂とは、1946 年から 1954 年の間、日本の総理大臣を務めた人物である。
- [4] 1999 年 8 月 19 日の *Final AFII Liaison Statement Concerning ISO/IEC 10036* (SC34N92) を参照 (<http://www.y12.doe.gov/sgml/sc34/document/0092.htm>)。
- [5] マークアップ言語で書かれているテキストに普通の言葉でコメントを書こうとする場合、それが明らかにテキスト本文ではなくメタ・レイヤーに属し、かつそれがそのテキストの電子化に関するコメントであるならば、そのためのマークアップ構文を用意した方がより役に立つだろう。上に引いた Example 2 には、「Example 1 と同じであるが、人間が読むための情報を含んでいる。人間が読むためのコメントが挿入されている。検索プロセッサはこの span 要素を無視することができるので、このコメントは検索の対象にはならない」という前置きがある。しかし、検索プロセッサが現在、こういう状況に対してどのような動作をするかを理解することは難しい。マークアップをする目的は、むしろこのようなヒントを持たせるということではなから

うか。もし、この Example をそこらにあるウェブブラウザで表示させたら、このコメントは無視されず、他のテキストと並んで表示されることであろう。

- [6] 詳細は *The Unicode Standard*, Chapter 4 を参照されたい。また、Unicode Character Database はオンラインでも閲覧できる (<http://www.unicode.org/ucd/>)。
- [7] GLOCOM のウェブページには、このなかの一部が AFII から引き継いだグリフ画像であると述べる箇所がある。しかし、この中のどれが相当するのかについてははっきりとしない。思うにこれは、最初の3つのカテゴリ〔英数字・漢字・ハングル〕がそれに相当し、GLOCOM に移ってからはカテゴリわけをすることを断念したことを意味しているのではないだろうか。繰り返すが、このことについての説明は一般公開されたウェブサイトには見当たらない。
- [8] GLOCOM のウェブサイトではこの件について何も述べてられていないが、冒頭にあげた Example 3 が文字

鏡研究会のウェブサイトにある GIF 画像を参照しているうえ、[GLOCOM の登録番号と文字鏡の] 2つの番号のあいだに機械的な対応関係があることから、事情を知らない者でもそれと知られる。

訳注

- [*] ISO/IEC 10036:1996, Cor.1:2001 の日本語訳である JIS X 4165:2002 では、writing system、registry の訳語として「表記システム」、「記録」が提案されている (http://www.y-adagio.com/public/standards/jis_fntobj_reg/tutr.htm)。しかしここでは、翻訳の都合上それぞれ「書記系 (writing system)」「レジストリ」とした。

附記

本論文の原文（英語）は、紙幅の都合上、Web 上に掲載している。<http://www.jaet.gr.jp/jj/egix.pdf> を参照されたい

JIS X 0213 の改正と UCS との 関係について

川幡 太一（かわばた たいち）

■ 1 はじめに

日本の基幹的な文字コード規格である JIS X 0208^[1]を代替・拡張する文字コード規格として、JIS X 0213^[2]が 2000 年に制定されてから、はや 3 年が過ぎた。しかし、JIS X 0213 規格はなかなか一般への普及の兆しを見せていない。その一方で、国語審議会による表外漢字字体表^[3]の答申に伴い、JIS X 0213 には早くも規格の大幅な改正が見込まれている。

本論文では、JIS X 0213 のこれまでの実装などの動向を俯瞰し、改正が検討されるに至った経緯と、改正の内容について概略を述べる。また、特に、JIS X 0213 と UCS^[4] との関係について考察し、その問題点について述べる。

■ 2 JIS X 0213 のこれまでの動向

■ 2.1 JIS X 0213 の実装

JIS X 0213 は 2000 年に制定されたものの、その普及はあまり進んではいない。Microsoft などの大手消費者向け OS ベンダが、JIS X 0213 を自社 OS 製品に対して正式に実装しない（またはできない）ことの影響が大きいと思われる。もともとこれらのベンダの OS は、JIS X 0208 の空き領域に、独自の文字を割り当てて、内部にシフト

符合化表現を使うことで、各々の日本語文字コード体系を構築していた。そのため、JIS X 0213 で追加された文字と、これらのベンダ独自の文字は、同じコードポイントで衝突してしまっていた。さらに最近では、各社とも OS の文字コードの UCS 化を推進しており、JIS X 0213 の多くの文字が UCS へ正式に対応していないことも、これらの OS による JIS X 0213 のサポートを困難にさせていた。

アップルコンピュータ社は、これまで MacJapanese と呼ばれる独自の JIS X 0208 の拡張文字集合等を使っていたが、Mac OS X からは、標準搭載される大日本スクリーン用のヒラギノ明朝等の OpenType フォントの文字レパートリに、JIS X 0213 の文字集合を含めている。当初、APGS と呼ばれていたグリフセットは、Mac OS X 10.2.3 から、正式に Adobe の集合 “Adobe-Japan1-5”^[5] となった^[6]。

Linux や FreeBSD 関連では、幾つかのテキストエディタやツールにおいて、JIS X 0213 をサポートしているが（詳細は [7] 等に詳しい）、これらは実験的な色合いが強い。テキストデータにおいては、青空文庫等において Shift_JISX0213 の実験的に採用されたことがある^[9]。

■ 2.2 JIS X 0213 の訂正

JIS X 0213 は規格票の出版当時にいくつかのバグがあり、それらのバグに対して、訂正表が

3 JIS X 0213 の改正

■ 3.1 表外漢字字体表の答申と、それに基づく JIS X 0213 の改正の計画

この表外漢字字体表における 1,022 文字は、JIS X 0213 で使われているのと同じ、平成明朝体で示されている。そこで、JIS X 0213 で示されている文字の字体を、この表外漢字字体表の漢字と比較すると、815 文字は同じ字体であるが、残りの文字については、デザイン差が微細なものも含めて、字形に違いが存在する。そのため、表外漢字と JIS 漢字の齟齬を早急に解消し、表外漢字への対応に関するベンダ間での混乱を避けることを目的に、JIS X 0213 の漢字の字体を、表外

[illegible]

改正予定の図形（JCS のホームページ^[10] より引用）

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
2	◆	◀	□	△	▲	▽	▼	※	干	一	↑	↓	≡	レ	ノ	ハ	ヘ	マ	ミ	メ	モ	ム	ミ	メ
3	▶	▷	◀	◀	▲	▽	▼	干	一	↑	↓	≡	レ	ノ	ハ	ヘ	マ	ミ	メ	モ	ム	ミ	メ	メ
4	あ	い	う	え	お	か	き	く	け	こ	さ	し												
5	ア	イ	ウ	エ	オ	カ	キ	ク	ケ	コ	サ	シ												
6	A	B	I	E	Z	H	K	A	M	N	E	G	P	E	T	G	O	F	X					
7	A	B	B	I	J	E	F	J	H	K	L	M	O	P	P	C	T	Y	F	X				
8	一	二	三	四	五	六	七	八	九	十	十一	十二	十三	十四	十五	十六	十七	十八	十九	二十	二十一	二十二	二十三	
9	€	£	¥	₹	₪	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	₱	
10	A	~	L	L	S	S	S	Z	Z	Z	A	I	S	S	S	S	S	S	S	S	S	S	S	
11	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	
12	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	
13	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰	⑱	⑲	⑳	㉑	㉒	㉓	
14	俱	丈	岳	一	手	半	丁	乞	份	仿	伋	你	佑	佐	佖	仉	佬	佻	伢	伥	伧	伧	伧	
15	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	
16	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	咩	
17	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	院	
18	押	旺	横	戊	王	界	冠	突	界	突	界	突	界	突	界	突	界	突	界	突	界	突	界	
19	院	旺	横	戊	王	界	冠	突	界	突	界	突	界	突	界	突	界	突	界	突	界	突	界	
20	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	病	
21	機	備	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
22	供	供	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
23	抽	抽	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
24	后	后	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
25	此	此	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
26	察	察	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
27	次	次	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
28	宗	宗	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
29	勝	勝	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
30	次	次	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	
31	次	次	疾	氣	汽	藏	祈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	屈	

漢字字体表に倣う形に変更することになった^[10]。

この改正では、大雑把に言えば、旧 JIS X 0213 にある新字体の表外漢字の多くを、UCS と矛盾がおきない形で字体表に合わせて康熙字典体（旧字体）に変更すると同時に、各文字の UCS との対応関係を最新の UCS 規格に基づいて規定し、規格全体の整合性を保つために、包摂基準や charset 名の変更なども併せて行うものである。

■ 3.2 人名漢字表の拡大の動きについて

しかしながら、この改正作業が行われる途中に、ある事件が起った。それは、2003 年 1 月の森山法相（当時）による、突然とも言える人名漢字の大幅拡充の表明である^[11]。ここで一番大きな問題となるのは、『「表」外漢字』の「表」には人名漢字表が含まれており、もし人名漢字が増えるならば、それに従って表外漢字表の対象となる漢字が減る可能性があることである。表外漢字表の漢字は、康熙字典体（旧字体）を主にしてきたが、一方の人名漢字表の漢字は通常、新字体の方を採用してきた。もし JIS X 0213 を表外漢字表に従って改正した後で、人名漢字表の拡充と表外漢字表からの新人名漢字の削除が行われ、かつ、新人名漢字と旧表外漢字で漢字の字体が変更されていれば、泥縄的に再び JIS X 0213 の改正が迫られ、混乱が起きることになる。

このような問題に直面し、JIS X 0213 の改正の審議会では、新しい人名漢字表が決まるまで改正を控えるべきか、または敢えてこれを踏まえた上で、表外漢字のコンピュータでの使用における混乱を避けるために、規格を早急に改正すべきかで、議論があった模様である。

2003 年の 3 月の公開レビュー審議を最後に、JIS X 0213 の改正に関する動向は外部に公表されていない。この問題に対する改正委員会の見解は統一されず、JIS X 0213 の改正作業が滞っている可能性がある。

■ 3.3 JIS X 0213 改正に伴う charset 名の変更について

JIS X 0213 の改正に伴い、MIME などを使わ

れる charset 名も従来の“Shift_JISX0213”や“EUC-JISX0213”等から変更されることになっている。EUC, SJIS, JIS はそれぞれ、encoding に、対象となる文字集合の制定・改正年の情報だけを組み入れた、“EUC-JIS-2003”、“Shift_JIS-2003”、“ISO-2022-JP-2003”となる予定である。

■ 3.4 JIS X 0213 改正に伴う新しいエスケープ文字について

改正される JIS X 0213 の 1 面の文字集合は、数多くの漢字のデザインの変更および包摂基準の変更が行われるため、従来の JIS X 0213 の 1 面とは非互換であるという認識のもと、あらためてエスケープ・シーケンス用の終端文字を取得することになっている。現在では、新たな 94 × 94 の文字集合を登録しようという動きが他に見られないことから、新しい終端文字は 5/1 になるであろうと予測されている。

これによって、JIS X 0208 が、4/0（1978 年）と 4/2（1983 年）の 2 つの異なる終端文字を持つと同様に、JIS X 0213 の 1 面も、4/15（2000 年）と、5/1（2003 年？）の 2 つを持つことになる。

ここで、（JIS X 0208 の 1990 年改正のように）更新文字を取得しないで終端文字を取得するのは、改正される JIS X 0213 は従来の JIS X 0213 と上位互換でないという認識しているためである。表外漢字の扱い方の安定を図るためには、当面の非互換性による混乱を受容するのは、やむをえないということであろう。

■ 3.5 JIS X 0213 の全ての文字に対する UCS 対応コードの付与について

2000 年の JIS X 0213 の制定当時は、JIS X 0213 で新たに追加された多くの文字が、UCS に存在しなかった。そのため、当時出版された規格書では、これらの文字の対応 UCS コードとして、当時の JIS 委員会が ISO/IEC JTC1 に提案したコード番号を、正式でないという意味で括弧つきで規格書に掲載していた。

その後、CJK Ideograph Extension B の制定や、ISO 10646-1 の改正を通して、JIS X 0213 の全

ての文字に対して、対応 UCS コードが決まった。今回の改正では全ての JIS X 0213 の文字に対して、対応 UCS コードが明記されることになった。しかし、一部の文字では、UCS との相互変換において、文字の解釈に関する問題、およびラウンドトリップの問題が見られる。これについては、次節で細かく見ていきたい。

■ 4 JIS X 0213 と UCS の対応について

Microsoft や Apple が自社 OS の文字コード基盤として UCS を推進し、またそうでない機器ベンダでも JIS X 0208 の未定義領域に独自のコードを割り当てている場合が多い現状では、残念ながら JIS X 0213 の付属書の符号化表現がそのまま使われることはあまりないだろう。むしろ、UCS 文字集合の日本語サブセットの文字レパートリとしての、JIS X 0213 が注目される可能性が高い。すなわち、今回の JIS X 0213 の改正において、ベンダ等にとって最も関心が高いのは、JIS X 0213 と UCS の対応ではないだろうか。

しかしながら、JIS と UCS の間には、文字の解釈に相違があったり、ラウンドトリップ変換が保証されない等の問題がある（従来からのベンダ間での UCS・JIS X 0208 対応表の違いによる問題は議論しない）。

■ 4.1 JIS と UCS における声調記号の対応

Unicode3.0 規格書の 7.8 節の冒頭で説明 (Modifier letters are ... used to indicate modifications of a preceding letter) しているように UCS では声調記号は、その前に現れる文字に応じて変形を行うような文字^[12]であるが、JIS X 0213 にある声調記号では、変形に関する規定が一切ない。

そのため、JIS X 0213 側では、1-11-64～68 の文字を並べても、そのままの文字の形として出力することになる。その意味で、1-11-64～1-11-68 は、厳密には U+02E5～U+02E8 には対応するとは必ずしも言えない。

さらに、このような対応では、例えば JIS X 0213 の、“1-11-64 1-11-68”の文字列は、UCS には“U+02E9 U+02E5”と変換され、これを再び JIS X 0213 に戻すと、“1-11-70”となり、ラウンドトリップが保たれない。

この問題を解決するには、声調記号の挙動を、UCS 側と合わせるよう改めて規定を JIS X 0213 に加えるか、または、JIS X 0213 の声調記号は合成を行わないことを明記し、UCS との対応に際して、対応 UCS 文字の前または後ろに、U+200C のような合成禁止文字を付加する必要があるだろう。

表 4.1.1 UCS 変換のラウンドトリップが維持されない声調記号

JIS コード	UCS コード	別の JIS コード
1-11-69 (ハ… 声調記号上昇調)	U+02E9(ハ) U+02E5(ハ)	1-11-68 1-11-64
1-11-70 (ハ… 声調記号下降調)	U+02E5(ハ) U+02E9(ハ)	1-11-64 1-11-68

■ 4.2 IPA の UCS 対応問題

JIS X 0213 と UCS の間では、IPA(発音記号)を表現する文字と従来のギリシャ文字等との区別が一部で異なる ([2] の解説表 4)。例えば、UCS では発音記号の“ε”とギリシャ文字の“ε”を区別しているが、JIS X 0213 では同じ文字と見なしている。そのため、JIS X 0213 で符号化した発音記号を UCS に変換しても、必ずしも同じ文字として解釈されるとは限らない。一部の教科書採用のアクセントつき母音も同様に、UCS との対応に疑問が残る。

また、JIS X 0213 では、一部の IPA の上つき文字や小さい文字については、その原型となる文字に組版指定で変形することを意図している。

一方、UCS ではこれらの文字に対して、すべて個別に文字コードを割り当てている。

そのため、発音表記に関しては、UCS を JIS X 0213 に変換したり、またはその逆を行う際に、文字コード以外のレイヤーを考慮する必要がでてくるだろう。

表 4.2.1 JIS X 0213 と UCS における IPA の対応
の相違

発音記号	JIS	JIS 対応 UCS	UCS
有声両唇ふるえ音	1-3-34	U+0042(B)	U+0299(ɸ)
無声両唇摩擦音	1-6-21	U+03A6(Φ)	U+0278(Φ)
有声軟口蓋摩擦音	1-6-35	U+03B3(γ)	U+0263(ɣ)
有声軟口蓋側面接近音	1-3-44	U+004C(L)	U+029f(L)
有声口蓋垂破裂音	1-3-39	U+0047(G)	U+0262(g)
有声口蓋垂鼻音	1-3-46	U+004E(N)	U+0274(n)
有声口蓋垂ふるえ音	1-3-50	U+0052(R)	U+0280(r)
有声口蓋垂内破裂音	1-11-9	U+0193(Ġ)	U+029b(ġ)
非円唇前舌広めの狭母音	1-3-41	U+0049(I)	U+026a(i)
円唇円舌広めの狭母音	1-3-57	U+0059(Y)	U+028f(y)
非円唇前舌広・中段母音	1-6-37	U+03B5(ε)	U+025b(e)
円唇前舌広母音	1-3-11	U+0152(E)	U+0276(æ)
無声口頭蓋摩擦音	1-3-40	U+0048(H)	U+029c(h)
有気	1-3-72	U+0068(h)	U+02b0(ɦ)
唇音化	1-3-87	U+0077(w)	U+02b7(ʷ)
硬口蓋化	1-3-74	U+006a(j)	U+02b2(ɸ)
軟口蓋化	1-6-35	U+03b3(γ)	U+02e0(ɣ)
側音破裂	1-3-76	U+006c(l)	U+02e1(ɭ)
非円唇舌広母音高	1-11-38	U+1F70(ɑ)	U+0251(a) U+0300(̀)
非円唇舌広母音低	1-11-39	U+1F71(α)	U+0251(a) U+0301(́)
非円唇前舌広・中段母音高	1-11-48	U+1F72(è)	U+025b(e) U+0300(̀)
非円唇前舌広・中段母音低	1-11-49	U+1F73(é)	U+025b(e) U+0301(́)

また、JIS X 0213 では、日本の教科書等で使われるアクセント付母音発音記号に対して、個別に文字図形と符号を定めている。これらは、UCS では、合成文字を使って対応することになるが、その合成文字自身が、JIS X 0213 に単独で存在する。すなわち、JIS X 0213 と UCS の対応は一意に決まらず、ラウンドトリップが保てなくなる可能性がある。

表 4.2.2 JIS X 0213 の複数の符号が同一の UCS
に対応する例

JIS コード	UCS コード	別の JIS コード
1-11-36(æ)	U+00E6(æ) U+0300(̀)	1-9-60 1-11-60
1-11-46(ɔ̃)	U+0254(ɔ̃) U+0300(̀)	1-11-34 1-11-60
1-11-47(ɔ̣)	U+0254(ɔ̣) U+0301(́)	1-11-34 1-11-58
1-11-48(λ)	U+028C(λ) U+0300(̀)	1-11-23 1-11-60
1-11-49(λ̣)	U+028C(λ̣) U+0301(́)	1-11-23 1-11-58
1-11-50(ə̃)	U+0259(ə̃) U+0300(̀)	1-11-16 1-11-60
1-11-51(ə̣)	U+0259(ə̣) U+0301(́)	1-11-16 1-11-58
1-11-52(ə̥)	U+025A(ə̥) U+0300(̀)	1-11-35 1-11-60
1-11-53(ə̦)	U+025A(ə̦) U+0301(́)	1-11-35 1-11-58

■ 4.3 漢字の UCS 対応問題

JIS X 0213 の改正案では、表外漢字字体表に関連して、該当する JIS 漢字の字形を表外漢字字体表の字形に変更した場合、その形をした文字が UCS で別コードにすでに存在してしまうため、あえて字形を変更せずに、新たに追加する漢字が 10 文字存在する。

表 4.3.1 UCS 互換のために追加する漢字

追加漢字	対象となる漢字
俱 (1-14-1)	俱 (1-22-70)
剥 (1-15-94)	剥 (1-39-77)
叱 (1-47-52)	叱 (1-28-24)
吞 (1-47-94)	吞 (1-38-61)
噓 (1-84-7)	噓 (1-17-19)
妍 (1-94-90)	妍 (1-53-11)
屏 (1-94-91)	屏 (1-54-2)
并 (1-94-92)	并 (1-54-85)
瘦 (1-94-93)	瘦 (1-33-73)
繫 (1-94-94)	繫 (1-23-50)

しかしながら、もし、ある文字の JIS における包摂範囲と捉えられる漢字が、UCS の別コードポイントにより適した字形があることが問題に

なるのならば、以下に示すように、実はすでに Extension B の制定時に同様の問題が発生しており、それを無視して、今回を措置を行うのは、片手落ちではないだろうか^[13]。

表 4.3.2

JIS コード	UCS コード	字形的に近い漢字の UCS コード
1-38-45(突)	U+07A81(突)	U+2592E(突)
1-65-41(疊)	U+07589(疊)	U+24D01(疊)
1-86-32(欄)	U+06B04(欄)	U+237EC(欄)

5 最後に

冒頭で述べたように、JIS X 0213 の制定から 3 年が経過した今も、その普及はあまり進んでいない。その最大の原因は、やはり、一般のコンピュータ OS が、サポートしない・できないことに拠るところが大きいと思われる。

JIS X 0213 規格は、文字集合と、その排列および符号化方法から成る。しかし、その両方の側面において、一般のコンピュータ OS にとって、JIS X 0213 は採用しにくいものであった。

すなわち、各ベンダが JIS X 0208 の空き領域に独自に文字を追加してしまったため、JIS X 0213 で定めた排列は採用しにくかったし、また JIS X 0213 の文字集合中の多くの文字が、UCS には対応する文字がなかったため、UCS をベースとした OS でも JIS X 0213 は採用しにくかった。

表外漢字字体表への対応するための JIS X 0213 の改正は、このような問題を解消するチャンスとも捕えることができる。このことを踏まえて改めて、JIS X 0213 の改正案を見てみると、残念ながら今回の改正を経ても、UCS との対応において、一部の問題は解決されていない。

(今回の JIS X 0213 の改正では、基本的に表外漢字字体表に関する事項に絞って議論されており、表外漢字に関係しない問題についてはコミットしない姿勢である。)

今後の JIS X 0213 には、排列としての JIS X 0213 の文字集合の役割とともに、UCS 文字集合

の日本語サブセット規定という役割が期待される。その際、UCS と JIS の対応関係をより細かく検証することが必要になると思われる。

引用文献

- [1] 日本規格協会. JIS X 0208:1997, 7 ビット及び 8 ビットの 2 バイト情報交換用符号化拡張漢字集合.
- [2] 日本規格協会. JIS X 0213:2000, 7 ビット及び 8 ビットの 2 バイト情報交換用符号化拡張漢字集合.
<http://www.itscj.ipsj.or.jp/ISO-IR/228.pdf>
- [3] 国語審議会. 表外漢字字体表. 2000 http://www.mext.go.jp/b_menu/shingi/12/kokugo/toushin/001218.htm
- [4] ISO/IEC JTC1/SC2. *ISO/IEC 10646 Universal Multiple Octet Coded Character Set*.
- [5] Adobe. *Adobe-Japan1-5 Character Collection for CID-Keyed Fonts*. 2002 <http://partners.adobe.com/asn/developer/pdfs/tn/5146.Adobe-Japan1-5.pdf>
- [6] 大日本スクリーン. ヒラギノ OpenType と Mac OS X のバージョン相関表. 2003 http://www.screen.co.jp/ga_product/sento/support/otf_ver.html
- [7] 矢野啓介. 新 JIS 漢字実践情報. 2000-2003 <http://www.asahi-net.or.jp/~wq6k-yn/code/zissen.html>
- [8] 福田 雅史. 「JIS X 0213 正誤票」を読む. 2003 <http://hp.vector.co.jp/authors/VA000964/x0213errata.htm>
- [9] 青空文庫. 青空文庫 新 JIS 漢字時代の扉を開こう. 2001 <http://www.aozora.gr.jp/newJIS-Kanji/newJIS1.html>
- [10] 財団法人 日本規格協会 符号化文字集合調査研究委員会. JIS X 0213 改正原案の公開レビュー. <http://www.jsa.or.jp/domestic/instac/review/0213review.htm>
- [11] 小形克宏. 人名用漢字の大幅増は、JIS 文字コードをどう変えるか. 小形克宏の「文字の海、ビットの舟」文字コードが私たちに問いかけるもの <http://internet.watch.impress.co.jp/www/column/ogata/sp17.htm>
- [12] Peter Constable. *Presentations of tone contours encoded as UCS tone characters*. 2001 <http://anubis.dkuug.dk/JTC1/SC2/WG2/docs/n2312.pdf>
- [13] 川幡 太一. 新 ISO/IEC 10646 と、Unicode を検証する. 漢字文献情報処理研究 2 号, 2001

Unicode 4.0

師 茂樹 (もろ しげき)

■ メジャー・バージョンアップの意味

■ 16 ビット原則の看板をついに下ろす

2003 年 4 月、Unicode 4.0.0 (以下、各バージョンを数字のみで表す) が公開された^[1]。2000 年に 3.0 が公開されて以来のメジャー・バージョンアップである。しかし、それにふさわしい派手な変化は見られず、基本的には 3.2 を踏襲していると言ってよい。文字の追加は 1,200 字あまり (全体の文字数は 96,248^[2]) で、3.1 における 8 万字にせまる大規模な追加を目の当たりにした者とすれば、これがなぜメジャー・バージョンアップなのか? と疑問に思っても当然であろう。

私見を述べれば、今回のバージョンアップは、Unicode の元々のポリシーと現実との乖離を公式に認め、現実に進み寄る形で規格書自体が修正されたことに大きな意味があるのであろう。

修正の中で最も注目されるのは、3.0 まで「Unicode Design Principles」の筆頭に上げられていた「Sixteen-bit character code」が 4.0 において「Universality」に置き換わったことである。よく知られているように、Unicode 開発当初、世界中の文字を 16 ビット = 65,536 字以内に収めようとしたために、これまで幾度となく批判と誤解を招いてきた。実際には早くからサロゲートペアなどの存在によって崩れ去っていたのであるが、「サロゲートペアも 16 ビット 2 文字を組み合わせたものだから 16 ビットコードの原則は崩れていない」などという苦しい言い訳をしていた。今

回とうとうそれを取り下げ、一方で 16 ビット (2 バイト) の原則を脅かすものとして 3.0 までは付録扱いだった UTF-8 (1 ~ 6 バイト)、UTF-16 (2 バイト+サロゲート)、UTF-32 (4 バイト) などの CCS が規格本文に取り込まれることになった。16 ビットの看板をはずしてしまった Unicode は、Universality ——あらゆる文字を収録する唯一のレパートリーを目指して歩むことになったのである。

■ 異体字処理の夢、再び

もう一つ、本誌読者の関心が高いであろう漢字とのからみで、Unicode のポリシーの変化として指摘しておきたいのは、Unification の原則が崩れつつあるということである。この原則は、現時点でも「Unicode Design Principles」に残っているが、特に 3.1 における CJK Extension B の追加に関してその実質的变化を川幡太一氏がするどく指摘したところである^[3]。

これまで Unicode は、Unification に対する批判を避けるべく異体字処理の方法を模索し続けてきた。例えば、日本の「骨」と中国の「骨」を使い分けられないのはおかしい、というような批判に対して、言語指定用の Language Tag を定義した。結局のところこの Tag は、4.0 本文で“The use of these characters is strongly discouraged.”となっていることからわかるように、XML 等と組み合わせると無用な混乱が起こることが容易に想像されることから^[4]、早くから問題を指摘され、現在では事実上あってなきがごとき存在となっている^[5]。

ところが 3.1 ～ 3.2 の段階で、Variation Selector (3.1 では Variant Selector) という異体字表現のための文字が追加されていたことは比較的注意されていないように思う。これは当初、モンゴル文字の異体字を扱うための 3 個 (180B..180D) と、それ以外の若干の文字に対応した 16 個 (FE00..FE0F) が定義されていた。何を異体字と見なすかは Standardized Variants^[6] という文書で定義されているが、そこに漢字は現在のところ見られない。今回、4.0 においてさらに 240 個 (E0100..E01EF) が追加されたが、これは “especially for future CJK variants”^[7] と述べられていることからわかるように、漢字の異体字を強く意識したものになっている。

このような異体字表現は新しいものではないが、XML ベースの “Embedding Glyph Identifiers in XML Documents”^[8] などが提案されており、Language Tag の場合と同様の問題が起きる可能性も否定できない。また、そもそも包摂基準がなし崩しにされ、異体字がどんどん収録されている現状との関連で新たな矛盾が生じたりしないか、今後注意深く見守ってかなければならないだろう。

■ 読みやすさの向上

以上のような内容的な面もさることながら、3.1 および 3.2 における様々な変更点が 4.0 として一冊にまとまったことは大きな価値と言えるだろう。3.1 および 3.2 は、書籍として発売されていた 3.0 に対する差分として提供されていたため、非常に読みづらかった。3.2 におけるある規定を読もうとしたとき、まず 3.0 の本文にあたり、次に 3.1、3.2 の順で該当箇所に訂正や追補がないかを追いかねなければならなかった。しかも、3.1、3.2 は、項目の並べ方が異なるため該当箇所を探すことが面倒であった。実際にこの作業を行った人にはわかるだろうが、相当のストレスを感じたはずだ。

■ その他の情報

■ 『易経』『太玄経』

中国学に関連することと言えば、『易経』の六十四卦 (4DC0..4DFF) と『太玄経』の八十一首 (1D300..1D356) が 4.0 で収録されたことが注目される^[9]。開発の際には、Richard Willhelm 氏のドイツ語訳 (1924) を Cary F. Baynes 氏が英訳したもの (1950) が主に参照されたようであるが、近年の出土資料による著しい成果を踏まえ、往年の Willhelm / Baynes 訳に基づくことに対する批判も見られた^[10]。

■ Shift_JIS 索引がなくなる

これまで、「Han Indices」として「Shift_JIS Index」と「Han Radical-Stroke Index」が載っていたが、前者がなくなり、結果的に後者が独立した一章となった。JIS X 0213 との対応問題が背景にあると考えるのは邪推であろうか^[11]。

■ Unicode 4.0.1 Beta

4.0 の書籍が出る前に、4.0.1 β に対するコメント募集のアナウンスがあった (10 月 27 日 ☞ 切)。未見であるが、内容は主に Unihan.txt のアップデートで、大幅な訂正、増補がなされた模様である。

Unihan.txt は Unicode に収録された漢字について、様々な情報を記した巨大な (現在 25MB 超) テキスト・ファイルである。その中身は、

- 各国・各メーカーの文字コード規格に対応するコード値
- 各種辞書 (康熙字典、宋本廣韻、諸橋大漢和、漢語大字典、辭海、大字典源 [韓国]、香港常用字字形表、*Fenn's Chinese-English Pocket Dictionary*、*Mathews' Chinese-English Dictionary*、*A Practical Cantonese-English Dictionary*、*Student's Cantonese-English Dictionary*、

The Modern Reader's Japanese-English Character Dictionary, Ten Thousand Characters: An Analytic Dictionary など) における該当ページ、文字番号など

- ・国別・辞典別の部首・画数・総画数
- ・各種発音(日本語の音読み・訓読み、普通話・広東語発音、韓国語発音、ベトナム語発音、唐代の発音〔Hugh M. Stimson. *T'ang Poetic Vocabulary* に基づく〕)
- ・漢数字であればそれに対応する数値
- ・異体字情報

などである。漢字は伝統的に形・音・義の三要素を持つと言われており、コンピュータ処理においても字形とコード値が定義されただけでは適切な処理ができない場合が多い。Unicode は早くからコード値以外の情報をこのような形で提供することを続けてきており、現在も増加中である。

Unicode に限らず、現在の大規模漢字処理ではこうした知識をいかに扱うかがトレンドと言ってもよい。Unihan.txt においては、すべての情報を同列で扱っているため、どれとどれが中国系の情報なのか、どれとどれが歴史的な情報なのか、といった場合わけをすることは、これ自体では難しい。

また、すべての漢字にこれらの情報が付されているわけではない。例えば、最新版の Unihan.txt を見てみると、漢字の「一」(U+4E00) には 50 個あまりの情報が記されているが、Extension B 領域の漢字には一文字あたり 3、4 個しかない。利用者としてはむしろ後者にこそ、出典情報などを含む充実した情報が欲しいわけであるが、現在ではまだ未整備である。とは言え、このような巨

大なデータベースが、機械処理が容易な形で自由に利用できるようになっていることは非常に意義深いことである。今後も国際的な連携の中で文字の知識情報が適切に構築され、自由に利用できる環境が整えられることが望まれるだろう。

注

- [1] Unicode 4.0 の書籍版 (The Unicode Consortium. *The Unicode Standard, Version 4.0*. Reading, MA, Addison-Wesley, 2003. ISBN 0-321-18578-1) は 8 月下旬刊行予定。執筆時点では入手できなかったため、本稿の内容はオンライン版に基づく。
- [2] これに制御文字等 2,313、私用領域 137,468 が加わる。
- [3] 川幡太一「新 ISO/IEC 10646 と Unicode の漢字を検証する」(『漢字文献情報処理研究』2、2001)
- [4] XML における Language Tag 等の使用を禁じた "Unicode in XML and other Markup Languages" (<http://www.unicode.org/reports/tr20/>、<http://www.w3.org/TR/unicode-xml/>) の公開は 1999 年ごろまで遡りうる。
- [5] ちなみに、Unicode とほぼ同じレパートリーをもつ ISO/IEC 10646 では、このような注意書きはない。
- [6] <http://www.unicode.org/Public/UNIDATA/StandardizedVariants.html>
- [7] <http://www.unicode.org/versions/Unicode4.0.0/>
- [8] 本号 Wittern 論文参照。
- [9] Unicode 本文よりも提案時の文書の方が詳細である。
<http://www.dkuug.dk/jtc1/sc2/wg2/docs/n2363.pdf>
- [10] <http://groups.yahoo.com/group/unicode/message/1773> 3 以下を参照。
- [11] 川幡前掲論文。本号川幡論文も参照。



実践レポート

勤務する大学に CALL 教室が設置され、それまでコンピュータに親しんだことのない教員が、ある日突然 CALL 授業を任されたら — 外国語教育改革の目玉として、多くの大学で大規模な CALL 教室が設置されている今日、決してあり得ない仮定ではない。

近年、市販の中国語教材ソフトが数多く出回るようになった。教育機関向けにハードとソフトをセットにしたいいわゆる「CALL パッケージ」の中にも、多言語に対応するものが出始めている。もし、こうした市販の製品や半製品の中に、自分の授業にぴったりのものを見つけ出せれば、少なくとも当面の問題は解決できるかもしれない。もちろん「ありもの」教材の限界は踏まえた上で、どれをどうやって使うか考えてみてはどうだろうか。

日本中国語 CAI 研究会について

本会（会長：田邊鉄北海道大学助教授）は「コンピュータ援用の授業方法を中心とした中国語教授法の研究・開発・普及を推進し、同時に教員・研究者・ソフトウェア開発者の交流をはかる」ことを目的とし、1996年11月に発足した。会員による研究発表・実践報告の場として、例会（年1～2回）、総会（秋、年1回）を開催するほか、常時メーリングリストで情報意見交換を行っている。参加を希望される方は、中国語 CAI 研究会 Web <http://moli.cims.hokudai.ac.jp/~ccai/> を参照していただきたい。

Contents

市販教材ソフトは使えるか	田邊 鉄	90
Blackboard Learning System について		
バーチャルユニバーシティ構築の可能性	高橋由利子	96

市販教材ソフトは使えるか

中国語入門／初級授業での実践

田邊 鉄（たなべ てつ）

はじめに

TOEIC や TOEFL と言えば、英語の好きな人や英語を仕事にしている人のものだと思っていたが、最近はどうでもないらしい。物理学が専門の知人など「英語が苦手な理系に進んだのに」とぼやいているが、英語でのプレゼンテーションやディベート能力は、これまで英語とは無縁と考えられていたような分野で、より求められているようだ。

英会話学校に通ったり語学留学したりするほど、英語学習そのものにモチベーションを持ってないが、とりあえず英語ができなければならない。市販の英語学習ソフトの多くは、そのような人々を対象とした自習教材である。PC ショップのソフトウェア売り場には、ダイアログを中心とした会話教材や、資格試験向けの語彙集、タイピング・ゲーム、名作映画など多彩な教材ソフトが並んでいる。対象としている学習者も入門／初級から、上級者まで様々である。

英語に比べると少ないが、他の外国語教材も製品が出揃いつつある。特に中国語はこの2年ほどの間に、初・中級を対象とした教材ソフトが増えている。

現在のところ、中国語 CALL 授業は、主に教員個人が自主開発した教材で行われている。自由度が高く、教員の技量とPCの性能次第で、本当に「使える」教材が完成する可能性が高い。反面、開発を行う教員の負担が大きい。市販の「決定版」教

材があれば、それを使いたいと考える教員も多いのではないだろうか。本稿は、いくつかの市販中国語教材ソフトを、大学のごく一般的な（いわゆる「教養」の）中国語初級授業で利用する、北海道大学における取り組みについて紹介する。

利用する教材は大きくわけて2種類ある。ひとつは通常の教科書と同じように、ダイアログを中心に、語彙、語法、作文ドリルなどを取り合わせたオール・イン・ワンのコースウェアである。もうひとつは、リスニング、語彙など特定の能力養成に特化した単機能教材ソフトウェアである。これらに加え、外国語教育一般、あるいは、コンピュータ利用教育一般向け教室システムの、多言語対応についても紹介する。

北海道大学の CALL 授業現況

北海道大学では、2000年に最初のCALL教室が設置された。情報メディア教育研究総合センター（当時）が、教育用電子計算機システムの一環として企画したもので、言語教育部門が仕様策定に関与しなかったため、外国語教員の意見が十分には反映されなかった。外国語教育用ソフトウェアは、外国語教育部門でサーバごと別途用意したが、計算機システムの仕様との不整合を起こす原因になった。その後、言語教育部門と情報教育部門の間で、「次期教育用電子計算機システム更新時に、言語教育システムの見直しを行う」という一応の合意を見た。

ところがその後の調査で、市販の CALL 教材ソフトウェアは、メディア・プレイヤーやブラウザのバージョン、ライブラリなど、特定の環境を要請するものが多く、情報教育のシステムと一元化するのが難しいことがわかった。そこで、教育用電子計算機システムと切り離れた、外国語教育専用システムの導入が決められた。

2002 年 9 月、LL 教室を改装した 60 ブースの CALL 教室が設置された。2003 年度中には、さらに 60 ブース×2 教室が運用を開始する予定である。

新設の 3 教室に、従来の 1 教室を加え、4 教室が揃った段階で、(1) 英語Ⅲ（2 年次配当）の 1 単位を CALL 化、(2) 独・仏・中の 3～4 クラスでの CALL 授業実施が可能になる。

中国語は、現在 2 名の教員が合計 3 クラスの CALL 授業を展開している。いずれも 1 年生配当の初級クラスである。授業では市販教科書『話す中国語北京篇』を用いているので、今回は、希望する学生 26 名について、各種ソフトウェアを使って復習や発展学習を行ってもらうことにした。

独習用コースウェア

サーバ・クライアント環境で使えるものとして、『ハイパー中国語・実用漢語課本』^[1]がある。北海道大学では、CALL 教室設置と同時に導入している。教科書の朗読、語彙、語法解説、置き換え練習などが、全てデジタル化されており、自習用には最強の教材の 1 つと言える。ただ、準拠する教科書の内容が多すぎて、特に週 2 回しか授業のない理工系クラスでは、「消化不良」を起こす学生が多かった。ペアで担当するネイティブ教員からも「CALL 教材があるという理由だけで、この教科書を採用し続ける意味があるのか」と疑問の声が上がった。やむなく、この教材の利用は自由登録制とし、カリキュラムから外すことになった。現在のところ、中国関係の専門を志望する 2 年生や、留学希望者を中心に、40 名程度が利用登録している。動画など見た目に派手な機能はないが、自分の発音を録音し、ネイティブの発

音と聞き比べる機能のように、独習用教材に必要な十分な機能を搭載しているので、対面授業の補助というよりは、いわゆるフリー・アクセスの授業に向いているのかもしれない。

これ以外のソフトは、基本的にスタンドアロンでの（個人の）利用が前提であり、学習履歴や成績をクラス毎に管理するようなことはできない。CALL を採用して成績評価の部分で「ラクをしよう」という向きには、採用をためらわせることになる。ただ、教科書を完全にデジタル教材に置き換えるつもりがあれば、使えるかもしれない。

この種のソフトウェアは、『王老師の標準中国語シリーズ』^[2]や『北京言人』^[3]、『中国語快速完成』^[4]、『中国語完全マスター』^[5]などがある。今回は、比較的新しい『中国語快速完成』の入門篇を 14 名の学生が 3 週間試用した。

■『中国語快速完成』の出来

発音と基礎語彙を学ぶ入門篇、構文と表現を学ぶ初級篇、ダイアログで会話を学ぶ会話篇の 3 つのグレードに分かれており、おおむね、1 年生から 2 年生前期の 1 年半のカリキュラムに対応している。いずれも、音声 CD の付いた教科書『やさしい中国語快速完成』が同梱されている。2 色刷りの教科書は、イラストを多用し、大きめの文字で見やすい^[6]。

入門篇と初級篇は、北京語言学院の『口語』を思わせる置き換え練習に重点を置いている。ダイアログがない代わりに、基本表現に 10～20 の単語を当てはめた文が並んでいる。基本構文の徹底した反復練習と、語彙の増強を基本とする学

図 1 学習画面





図2 中国語学習の案内人?キャラクター

習法を想定しており、収録語彙数は入門篇だけで700を超える。日本で出版された教科書としては、トップクラスの語彙数であろう。

ソフトウェアは、アドベンチャー・ゲーム形式になっている。中国にゆかりのある主人公が、家に代々伝わる秘宝の謎を追って、中国に渡るというありがちなストーリーである。学習者はこの主人公に扮して、中国語の練習問題にチャレンジする。全問正解すると、謎解きのカギが与えられる、という具合だ。それぞれの練習問題は、教科書に完全準拠しているが、ゲームの内容は教科書とは全く関係ない。そもそも謎解きのためにチャレンジする問題が「ホテルにチェックインする」や「お土産を買う」といった旅行会話である。全問正解のご褒美である「謎」も（ネタバレはまずいのでここでは言わないが）ショボすぎるように思う（図2）。

ナムコのゲーム・デザイナー中村隆之氏は、いわゆるエデュティメント・ソフトについて、次のように述べている^[7]。

ためになる、とか、勉強の役に立つ、というのは、後から付いてくるもので、まず「面白いゲームを作る」ところからスタートしないと、いいものはできないと思います。その点で、教師の方の提案される学習用ゲームと、我々が考えるゲームとの間にはまだまだ開きがありますね。

ゲーム・メーカーの作るゲームと一緒にすることはできないが、市販の「教育・学習」ソフトに

搭載されるゲームは、面白いものが少ない。『快速完成』のゲームも、「教科書のオマケ」的色合いが濃い。ゲーム仕立てにすることによって、不必要な効果音や画像が増え、軽快な操作感は失われる。しかも CD-ROM で学習を進めるには、ゲームに参加し、キャラクターの「励ましの言葉」を聞かなければならないのだ。『完全マスター』が、高機能を目指しながら、学習画面は無駄な装飾を排し、ゲーム・モードは（『快速完成』以上につまらないゲームだが）完全にオプション扱いであるのと対照的である。

そもそも、これだけのボリュームの教科書であれば、対面授業で詰め込み、テストで締め上げれば一定の成果は上がる。授業でソフトウェアを利用するのは、「効率よく端から詰め込んでいく」とこと、「学習の進捗状況をテストで見る」ことをラクに行うためである。『快速完成』のゲームは、教科書の説明をそのまま電子化したものに、小問4～5問の「通過テスト」を付け加えただけのものである。教科書の補助に使うとしても、あまりに物足りない。90分の授業なら、練習問題だけで50問は楽々こなす。せめてそのくらいの分量が入っていなければ、大学の授業で、学生に一斉利用させるのは難しい。

■反応は上々だが…

今回は、入門篇の教材を用いて、既習の発音や基本表現を復習するようにした。対面授業はソフトウェアの操作や内容の説明のために、最初の1回だけ60分行った。その後、週3回90分の授業を行った。自習形式の教材であるが、今回は現行の授業に教材ソフトを用いることを想定し、学習はフリー・アクセスとせずに、決められた時間に教室に集まって一斉に行った。授業時間以外の自習については特に制限しなかった。

最終回には面接による発音テストと、語法のペーパーテスト、自由記入のアンケートを実施した。

教科書は全40課の構成で、週2回、半年間ではちょっと学習しきれないかな、という分量である。にもかかわらず、ほとんどが既習の内容とは

いえ、3週間で「全部やった」という学生が2人いたのには驚いた。

学生のほとんどは、この教材を用いて「役に立った」「役に立ちそう」と認識している。また、「自分の中国語能力が上がった」と感じる学生も半数を超えている。その一方で、「面白かった」「楽しかった」と感じた学生は少なく、学生のモチベーションを向上させるには、「面白いゲームを作る」こととは別の方策が必要であることをうかがわせる結果が出た。

発音は目に見えて向上した。特に母音の発音が「中国語に聞こえる」ようになったことは、全ての学生に共通して見られる、最大の成果である。ただし、発音能力の向上は、コンピュータ・ソフトの利用とはあまり関係ない。授業以外でこのソフトを利用しなかった学生と、自習を盛んに行っていた学生との間に、差はほとんど見られなかった。発音が向上した学生は、教科書付属の音声CDをよく聞いていた。音声CDを全く聞かなかった4名は、コンピュータ・ソフトの利用には熱心だったが、発音能力はさほど向上していない。もしかすると、コンピュータ・ソフトの利用は「あまり」どころか「全く」関係がないのかもしれない。

語法テストは、北海道大学の中国語統一試験と同じ、多肢選択形式で行った。中国語検定試験など、多肢選択試験の対策にコンピュータが役立つことについては、まず異論のないところであろう。今回もおおむね、好結果が出た。ただし、テストの範囲が教材全体のごく一部なので、印刷教材などと比較して、本当に有効性が高いのかどうかは、もう少し検証してみないと何とも言えない。

図3 正解すると爆竹が・・・



いずれにせよ『快速完成』を授業の全部または一部と置き換え、またはネタとして使うのは難しいものの、印刷教材と比較して圧倒的に高い効果を上げられるものではない。CALL 教室に全面的に導入することには慎重にならざるを得ない。

☞ 単機能教材ソフトウェア

語彙や発音など、特定の能力養成に特化した教材ソフトウェアは、「素材」「部品」として授業にアクセントを添えることができそうなものだが、ソフトウェアによっては、結構難しいこともある。

中国語 CAI 研究会でも紹介された『中国語発音講座』^[8]は、「授業でも自宅でも学習できる教材」を意図して作られたという^[9]。中国語入門時に必ず学ぶ音節表が、口元を見るためのビデオ付きで参照できる。これは、入門時にまとめて学習する以外に、中級者が発音を確認するための、いわば「音声辞書」として利用する、という使い方もある。発音を忘れさせないために、「いつでも音節総表に立ち返れ」と言ってみるものの、放っておいたら教科書の発音編など2度と開かない。パソコンでいつでも参照できる教材を用意することには意味がある。この教材を使った授業の進め方については、本誌創刊号の CAI 特集の中で村上公一氏自身が詳細に述べているので、そちらを参照してほしい。

『タイピングで覚える中国語の基礎 1・2』^[10]は、それぞれ HSK 甲級と乙級の単語を学ぶ、語彙学習ソフトである。学習は表示された漢字をピンインでタイプしていくだけ。成績表示や、間違えた語を集中的に学習する機能もある。正解すると鳴り響く爆竹の音がご愛敬（うるさい、という学生が多い）。このソフトは起動すると、3分から18分の設定した時間、延々問題を出し続ける。何問、という目標は示されないし、出題内容は前回間違えた語以外はランダムに決められる。独習ならいいが、授業で使うなら授業前5分ほどで、前回習った単語を10～20問だけ復習する、という使い方をしたい。「始め」と「終わり」がはっきりしない、だらだら続く単調な練習では、その

まま授業のアクセントとして使うことはできない。授業の進行に合わせて、ソフトの扱う語彙の範囲を細かく調整する必要がある。

ところが、このソフトウェア（および多くの同種のソフト）は自分でデータを作って読ませることができない。市販教材がまだまだ少ない今の状況では、自分の授業に合った教材を網羅的に集めようとすると、相当の部分が抜け落ちる。この部分は教材を自作するしかない。教材ソフトが出回るまでには、まだまだ時間がかかるのだから、このソフトのシステムだけを使い、単語データを入れ替える、クリア条件を自由に変更する、など授業用途の機能にぜひ対応してほしい。

このソフトは 19 名の学生が試用したが、たちまち「飽きた」と言ってくる学生が多く、あまり活発に利用されなかった。

ただし、販売会社であるフーシャン・メディア・ジャパンの担当者からは、「大学・教官・授業に合わせてカスタマイズしてもらえるのか」というこちらの質問に対し、前向きな回答を得ている。「サーバ——クライアント型の教材管理」や「学籍管理」にも、有償または無償で対応可能、ボリューム・ライセンスあり、ということなので、これから CALL システムを導入する、という大学ならば、開発に深くかわかり、ソフトを「叩き」まくる苦労を惜しまないという前提で、検討する余地はあるだろう。

❷ 汎用教材プラットフォームの多言語対応

北海道大学では、2002 年 10 月から ALSI の CALL ソリューション、CaLabo2000 を導入している^[11]。従来の LL 機能とコンピュータによる学習支援機能を組み合わせた製品である。CaLabo シリーズには、言語学習用のシステムであるにも関わらず、多言語に対応していないという欠陥があった。出欠やペア・レッスンといった授業管理機能はともかく、学習機能の中核である、汎用ドリル教材作成実行ツール「SmartHTML」と、音声教材作成ツール「ソフト・テレコ」で中国語

の文字が表示できないのは致命的である。この点について、2002 年 9 月の ALSI 担当者を交えた打合せの際に、中国語を含む多言語対応を要望事項として示した。同様の要望は、既に他の大学からも出されていたようで、「2003 年 3 月リリースの新バージョンで対応する」との回答を得た。

やや遅れたが、2003 年 7 月にリリースされた新バージョンでは、SmartHTML の教材サーバに Solaris を加え、データベースを Oracle にするなどの変更を行い、正式に多言語に対応した。

SmartHTML は、動画や音声を組み合わせた四択問題や穴埋め問題を簡単に作成でき、学習履歴や成績の管理に対応した、Web ベースのオーサリング・ツールである。単体でも利用可能だが、CaLabo シリーズの学習機能の 1 つとしての利用が想定されている。

北海道大学では 2003 年 9 月末から、新しい CALL 教室が稼動する予定で、従来のシステムも全てこの新バージョンに置き換えられる予定である。本稿執筆時点（2003 年 8 月）では、まだ稼動していないため、多言語対応の状況や中国語授業での利用について詳細なレポートはできないが、タイトル、問題文、選択肢等全ての入力と表示に中国語が利用できることを確認している。

多言語対応以外の部分については、目に見える変更はあまりない。「ワープロ感覚で問題作成」という謳い文句通り、動画・音声などのメディアファイルを選択し、テキスト・ボックスに必要な事項を入力するだけで問題が完成する。だが、問題をテキスト形式／XML でまとめて作って、読み込ませるといった、生産性を上げる仕組みは、前バージョンに引き続き用意されていない。

また、端末単位・時間単位での細かなアクセス制御に対応していないため、試験には使いにくい。たとえば、複数のクラスで同じ試験を時間差で実施するような場合、後のクラスの学生は、他の教室にある端末で、実施中の試験問題をこっそり参照できてしまう。ALSI によると、同社のサーバ製品 eNetLibe の機能を用いることにより、目的を達することができるそうだが、CaLabo と eNet は、本来異なるソリューションであり、その「継

ぎ接ぎ」で対応、というのは CALL システムをトータル・ソリューションとして提案する企業のあり方としてスマートとは言えないだろう。

ソフト・テレコは LL 教室のテーブルレコーダの機能を、PC 上で実現したソフトウェアである。教師卓で、CD などのソースからサーバに音声をアップロードすると、直ちに学生用 PC から参照できる。自分の声を録音して聞き比べるなど、LL 用テーブルレコーダの一通りの機能を、ソフト的に実現している。新バージョンでは、音声と文字表示をシンクロさせる機能が追加された。また、MD にダビングしたり、フロッピーなどのメディアに再生用ソフトとともにデータをコピーしたりすることが可能になり、教室外での学習にも対応した。

ソフト・テレコの新バージョンは既に導入済みで、2003 年度前期から初級授業 2 クラスで利用している。ただ、目玉の文字表示機能とメディアへのコピーはほとんど利用していない。

文字表示機能では、中国語の入力・表示ができる。原文と訳文、もしくは原文とピンインなど、2 種類の文字表示を切り替えることもできるが、原文とピンインと訳文といった 3 つの「字幕」を切り替えることはできない。結局、初級教科書の 8 文～10 文程度のダイアログであれば、ソフトの文字表示機能を使うよりも、ピンインと漢字を併記した教科書を見ながら聞く方がわかりやすい、というのが大半の学生の意見である。教員としても、教科書に書かれている内容を、わざわざ入力し直す手間はかけたくない。当初は漢字とピンインを貼り付けていたが、メリットがあるとは思えないので、学期途中で止めてしまった。

北海道大学情報基盤センターでは、オリジナル中国語教材の作成を予定しており、音声 CD の代わりに、ソフト・テレコ教材を収めた CD-ROM を配布することを検討している。ソフト・テレコは既存の教材を利用する「テレコ」ではなく、音声とテキストを含んだ、オリジナル教材を簡単に作成できる「簡易オーサリング・ツール」として利用するのが最も適当だと思われる。

📎 おわりに

以上、市販の教材を中国語授業に取り入れる試みについて紹介した。大学での中国語 CALL 授業は、まだまだ自作教材を用いることが多い。いくつかの教材を使ってみてその理由がよくわかる。

コースウェアは個人向けの製品がほとんど全てを占めており、大学の教室で用いるメリットは少ない。発音や語彙の単機能教材ソフトは、うまく取り入れれば効果を上げることができるが、欲しい教材が展開されているかどうかかわからない。また、汎用プラットフォームの多言語対応は進みつつあるが、言語教材の実行環境としては、不足する部分が多すぎる。

「いっそ使わない」というオプションが認められていない場合は、おおよそ、上記のような問題点を十分認識した上で、単機能のソフトを授業のネタに用い、足りない部分は、そのソフトをマネして自分で作る、というのが現実的な対応であろう。

注

- [1] 東方書店／クリエイト大阪、林要三企画・監修
- [2] マネージ <http://www.manage.co.jp/>
- [3] アークビレッジ <http://netpal.co.jp/china/>
- [4] 創育 <http://www.soiku-mm.co.jp/>
- [5] メディアファイブ <http://www.media-5.co.jp/>
- [6] ただし、課文がゴシック体であること、カナが併記されていることなどは、個人的にあまり好まない。
- [7] 2003 年 1 月北海道大学で行われた座談会での発言
- [8] 村上公一編、早美出版社、1999
- [9] 村上公一、「コンピュータによる中国語学習支援」、『大学教育と情報』Vol.10 No.4、私立大学情報教育協会、2002
- [10] フーシャン・メディア <http://www.fushan.com/jp/>
- [11] アルプス・システム・インテグレーション (ALSI) <http://www.alsi.co.jp/>

Blackboard Learning System について

バーチャルユニバーシティ構築の可能性

高橋 由利子（たかはし ゆりこ）

1. Blackboard について

1.1. Blackboard とは

Blackboard とは略称で、正式名称は Blackboard Learning System ML (Multi Language の略)、Blackboard 社^[1]のソフトウェアである。

Blackboard は、単にドリルやテストを行うだけの単純なソフトウェアではない。また、それ自体に教材が含まれているわけではない。授業を中心とする大学での諸活動を支援する枠組みをネット上に構築するソフトである。

例えば、教員は授業のためのシラバス、教材、宿題、テストなどのさまざまなコンテンツを、素人にも使いやすいコンテンツ作成ツールによって、ネット上に簡単に構築することができる。そして学生は、それらを自由に利用することができる。

さらには、電子会議システムなどのグループウェア機能も備えている。

このように、ネット上のスペースの提供とコンテンツ作成の 2 つの機能が一体となる形で、インターネット上に、いわゆるバーチャルユニバーシティを立ち上げるのが、Blackboard の特徴である。

筆者の所属する上智大学比較文化学部でも何人かの教員が海外の Blackboard を視察したり、トライアル版を試用した結果、前述の特徴が我々の要求にかなっていることがわかったので、2002 年の秋学期から導入した。そのため、まだ利用期間が短く十分に活用しているとは言い難く、コンテンツ作成も途上にあるが、以下の大学は十分な運用経験を持ち、コンテンツも充実しているので、参照されたい。

<http://aims.crdc.gifu-u.ac.jp> (岐阜大学)

<http://bb.nucba.ac.jp> (名古屋商科大学)

<http://www.e-nufs.net> (名古屋外国語大学)

<http://bb.tamagawa.ac.jp> (玉川大学)

<http://bb.kobe-c.ac.jp> (神戸女学院大学)

<http://blackboard.duke.edu> (デューク大学)

<http://blackboard.Princeton.edu> (プリンストン大学)

<http://edventure.ntu.edu.sg/> (南洋工科大学)

Blackboard は多言語への対応も進んでおり、コンテンツ作成機能はもとより、インターフェースも多言語に対応している。

Blackboard にはベーシック版とエンタープライズ版がある。ベーシック版はインターフェース言語の数が 2 種類 (上智大学比較文化学部は英

語と日本語)であるのに対して、エンタープライズ版は中国語、韓国語を含む12カ国語から選ぶことができる。

また、エンタープライズ版は大学の事務運営のための学生管理システムも備えている。

1.2. Blackboard の導入

1.2.1. Blackboard の動作環境

Blackboard は Web アプリケーションソフトであるので、専用サーバの設置が必要となる。このため十分なサーバ管理体制が必要であり、小規模な授業や学科単位での導入には向かない。

現在は Windows Server 版と Solaris 版が提供されている。動作環境の詳細についてはブラックボードジャパン社の Web サイト^[1]を参照していただきたい。

1.2.2. ユーザの利用環境

Blackboard は、インターネットを通じてコンテンツを提供する Web アプリケーションソフトであるため、ユーザの側で特別なソフトを用意する必要はない。インターネット接続環境と、多言語処理機能が設置されたパソコンがあればよい。

もちろん、授業時間中にパソコン教室や CALL でも運用することができるが、学生が時間や空間に制約されず、自宅からでもアクセスできる予習復習、授業フォローアップツールとしての運用に適している。筆者も教室外でのコミュニケーションの一つとして、特に学生に利用を強制はしていない。

1.2.3. Blackboard が提供するもの

前述したように Blackboard が提供するものは、インターネット上の空間とコンテンツ作成ツールであり、教材そのものではない。

従って、各大学のカリキュラムに応じた授業科目の設定、各科目の教材の作成等は、それらのツールを使って、各教員がやらねばならない。

だからこそ、それぞれの大学や教員の特色を反映したさまざまなコンテンツが、Blackboard 上に立ち上がってくるのである。コンテンツ作成の詳細については第3章で述べることにする。

1.2.4. ユーザの種類と権限

Blackboard ではその権限によって次の4種類に設定されている。

- システム管理者：Blackboard 上に設置する授業科目名、Blackboard の細かい仕様、ユーザ管理などの権限を持つ。
- 教員ユーザ：自分の担当科目について、通知事項、授業予定表、シラバス、スタッフ情報、参考資料、教材、テストなどのコンテンツを作成したり、コンテンツへのアクセスログや試験の得点の一覧表を閲覧する権限を持つ。
- 学生ユーザ：各科目のコンテンツを利用することができる。ただし利用に際しては、あらかじめ Blackboard 上でその科目に登録しなければならない。
- ゲストユーザ：特別な ID を持たない。コンテンツをどこまでゲストユーザに開放するかは、システム管理者と教員ユーザが自由に設定できる。

上述のように大学や各教員によって、ゲストへの開放度を自由に設定できるため、大学によっては外部への開放を積極的に推進しているところもある。

2. ユーザインターフェース

2.1. 教員ユーザ

図1以下は私が教員ユーザとして上智大学比較文化学部の Blackboard (<http://blackboard.sophia.ac.jp/>)にログインした後に表示されるユーザページである。画面右側にシステム管理者によって設定された担当科目一覧が表示されている^[2]。

各科目へのリンクをクリックすることによって、その科目のコンテンツを作成するページに移動することができる。

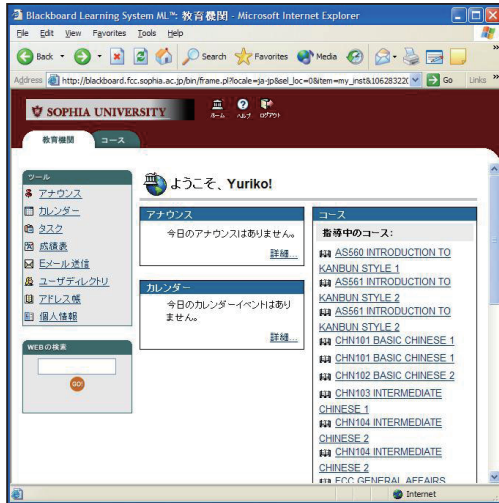


図 1

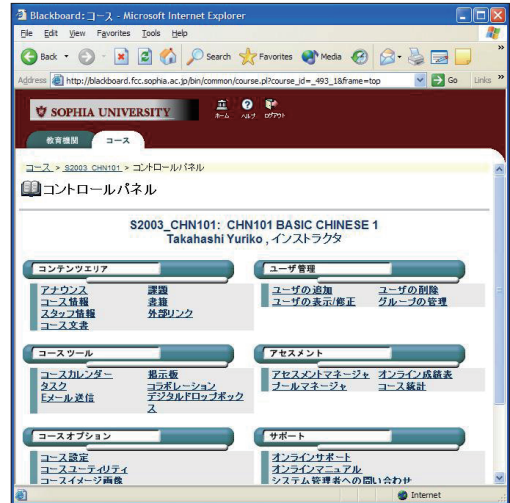


図 2

2.2. 学生ユーザ

学生ユーザの場合は初回ログイン時に科目を登録しなければならない。画面に表示される授業科目一覧から履修科目を選択して登録すると、次回ログインからは、それが履修科目として自動的に表示される。ただし、それはあくまで Blackboard 上での登録で、上智大学比較文化学部の場合は事務上の実際の登録とは連動していない。そのため、学生は実際には履修していない科目を Blackboard 上で登録して、そのコンテンツを利用することもできる。

3. コンテンツの作成

3.1. コンテンツの種類

Blackboard の作成ツールによって提供できるコンテンツは以下の 3 種である。

教材コンテンツ、コミュニケーションツール、テスト

これらのコンテンツを作成するには教員ユーザの担当科目名をクリックする。すると、科目ごとのコントロールパネルが表示される。そこでメ

ニューから作成したいコンテンツの種類を選んでクリックすると、それぞれのコンテンツ作成ページに移動する。以下、コンテンツの種類ごとに作成方法を紹介していく。

3.2. 教材コンテンツ

上智大学比較文化学部の Blackboard 上では教科書や授業資料などをコース文書と総称している。教員は授業にあわせて自由にコース文書を設置することができる。実際の設置作業は科目ごとのコントロールパネルから行う。(図 2)

上の画面は春学期初級中国語 1 のコントロールパネル画面である。画面左上のコンテンツエリアの中のコース文書をクリックすると教材コンテンツの作成ページに移動する。

コース文書は Web ページ上のテキストボックスに直接入力して作成することができる。入力形式はプレーンテキストの他、HTML 方式にも対応している。Blackboard はユニコードに完全対応しているため多国語（例えば日本語と中国語と英語）を混在して入力することもできる。

コース文書はまた教員が作成したファイルをアップロードして配布することができる。ファイルの形式はワード文書、音声、動画など何でもよく、Web ページ上でローカルファイルを指定するだけで簡単にアップロードする機能が備わって



図 3

いる [3]。

さらにこれらを組み合わせて、例えば解説文付きの動画ファイルとして提供することもできる。

コース文書はフォルダを作り、フォルダごとに分類して配置することもできる。授業の進度や内容にあわせて分類することもできるし、容量が大きくなりがちなコース文書の設置場所を分散することで、アクセシビリティを向上させるために使うこともできる。(図 3)

上の画面はあるフォルダの中のコース文書を一覧表示させたところである。フォルダの中のコース文書が順に表示され上にはアップロードしたファイルのリンクが、下にはテキストボックスから入力した文章が表示されている。学生はこのリンクをクリックしてファイルをダウンロードして利用することになる。

3.3. コミュニケーションツール

コミュニケーションツールとしては、掲示板、Eメール送信、デジタルドロップボックスなどの機能が提供されている。

掲示板はスレッドノート形式である。設置できる掲示板の数の制限はない。学生ユーザは自由に書き込める他、許可されていれば掲示板に新たな

スレッド(話題)を立てることもできる。

デジタルドロップボックスは教員と学生がファイルをやりとりするための機能である。例えば、教員が映像データを送り、その聞き取り結果をワードファイルで提出させることも可能である。

3.4. テスト

Blackboard のテスト作成機能は非常に充実している。テスト作成ツールはコントロールパネルのアセスメントマネージャで呼び出すことができる。

対応する出題タイプは次の 7 つである。

穴埋め問題、マッチング、複数回答、多肢選択、並べ替え問題、短答・エッセイ、正誤問題

マッチングとは A 群と B 群の選択肢の関連するものを結ばせるタイプ、複数解答とは記述問題で複数の正解があり得るもの、短答・エッセイとは、自由記述問題で、解答は指定せず、教員が画面上で採点していくことになる。

これら複数の形式を組み合わせることで試験問題を作成することができる。また、同一タイプの問題をプールしておき、それをランダムに呼び出して、試験問題を作成することもできる。特にドリル問

図 4





図 4

題に適しているといえる。

以下は多肢選択と穴埋め問題で作ってみた試験問題の例である。(図 4)

これらの試験問題は各設問ごとに時間を制限できる他、試験にアクセスできる時間帯を指定することにより、学生に一斉受験させることも可能である。また、試験は 1 回限りにすることもできるし、練習問題として何度でも受けられるようにすることもできる。後者の場合、データベースには最終回の点数が保存される。配点は教員が自由に指定できる。解答が指定されている問題については、採点が自動的に行われ、受験者はただちにその得点を見ることができる。試験結果は Blackboard によって自動的に集計される。教員はコントロールパネルのオンライン成績表をクリックすれば、それを閲覧保存できる。

3.5. その他

以上に紹介したほかに Blackboard はさまざまな機能を備えている。例えば、受講者へのアナウンス機能、各コンテンツへのアクセス統計などがある。すべての機能を使えば効率的な授業運営が可能になるように設計されている。(図 5)

上の画面は一つの科目のアクセス統計画面であるが、スクロールしていくと、アクセスの詳細が、授業に登録した学生の一覧表とともに表示される。

4. 初級中国語春学期における使用例

ここで筆者の実際の経験について書いてみたい。筆者のコンピュータスキルは Windows の基本操作を一通りマスターし、ワープロによる文書作成をこなす程度である。そのため色々と不十分な点があると思うが、逆に、この程度のコンピュータスキルでも、この程度のことはできるという風に理解していただければ幸いである。

4.1. 使用クラスについて

上智大学比較文化学部は完全セメスター制を採用しており、週 2 回 2 コマを同じ教員が担当しており、半年で成績をつける。すべての授業は英語で行われているので、教科書は Elementary Chinese Readers (revised version) book 1 (基礎汉语课本, 华语教学出版社) を使用している。

4.2. 作成したコンテンツとその運用方法

以下の 3 種類のコンテンツを作成した。

- 中国語の基本母音と子音、声調のビデオファイル
- 教科書の第 1 課から第 6 課までの新出単語と会話のビデオファイル
- 聞き取りテストの回答のビデオファイル。

これは、いずれも、授業のフォローアップとして予習や復習に使い、初級段階の発音の定着率をあげることを目的としたものである。発音については市販のテープもあるが、知らない中国人のネイティブスピーカーの発音よりも、普段接している日本人教員の姿や発音、口の形が映し出されるほうが、授業との連携性が保たれ、リピート効果も大きいと思われたからである。

学生は 32 人中 24 人がアクセスした^[4]。アクセス数と個人の成績との間に特に相関はないが、全体的にみると授業の前の日のアクセス数が多い。

ただ、主観的な印象ではあるが、発音だけに限

ていえばかなり定着率もよいようであり、授業以外の日でも教員の授業での顔に接することができるせいか、学生のモチベーションが例年になく高く、授業時間の始まる前から中国語の勉強をするのだという雰囲気ができていて、すぐに集中することができ、より有効な授業ができた感触があった。

今後は中級の中国語や漢文の授業などにも、また違ったかたちで生かせる方法を模索してゆきたい。ただ、使ってみて感じたことは、結局自分自身が個々の授業についての運営をどのように行うかについてきちんと考えていかねば、Blackboard を有効に生かす方法も模索できないということである。Blackboard を使うことによって自分がどのように授業を運営していくかを考えざるを得ず、実際の授業の組立てや内容についても、考え直すよいきっかけとなった。今後はそのようなことも含めて Blackboard をどのように授業に生かしてゆき、有効に利用するかを考えてゆきたい。

📁 5. 終わりに

いままでに述べたように Blackboard は大学にとって非常に有効なシステムであるが、やはり、そのシステムの一部が不安定になることがあった。これはおそらくサーバ設定上の問題と思われるが、学生に常時提供するサービスが不安定であるのは、

やはりよくないことであるので、今後その原因をきちんとつきとめて、解決の道をさぐっていかねばならない。

また、システム管理者はすべてについて情報を見ることができるため、色々な方面から教員の授業に対する過度な干渉に利用される危険性を指摘する教員もいる。

また、教員のコンピュータスキルも一定ではないため、学部として導入したとはいえ、ただちにすべての教員が使い始めたわけでもない。今後それらの色々な問題を少しずつ解決し、Blackboard を活用するためのサポート体制を確立し、より有効にシステムを運用できるようにすることが、今後の課題である。

注

- [1] 本社はアメリカ、ワシントン D.C.。日本にブラックボードジャパン社があり、2003 年 8 月 28 日に第 1 回日本ユーザ会が、青山本社で開かれた。<http://www.blackboard-japan.com/>
- [2] 上智大学比較文化学部 of 授業はすべて英語で行われるため、授業科目名はすべて英語で入力されている。
- [3] ただしアクセス集中が予想される授業などでの一斉利用には適さないものと思われる。
- [4] その他、正式履修登録者以外に 3 人が Blackboard 上での登録を行い、コンテンツを見ている。

ソフトウェア レビュー

本レビューは、一般的なソフトウェアの機能紹介よりもむしろ多言語（多漢字）対応に対する論評を中心とする。また、中国・台湾など漢字圏のソフトウェア製品のレビューにも多くの誌面を割いている。これらは他の雑誌媒体では見られない本誌独自のコンテンツである。

オペレーティングシステムや WWW ブラウザの Unicode 対応は一段落したものの、各アプリケーションの Unicode 対応は一部のソフトウェアを除いて未だ進んでいない。その中でも Chinese Writer 7 の GB18030 への対応や、定番エディタ秀丸の Unicode 対応が遅ればせながらも進みつつあるのは喜ばしい限りといえよう。

本レビューは、各レビュー執筆者が本会の主要な活動であるインターネット掲示板上で日々行われている情報交換や討論をベースとして、それぞれ独自にソフトウェアの検証作業をすすめた末に執筆されている。ただ、各執筆者の好みなどもあって、必ずしも全てのソフトウェアが検証されているわけではない。もしかしてこのレビューから漏れているすばらしいソフトがあるかもしれない。「こんなソフトウェアを知っている」という方がおられたら、是非本会の掲示板に書き込んでいただきたい。

Contents

OS	総括.....二階堂善弘 104	Linux.....師 茂樹 105
ワープロソフト	一太郎 13 & ATOK16....山田 崇仁 108	EGWORD.....二階堂善弘 110
IME と 中文ツール、フォント	cWnn（楽々中国語）.....二階堂善弘 111	Chinese Writer 7二階堂善弘 112
	フォント.....千田 大介 113	翻訳ツール.....千田 大介 115
ネットワーク	WWW ブラウザ秋山洋一郎 121	その他.....千田 大介 124
その他	テキストエディタ.....山田 崇仁 127	その他.....千田 大介 130
	中国のソフトとネットの動向.....千田 大介 158	



総括

二階堂 善弘

□ Windows と Mac は大きな進展無し？

これまで毎号、主に各 OS の多言語対応を中心にレビューを行ってきたが、今回はいささか困ってしまった。

その理由は他でもない、Windows や Mac OS などの OS に、メジャーなバージョンアップが無かったからである。例えば、Windows の次期バージョン、Longhorn や、次の Mac OS の Panther (Mac OS 10.3) などは、様々な媒体によりその概略が伝えられているが、双方とも現時点 (2003 年 8 月) ではまだ発売されていない。

また多言語対応などについてどれくらい進展しているのかについては、不明な点が多い。おそらく、双方とも次期バージョンについては、Unicode への対応が進むなど、実装面での期待はできるが、それほど大きな進展は無いと予想している。

むしろこの間は、ワープロやブラウザなど、アプリケーションソフト側の多言語対応が進み、それによって OS の能力が引き出されてきた、というのが傾向としては強かったように思える。

□ Linux や超漢字における変化

一方 Linux においては、カーネルは間断なくアップデートされてきた。しかしこちらも目立った変化には乏しい。とはいえ、Linux は当然ながら各ディストリビューションで大きな差があるので、注意が必要となる。

中でも RedHat Linux のバージョン 9 へのバージョンアップは特筆すべきで、ログインとログアウトのみで簡単に使用言語を切り替えることが可能となった。

以前に UNIX の Solaris において実現されていた多言語対応が、ほぼ Linux においても備わったといえる。この変化は大きい。

超漢字については、メジャーなバージョンは 4 のままであるが、細かいアップグレードは行われている。特に、R4.100 のバージョンアップでは、GB18030 に対応したことにより、中国漢字の領域に大きな変更が加えられている。しかし、現在はまだそのほとんどの領域は未定義のままである。GB18030 の漢字はほとんど Unicode と一致するので、そういった意味では形を変えた Unicode 採用と言えよう。また漢字検索の時に Unicode の番号が表示されるようになったことから、統合漢字的な方向への傾斜は明らかである。ただ、これはむしろ好ましい変化であると考えられる。

□ 総じて乏しい変化

どうもこの 1 年くらいは、Windows と Mac OS のメジャーなバージョンアップの谷間となり、総じて OS 面での変化が乏しい時期に当たってしまったようである。

また、OS 自体の選択肢が少なくなっていることも、別の意味で問題かもしれない。かつては、OS/2 Warp や Be OS など、特徴を有した OS がそれなりに存在した。むしろこれらの OS も引き続き使われており、それなりのアップデートも行われてはいるようだが、それは発展というよりは、保守点検に近いものとなっている。

もっとも、Be OS については、これを改造し、Zeta という新 OS として開発するプロジェクトもある。多様化という面では、こういった動きがもっと出てくることが望ましいと考える。

Linux

師 茂樹

□ Linux のメリット／デメリット

■ free ということ

Linux や FreeBSD をはじめとする所謂 PC UNIX に注目が集まって久しい。筆者の周りに多くいる東洋学の研究者の間でも Linux に対する関心は高いが、Windows や Macintosh ほど普及しているわけでもないで、二の足を踏んでいるという人が多い。

"free" には「自由」のほかに「無料」という意味がある。Windows の世界で「フリーウェア」というと後者の意味で使われる場合が多いが、Linux の世界では明確な定義^[1]に基づいた前者の意味で用いられる。すなわち、プログラム（が書かれたソースコード）を入手、実行、研究、修正、再頒布、公開する自由が保障されたものを Free Software と呼ぶのであり、ここには「コピーを売る自由すら含まれ」る。

Red Hat Linux などのディストリビュータは、この定義に従って Linux を売り、同時に OS 本体を含むほとんどの部分を FTP などで公開しており、ダウンロードすればそのまま使えるようにしているのである。

■ free であるメリット

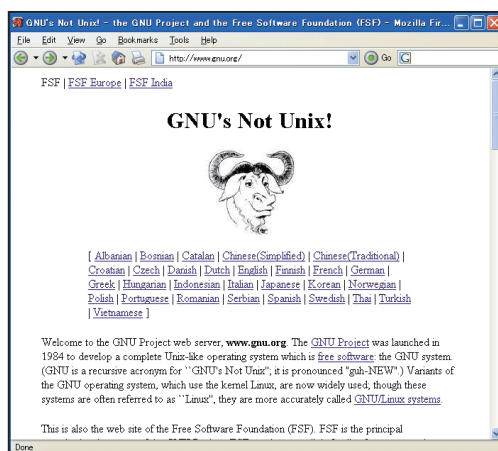
Web 閲覧やメール、ワープロなどの作業について言えば、今日、どの OS でもあまり違いはないと思ってよい。各 OS それぞれに長短と作法があることを認識し、選択にあたってはこれまでの経験や操作感、求めているツールがあるかないか（中国語で文書を書くツールについては Windows 系がもっとも充実しているだろう）、まわりに使っている人が多いか少ないか、などで判断すべきであって、Linux を特別視する必要は（他の OS と同様）ない。

しかし、特に研究・教育分野において Linux を積極的に選択する理由は少なくない。まず、あらゆる面で改造がしやすい点をあげたい。例えば、テキスト・ファイルの文字コード変換などを行っている際、ある文字の扱いが自分の思う通りにならないため、その部分だけ毎回手で直す、というようなことは少なくなろう。

こう言った変換テーブルは、ツール自体に含まれていたり OS に組み込まれていたり様々で、ツールの場合、それを部分的にカスタマイズする機能を備えているものもあるが、そうでないものも多い。

また、Windows に組み込まれた変換テーブルを自分なりに直すことは、ライセンス的にも技術的にも非常に難しい。一方、Linux という OS は徹底してテキスト・ファイル主義なので、改造はテキスト・エディタさえあればほとんどの場合事足りてしまう。もちろん、ライセンス的には何の問題もないし、それを公開することも可能だ。研究・教育活動においては汎用ツールではカバーできない特殊な作業をすることも多い。その点、Free Software のこうした点は大きな魅力である。

また、近年、インターネットを通じた学術情報の発信が強く求められるようになった。例えば、Microsoft Word で作った文書を PDF で配布することを考えてみよう。変換結果を見ると、フォントが別のものに置き換わっていることに気づくはずだ。場合によっては埋め込みフォントの変換ができずに失敗してしまうこともある。これは、Windows 上で用いられているフォントの多くが再配布を禁止しているため、PDF への埋め込みが拒否されるからである。しかし、Linux に収録されているフォントはほとんどが free であるので^[2]、このような問題が起きることはほとんどない。



それどころか、研究活動に便利なシステムを構築した場合、それを OS ごと他の PC にコピーする、などということも可能だ。実際、自分なりにカスタマイズした Linux を CD-ROM 等で配布するということは広く行われている（後述の Knoppix など）。

Free Software はしばしば「伝播性」を持つと言われ、発信した情報やツールの寿命が最も長くなることが期待できる。学術情報にとってそれは得がたいメリットではないだろうか。

■ free であることのデメリット

このような自由度の高さが、かえってデメリットになってしまうこともある。実はユーザー・インターフェースや多言語がらみは、それが顕著に出してしまう部分でもある。

例えば近年 Unicode が主流になると、Microsoft や Apple などのベンダではトップダウンによって OS 全体を一気に Unicode 対応してしまい、サードパーティーもそれに追随した。ところが Linux のように各自が自由に開発した部分が集積した形でできあがっている OS の場合、このようなトップダウン的な変化はなかなか起こらない。Li18nux（現在は OpenI18N.org）^[3] などの努力によって国際化の統一指針は用意されたものの、全体が一気に変わることはあり得ないだろう。多言語に強い Emacs は、Linux が誕生する遙か以前から存在してきたものである。そこで蓄積されてきたものを捨てて Unicode ベースに生まれ変わるのとはなかなか難しい。トップダウンをよしとするか、伝統や多文化的な世界を大切にしたいと思うか、甲乙をつけることは簡単ではなからう。

□ オルタナティブとしての Linux

とは言え、最近では、Windows や Mac OS に迫るような統一され洗練されたインターフェースを持つものも増えてきている。メジャー・ディストリビュータに大きな動きがなかったこともあり、以下、最近注目されつつある手軽なディストリビューションについて簡単にレビューしたい。

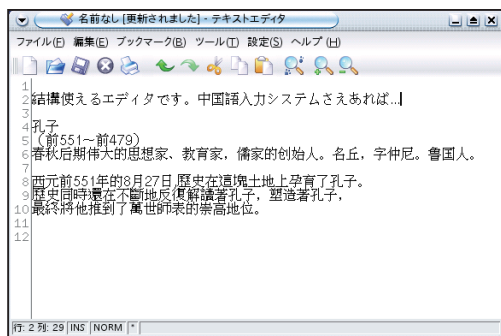


■ Lindows OS 4.0

その名前からして、Windows への対抗意識を強く打ち出す Lindows OS であるが、中国のディストリビュータ Xteam（冲浪平台）が出していた同名のものとは異なる、アメリカで作られた Debian GNU/Linux (sarge) ベースの新しいディストリビューションである^[4]。なお、バージョン 4.0 日本語版の発売が 8 月 29 日ということもあり、本レビューはほんのわずかな期間の使用に基づくものであることをお断りしておく。

Windows とほぼ同じショートカットを備え、ホイールマウスにも対応するなど、操作性は Windows とほとんど変わらないといってよい。また、Click-N-Run というソフトウェア・ライブラリの有料サービスがあり、一定料金（年間契約）で StarOffice などがインストールし放題となる^[5]。





ブラウザとメーラは Mozilla / 和ジラをベースにカスタマイズしたものであるため、多言語環境も Mozilla のそれと同等である。フォントの種類が不十分であるため Extension B 領域などは表示されないが、きちんと処理されているようである^[6]。

また、付属のテキスト・エディタが Unicode ベースで多言語に対応しており、言語を混在させることも可能だ。

日本語環境は ATOK 12 が搭載されて、一応合格点をあげられると思うが、惜しむらくはその他の言語の入力システムがデフォルトで提供されていないことである。ただし、Click-N-Run などから中国語入力システムを入手可能なようである^[7]。

■ Knoppix 3.2 日本語版

CD-ROM だけで起動する Knoppix は、やはり Debian GNU/Linux ベースの所謂 1CD Linux のひとつで、AIST が日本語化したものを公開している^[8]。Open Office を始めとする一通りのアプリケーションが日本語化されて収録されており、また非常に強力なハードウェア自動認識機能を備えているので、ほとんどの場合 CD-ROM を入れてしばらくすれば KDE の洗練されたデスクトップが表示されるだろう。Knoppix 側からハードディスクを読むことができるが、書き込みはできないので Windows が壊れるということはない。Linux を入れるマシンの余裕がない場合、あるいはどんなものか少し試してみたい場合などに便利だろう。

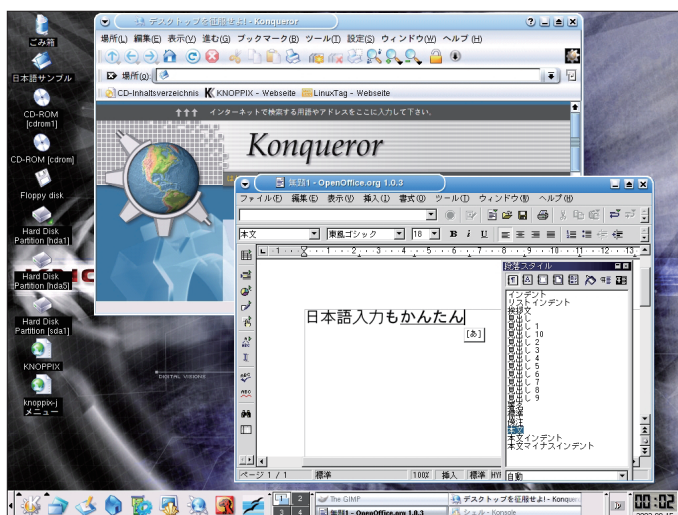
ちなみにこの Knoppix は、ハードディスクにインストールすることも可能であ

る。CD-ROM で試し、気に入ったらハードディスクに入れる、ということもできる。

CD-ROM で動いているうえ、商用ソフトウェアを利用していないことから、Lindows よりは若干劣るものの、まずまずの操作性と言ってよいのではないかとと思う。多言語環境もデフォルト状態の Lindows と同程度であるが、こちらは追加インストールが基本的にはできないため、日本語以外の文字を書きたい場合には XEmacs などを使うか、自分で CD-ROM を再構築しなければならない。

注

- [1] <http://gnu.ghks.de/philosophy/free-sw.ja.html>
- [2] これに関連して、従来フリーのフォントとして配布されてきた所謂 Watanabe フォントが、実は無断複製されたものだったことが発覚し、公開の中止が相次いだ。
<http://khdd.net/kanou/fonts/stolenbitmap.html> 参照。
- [3] <http://www.li18nux.org/>
- [4] <http://desktoplinux.com/articles/AT2152579590.html>
- [5] もちろん元は Debian なので、apt によるパッケージの入手も可能。
- [6] 本誌所収の秋山陽一郎氏によるレビューを参照。
- [7] http://www.lindows.com/lindows_products_categories.php から Chinese で検索すれば中国語関連のソフトウェアを確認できる。ただし、ほとんどが apt-get で入手可能だと思われる。
- [8] <http://unit.aist.go.jp/it/knoppix/>



❖ ワープロソフト

一太郎13 & ATOK16

山田 崇仁

□ 一太郎 13

■ 一太郎に追い風は吹かないのか？

評者は自分の Web サイト^[1]では、「MS Word を使いこなす」とか「MS Excel を使いこなす」とかのページを開設しているが、これはあくまで授業で使用するレジュメとして作成しているものであり、評者本人は 10 年以上一太郎と ATOK を使い続けている。

一太郎と言えば、評者が使い始めた 1990 年代初頭には日本語ワープロの王者だったはずだが、今では Word に押されっぱなしの印象を受ける。雑誌の取り上げ方も、以前なら新バージョンの一太郎と言えば特集記事や見開きページでの紹介が当たり前だったのが、最新版の一太郎 13 の紹介記事といえば、その他大勢の新製品と一緒に雑誌の新製品コーナーの片隅で見かける程度であった。この扱いは、OpenOffice よりもはるかに冷遇されているといっても過言ではない。

しかし一太郎は長い間に蓄積された日本語に対する膨大なノウハウがあり、最近のバージョンでは「一太郎プロンプト」を利用したのスク립トや正規表現まで利用可能になっている。むしろこの点に関しては日本語の扱いにまだ難がある OpenOffice はもとより Word にすら勝っている部分だろう。

それなのになぜ「脱 Word」の風が OpenOffice ではなく一太郎に吹かないのか。一太郎には Windows マシンにプリインストールされている Word を押しのけてまで利用したいというほどの魅力が無いのだろうか。あるいは有料かつ Office パッケージタイプではないというのが大きいのだろうか。

■ MingLiU が使えるようになった一太郎 13

さて、本誌創刊以来歴代一太郎のレビューで評者

が問題にしてきた「Windows 2000 や XP 上で一太郎を利用する際に MingLiU や SimSun など日本語以外の .ttc 拡張子のフォントが使えない」という不具合が、バージョン 13 でついに解消された。

漢字文献を日常的に取り扱う機会が多い評者にとって、上記の不具合は致命的だったため、毎年この時期になると制作元のジャストシステムにこの点について不具合報告を出し続けていた。このために、評者は一太郎で論文を書く事を避けていたが、ようやく今年からは一太郎で論文が書けそうである。

ただし、「ラテン拡張 B、IPA 拡張領域の文字の一部を入力するとフォントが日本語用のものに固定されてしまい Times New Roman、Lucida Sans Unicode などの欧文フォントに切り替えられないため、文字が欠けたり文字幅がダブルバイト文字の幅に固定されてしまったりする」不具合が解消されていない。今後の改善を望む次第である。

■ 相変わらず中途半端な Unicode 対応

MingLiU の不具合は解消されたものの、Unicode 関連では相変わらずという面も多い。例えば UTF-8 で保存されたテキストの読み書きに対応していない。そのためこれらのファイルを利用するには、一端 UTF-8 のファイルの読み書きに対応したエディタでファイルを開いてコピーし、一太郎上に貼り付ける必要がある。

拡張漢字関連^[2]では、Unicode 2.1 への対応で留まっており、Unicode 3.x 関係となると BMP 領域内の Extension A には対応するものの Extension B には対応していない。ジャストシステムは日本から Unicode コンソーシアムに参加している貴重な存在なのだから、Unicode 3.x あるいは Unicode 4 への対応について積極的になって欲しいところである^[3]。

□ ATOK16

■ 相変わらず元気な ATOK

方言への対応・各種拡張辞書の提供など、一太郎に比べて相変わらず元気な ATOK である。

これはやはり長年の積み重ねによる圧倒的な変換効率の良さと、方言への対応やユーザーによる人名辞書の構築などの遊び心を含んだ柔軟性が、システム標準の MS IME を押しのけても使いたいというユーザーを開拓する要因となっているのだろう。

本誌読者の中には、山川出版社の日本史・世界史用語集の用語が辞書化された「歴史用語変換辞書」への興味をもたれる方が多いと思われる。

中国史関連では評者の「ATOK 用中国史辞書」シリーズもあるが ATOK16 に対応した以外は最近更新をさぼっている（これについては、評者の不徳の致す所である…）、本会会員の秋山氏による「漢語詞典」もおすすめておきたい^[4]。

■ こちらも中途半端な Unicode 対応

ATOK の Unicode 多漢字関連に目を転ずると、相変わらず文字パレットに Extension A・Extension B の項目が存在しない。Extension A については、前バージョンで見られた未収録字が表示されてしまう不具合は解消され、BMP 領域の該当コードエリアを指定すれば表示されるものの、Extension B を含む BMP 領域以外の Unicode 拡張部分については未対応である。この辺りは一太郎と全く同じなので、早急に改善していただきたい。

また、手書き入力認識領域に関しては、相変わらず JIS の第一・第二水準止まりであり、補助漢字まで認識可能な MS IME の手書き文字パレットには及ばない。加えて、MS IME には補助漢字まで登録された単漢字辞書がある。これは ATOK にはないアドバンテージである。この辺りは早く追いついて欲しい。その他、部首 + Shift + F6 で行う JIS 第二水準の変換も、Unicode CJK 統合漢字領域の部首で表示されるようになるとうれしい。

■ 効率的な入力手段の提供こそが IME の売り

Unicode に収録される漢字が『大漢和辞典』の収録

範囲をも超えてしまった現在、文字を効率よく入力するという点で、IME の存在価値はますます大きくなっている。

大規模な文字表（たとえ部首・画数をキーとして整理されていたとしても）から文字を選択して入力するというのは、決して効率的な入力環境ではない。あくまで文字検索機能を利用した文字入力は、効率の良い通常の入力手段が確保されて初めて生きる補助的な手段にしか過ぎない。その点からすると、通常の入力手段 + 補助的な入力手段をバランスよく兼ね備えた ATOK の強みというのが際だつのである。

□ おわりに—— Linux への移植——

ジャストシステム社は 2003 年の 5 月 15 日の記者発表で、Linux 用デスクトップ環境向けの製品として、一太郎と ATOK16 の Linux への移植を発表した^[5]。この製品は、従来 Linux 向けに提供していた JAVA ベースの一太郎 Ark や ATOK X ではなく、Windows 用のそれと同等の機能を移植したものとなる予定である。

これら両ソフトが Linux コミュニティ上でどう受け入れられるか、また一太郎が再浮上するきっかけとなるのか、興味深く見守っていききたい^[6]。

注

- [1] <http://www.ritsumeai.ac.jp/kic/~tyv07679/index-j.html>
- [2] また、ジャストシステム独自の Unicode 拡張設定である拡張漢字設定オプションは、Windows NT/2000/XP 環境で使用すると不具合が起こるので、それらのを利用している場合は、この機能を使用しないかアンインストールしておくことをおすすめする。
- [3] もっとも Unicode に対して冷淡なのはジャストシステムだけではなく、オンラインソフトを含めた日本のソフト業界がそうだといえよう。この原因には、それらのソフトウェアを利用する層が Unicode の多言語それ自体を必要としない点があるのだろうが、何年か前の神学論争的な Unicode 批判もその背景にあるのかもしれない。
- [4] <http://www.karitsu.org/tools/zidian.htm>
- [5] <http://www.justsystem.co.jp/news/2003f/news/j05151.html>
- [6] また、何年か前に ATOK の中国への対応というニュー

スが流れたが、これについては現在どうなっているの

だろうか。こちらも興味深く見守っている。

EGWORD

二階堂 善弘

■ エルゴソフトの EGWORD13

エルゴソフト (<http://www.ergo.co.jp/>) の EGWORD は Mac OS 用の多言語対応ワープロである。ワープロだけでなく、優れた IME である EGBRIDGE も発売している。実際に EGWORD を使いこなすには、他の IME よりも EGBRIDGE を使った方が、相性がよいと思われる。もっとも最新版の EGWORD13 を購入すると、ちゃんと EGBRIDGE14 が付いてくるので、この点は安心であろう。

EGWORD は、現在のところ Mac OS X (或いは以降のバージョン) にもっとも適応したワープロと考えられる。マイクロソフト社の MS Office v. X for Mac も、実際には OS X には不十分な対応しかしていない。特に、Unicode を使った処理を行うと、その不完全さが浮き彫りになる。他の Unicode 対応アプリからのデータの受け渡しが出来ないのである。

ところで Mac OS X に対応した多言語ワープロとしては、ナイスライター (Nisus Writer) が有名であるが、その最新バージョン Nisus Writer Express 1.0 については、現時点 (2003 年 8 月) ではまだ日本語版が出ておらず、試していない。

このため、現在一番 Mac OS X の属性を活かせるワープロは、EGWORD ということになると思われる。

■ 優れた Unicode 対応

EGWORD と EGBRIDGE の特色としては、Mac OS X へのフル対応が挙げられる。このため、Unicode のバージョン 3.2 以降の拡張領域の文字を扱うことも可能である。

また、中国語などの多言語にも対応しているため、Mac OS に付属の IME を使用して、多くの言語を混在することが可能である。MS

Office for Mac が、そもそも UCS-2 の領域すら不完全にしか扱えないのに比べると、EGWORD の方がはるかに勝っている。

また、IME である EGBRIDGE も、Unicode への対応が進んでいる。この点では、ジャストシステム社の ATOK 15 for Mac OS X や、Mac OS 付属の文字パレットよりも優れている面がある。特に、Unicode 拡張漢字の空き領域を明示的に表示できるということでは、Windows 系の IME を含めて、唯一のものであると言ってもよい。

但し、これは EGBRIDGE13 にて行っている。新バージョン 14 だと、Windows 系のフォントの表示には問題があるのか、うまくいかない。

EGBRIDGE で拡張漢字 B を表示



❖ IME と中文ツール、フォント

cWnn (楽々中国語)

二階堂 善弘

□ 総合ツールとしての「楽々中国語」

実のところ、Windows 2000/XP には、中国語のフォントや IME が標準で備わっている。単に入力するだけなら、特に別のソフトを購入する必要は無い。そういった意味からすれば、以前と異なり、Chinese Writer や cWnn といった、中国語 IME ツールを導入する必要性は少なくなっているかもしれない。

しかし、一方で Windows XP に付属する「微軟拼音輸入法」などは、日本人が使うには不便に感じられる面もある。例えば、文字変換の場合、1 度目がスペースキーで、2 度目に候補を選択する時は矢印キーで行うことなどである。これも慣れればなんともないが、中国語入力を始めて習う者などにとってはとっつきにくい面があるだろう。cWnn などであれば、普通にスペースキーを押し続けるだけで済むので、やはり学習が容易である。

オムロンソフトウェアの「楽々中国語 v.2.0」が、ますます中国語総合ツールとしての性格を強くしているのには、そういった背景もあるかもしれない。

この「楽々中国語」は、有名な IME ツールである cWnn6 の他、多言語エディタソフトの「楽々エディタ」、マウスを当てるだけで意味を表示する「辞書ブラウザ」、中国語学習ソフトである「陳さんのやさしい中国語教室」など、幾つかのソフトにより構成されている。

このうち、使えるのはなんといっても辞書ブラウザであろう。小学館の『日中辞典』『中日辞典』に収録される語彙を、日本語からでも中国語からでも、どんなに訳してくれる。この機能は文章を入力している最中にも使うことができ、重宝する。ただ、マウス操作による翻訳は実はちょっと使いにくい。コピー＆ペーストで行った方が便利に感じる。

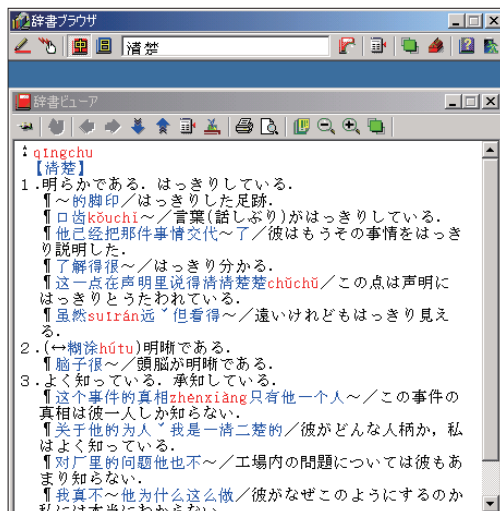
また楽々エディタも、様々な GB/Big5/Unicode を自在に使えるので、使い勝手のよいエディタである。ただ、フォント指定が限られるのが若干気になるのと、Windows 2000/XP のメモ帳が高機能になっている現在では、エディタとしてよりもコンバータとしての役割が強いように思える。

この楽々中国語を全体として見た場合、やはり「単なる IME」ではなく、日本人が中国語を総合的に使いこなす機能を備えているのが特徴と言えるだろう。特に大学などの中国語学習の現場においては、オススメできるソフトではないだろうか。

□ cWnn6

むろん、こういった総合性を備えるとはいえ、その中心となるのはやはり cWnn6 である。その変換のしやすさは定評がある。Unicode への柔軟な対応により、Word や Outlook など、様々なソフトで中国語を入力

辞書ブラウザと辞書ビューア



することができる。その語彙の豊富さは言うまでもない。ほとんど気の向くままに文章を打っていく形でかなり正確な入力が可能である。また、使いこなすほど変換効率がよくなっていくが、これは Wnn の「入力予測」というプログラムによるものである。

むろん、簡体字を入力する場合は GB 系に、繁体字を入力する場合は Big5 系に切り換える必要がある。

しかし、GB 系については、GB2312 だけではなく、語彙面からいっても、GBK にももう少しシフトしてもよいのではないだろうか。或いはこれだけ Unicode 系が広まっている現在、GB 系 Big5 系とせず、一つのバーで、簡体字と繁体字が切り換えられるとなると、もっと便利かもしれない。それから、Big5 の辞書がいまだにちょっと弱いのが気になる。

Chinese Writer 7

二階堂 善弘

□ 大きな変化

日本における中国語処理ソフトの定番といえば、オムロンソフトウェアの cWnn と、高電社の Chinese Writer の両者であろう。それぞれ、これまで幾多のバージョンアップを重ねて、機能を向上させてきた。

しかし、今回の Chinese Writer のバージョンアップは、これまでのものとは一線を画すほど大きな変化であると考えられる。これは Unicode に対応した V4 の時をも上回る変化かもしれない。

また、このバージョンから呼び名も変わり、「Chinese Writer 7」という名称になった。それまでの「V4」「V5」という表現に比べ、よりシンプルなものになっている。

□ GB18030 への対応

まず特筆すべきは、何と言っても GB18030 への対応であろう。これまで、日本の中国語処理ソフトが対応している文字コードは、ずっと GB2312 と Big5 の

4 バイトコード (拡張漢字 A) を入力する



みであった。Unicode に対応していたとはいえ、「入力可能」な文字数は、これらのコードの範囲を出るものではなかったのである。

その間に、中国の IME は、GBK への対応を行い、入力可能な文字数や語彙数を増やしていった。しかし今回の改良により、Chinese Writer は一気にそれらの IME を凌駕することとなった。

とはいえ、現時点で GB18030 に定義されている漢字数は約 2 万 7 千であり、実際に Chinese Writer 7 で使える文字数もその数となっている。事実上、この漢字は、Unicode の「UCS-2 プラス CJK 拡張漢字 A」に等しい。むろん GB18030 対応の明朝体とゴシックの 2 書体のフォントが付属している。

ただ、多くの漢字は「コード入力 (文字パレット)」からしか入力できない。とはいえ、単語登録は可能であり、これにより拡張漢字 A の使用はかなり楽になった面がある。また、漢字の増加にともなって、辞書の語彙数もかなり増加している。GB 入力には、「簡体字・繁体字」切り換えの機能もついた。

日本語入力には、例によって「CW」形式がデフォルトでセットされる。しかし、現在では CW のような独自形式を使うニーズは少なくなっているはずである。選択してインストールする方がよいのではないか。

□ 付属ツールの充実

Chinese Writer 7 においては、V6 よりも大幅にツールの増強が行われている。「日中・中日辞書」や「ビジネス文例」「チャイニーズボイス」「中国語手書き入力」「日中翻訳 C-TIME」などは変わらないが、この他に多くのソフトが追加されている。全部で 44 書体の

フォントも付属する。

特に充実したのは、ピンイン関連のツールである。「漢字ピンイン変換」を使えば、任意の中国語の文章を、自動的に声調付きピンインデータに変換することが可能だ。これはマイクロソフトの Proofing Tools などにも似た機能を持っているが、Chinese Writer の方はルビで行うのではなく、完全に別データとして処理する。

また「ピンイン付きフォント」が付属している。これはダイナコムウェアに同趣旨の製品があった。フォントを変えるだけで、ピンインのルビが振られると同じような効果があるわけである。

また中国語のテキストファイルを読み上げる「中文ファイルプレーヤー」も付属するようになった。中国語教育の面ではかなり応用が可能であろう。

「ファイルコンバータ」で処理可能な文字コードに、

GB18030 が加わった面も大きい。

□ 今回はバージョンアップすべき？

細かい面での改善が行われている点も好感が持てる。たとえば、以前から批判していた、繁体字辞書に「皇后」が出てきてしまう問題も、今度の版でようやく解決されている。

むろん、拡張漢字 A の漢字が部首引きで出てこなかったり、音声データがふくまれていなかったりと、GB18030 に完全に対応とは言い難い面もある。しかし、これだけの改良がなされている点からすれば、これまで使ってきたユーザは是非バージョンアップすべきであろう。むろん新たなユーザへの購入もオススメである。

フォント

千田 大介

□ ダイナコムウェア『中文字型』

台湾に本拠を置く有力フォントベンダー、ダイナコムウェア^[1]は、中国語(簡体字・繁体字)TrueType フォント集『中文字型』を発売した。簡体字フォント・繁体字フォント各 26 書体を収録したもので、Windows 版と Mac 版とがあり、定価はいずれも 24,800 円、実売価格は一万円台後半である。

日本国内で入手可能な中国語フォント製品は、数が非常に限られている。MS Office Proofing Tools に収録されるものが、価格・内容の面で最も充実しているが、しかし、繁体字 Big5 のフォントは収録されていない。そのため、筆者などは台湾でフォントを買ってきて、それを日本語 Windows にインストールして利用していたが、フォントの英文名称が設定されておらず日本語環境では使えないものも多かった。この点、『中文字型』はダイナコムウェア社が中国・台湾向けに発売したフォントをセレクトしてパッケージしたものではあるが、すべてのフォントに英文名称が設定されているおかげで、日本語環境でも問題なく利用できる。また収録書体には同社の日本語フォントと共通するものが多く、日中混在文書のデザインなどでさまざ

まな活用が可能となろう。

しかし、不満な点もある。簡体字フォントは GB2312 対応で 7,000 字程度しか利用できない。せめて基本的ないくつかの書体は、GBK もしくは GB18030 対応のものを収録してほしい。繁体字フォントも Big5 対応であるが、これも同社の製品である MingLiU と同様の Unicode 対応版がほしいところである。

ダイナフォント製品の多くは、Unicode 対応フォン

『中文字型』中国語フォント例

一马离了西凉界
一马离了西凉界
一马离了西凉界
一马离了西凉界
勸千歲殺字休出口
勸千歲殺字休出口
勸千歲殺字休出口
勸千歲殺字休出口

トなど附録が充実しており、それが多漢字処理を必要とするユーザーにとってはたまらない魅力となっていたが、『中文字型』には附録フォントが一切収録されていない。従来の製品で附録扱いだった中国語フォントの独立したパッケージであるから、それも仕方のないところであろう。

細かい注文はあるにせよ、中国語を使った文書・ポスターなどをデザインする必要のある人にとって、『中文字型』は最良の選択肢となろう。同社のフォント製品としては若干高め価格だが、それだけの価値は十分にあると言える。

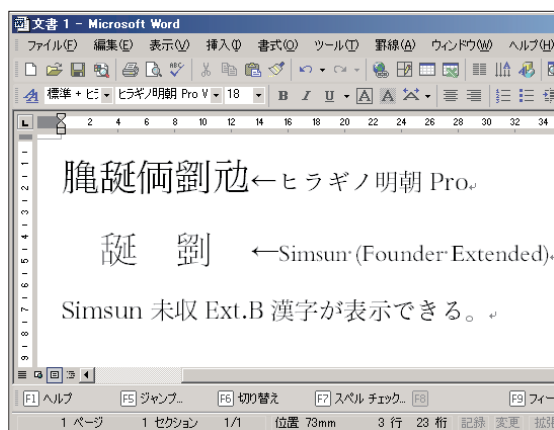
□ ヒラギノフォント

Mac OS X の標準フォントに採用されている大日本スクリーン社^[2]のOpenType フォント、ヒラギノのWindows 版が発売されている。従来の Windows 向け OpenType フォントは Adobe-Japan1-4 対応のものしかなかったが、ヒラギノは Adobe-Japan1-5 に対応しており、これで JIS X 0213 のグリフを含む 20,317 字を、Windows 環境でも、Mac OS からフォントを移植するというライセンス違反を犯すことなく利用できるようになった。

OpenType フォントの機能を使いこなせる Windows ソフトが InDesign のみであるという状況は変わらない。しかし、JIS X 0213 収録文字の多くは Unicode にマッピングされているため、MS Office XP でも利用することができる。しかし Office XP の「記号と特殊文字」では OpenType フォントを上手く表示できないので、JIS X 0213 にしか収録されない文字を入力するには Word で、その文字に対応する Unicode のコード番号を入力→Alt+X で文字に変換→フォントをヒラギノに変更、という手順を踏むしかない。一度入力してしまえば、それを他の Office アプリにコピーして利用することもできる。

問題はその価格である。2003 年 7 月に発売された、Mac OS X 搭載フォントと同等の 6 書体パックは、相当安くなったとはいえ定価 59,800 円もする。前号で紹介したモリサワフォントと比べると、収録文字数の違いはあるものの、割高感は否めない。

Windows 版ヒラギノフォントの主なターゲットは、Mac 環境とのフォントの共有化が必要な印刷事業者であると思われる。しかし以上のように、Windows



Word でヒラギノ明朝を使う

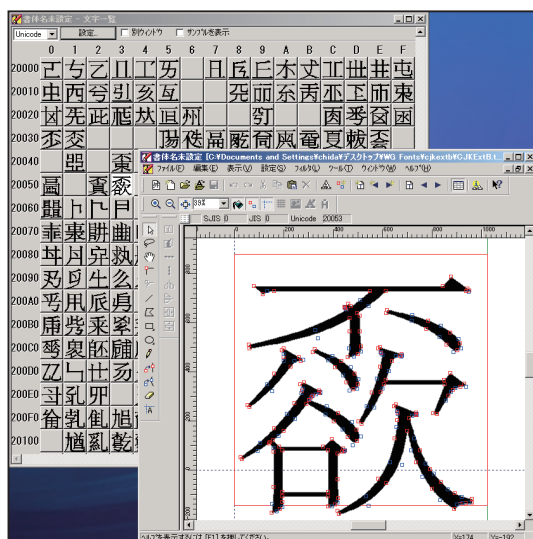
上で JIS X 0213 収録文字をいち早く利用したいユーザーにとっても、十分購入の価値はあるといえよう。

□ TEdit

TEdit は武蔵オンラインソフト^[3]が提供する、シェアウェア（3,000 円）の TrueType フォント編集ソフトである。安価ながら非常に多機能で、Windows 環境でのフォント定番フォント作成ツールになっている。

TEdit は、従来から CJK 統合漢字、約 21,000 字に対応していたが、この八月に Unicode の Extension B

TEdit で Ext.B を編集 (SimSun (Founder Extended) を試験的にインポートして表示)



領域をも編集できるα版が公開された。対応文字数は58,227字、内訳はCJK 統合漢字のうちJIS X 0208 + 0212 相当の約14,000字と、Ext.B 領域の約43,000字である。

しかし、現在のα版ではJIS 拡張漢字未収録のCJK 統合漢字およびExtension A 領域の漢字を編集することができない。A 領域を飛び越してB 領域だけ、というのはいかにも片手落ちなので、すべてのUnicode

の漢字編集ができるように改良を進めてもらいたい。

注

- [1] <http://www.dynacw.co.jp/>
- [2] <http://www.screen.co.jp/>
- [3] <http://www.interq.or.jp/www1/anzawa/>

翻訳ツール

千田 大介

□ はじめに

『漢情研』第2号から毎号、機械翻訳ツールのレビューを執筆している。これまでは、Web 翻訳およびパッケージソフトのうち比較的安価なものを中心に試験してきたが、2003 年前半に代表的な日中翻訳パッケージソフトである蓬萊とj・北京が相次いでバージョンアップした。また、インターネット上で無償翻訳サービスを提供するサイトも現れている。

そこで今年は、それらのパッケージソフトと無償オンライン翻訳サービスの機能を検証する。

■ 検証対象

● 蓬萊

クロスランゲージ社 (<http://www.crosslanguage.co.jp/>) の製品。2003 年3月に発売された最新のバージョン4 (パッケージ名は蓬萊II) には、蓬萊のみのパッケージ (48,000 円) と、中国語 OCR ソフト「Asian

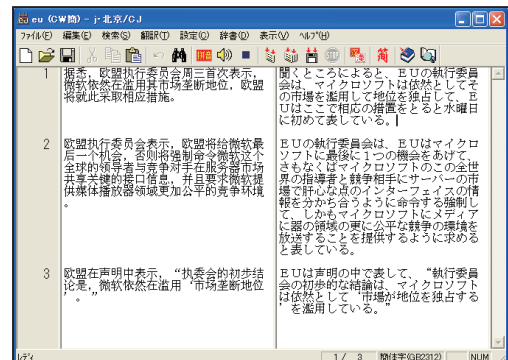
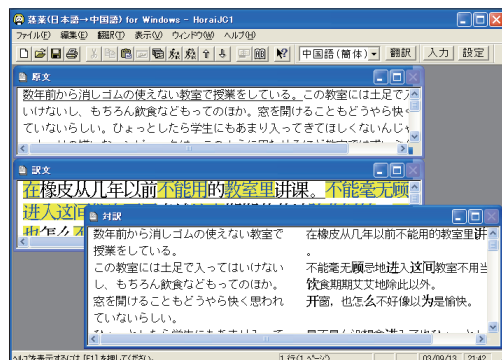
Reader」・中国語 IME「ロボワード」をバンドルした「フルパック」 (58,000 円) とがある。

日中・中日の双方向翻訳が可能。Unicode に対応しているため、中国語・日本語のIME による入力やコピー & ペーストも簡単である。

テキスト翻訳の他、Internet Explorer のアドイン機能を持っており Web 翻訳もできる。テキスト翻訳では、全文を一括翻訳するだけでなく、原文に関する質問に答えながら翻訳をすすめる対話翻訳も可能。簡体字・繁体字に対応し、訳文の文体設定、ユーザー辞書機能、翻訳テンプレートなども備える。インターフェースでは、原文ウインドウ、翻訳結果ウインドウ以外に、原文と翻訳結果を左右対照表示する対訳ウインドウ、以上三種類のウインドウが用意されるのが特色である。

● j・北京

Chinese Writer で知られる高電社 (<http://www.kodensha.jp/>) の日中・中日翻訳ソフト。最新バージョンは2003 年4月発売のV.4。こちらもOCR「China



Scan 2」を含む「j・北京 V 4+ 中国語 OCR」(68,000 円、アカデミック 45,000 円)と、単体パッケージ (58,000 円、アカデミック 38,000 円) がある。

翻訳ソフト本体は翻訳サーバと呼ばれ、左右に日本語・中国語テキストが一文ずつ対照表示される設計になっている。また、j・北京コントローラというツールバーを起動させておくと、クリップボードのテキストをクリック一つで翻訳することができて便利である。ユーザー辞書や Web ページ翻訳のほか、Word・Excel・Outlook のアドイン機能もある。このほか、簡体字・繁体字 IME、中国語のピンイン変換、読み上げ、日本漢字と簡体字の変換など、中国語を扱うためのツール類も充実している。

● Excite 翻訳

<http://www.excite.co.jp/world/>

日中・中日のテキスト翻訳、URL 指定による Web ページ翻訳機能を提供する。翻訳エンジンは、高電社のものを使用。文字数制限については明記されていない。翻訳結果は j・北京と一致した。

● infoseek マルチ翻訳

[http://www.infoseek.co.jp/Honyaku?pg=](http://www.infoseek.co.jp/Honyaku?pg=honyaku_top.html&svx=100302&svp=SEEK)

[honyaku_top.html&svx=100302&svp=SEEK](http://www.infoseek.co.jp/Honyaku?pg=honyaku_top.html&svx=100302&svp=SEEK)

こちらも、日中・中日のテキスト翻訳 (全角 2,000 字以内)、URL を指定しての Web ページ翻訳に対応

している。AMIKAI の翻訳エンジンを使用。

■ 検証方法

全てのツールに、同じ例文を読み込ませた。パッケージソフトに付属する訳語の調整機能や対話翻訳は一切使用せず、翻訳された結果そのままを対象とした。また、例文も短くさほど吟味されてもいないので、以下の検証はあくまでも一つの参考として読んで頂きたい。

なお、j・北京と infoseek では、翻訳が一文ごとに行き改行されて出力されたが、引用にあたって整形した。

□ 中→日翻訳

■ 試験 1 教科書

まず、毎年のレビューで使用している、中国語教科書『話す中国語・北京編』本文の一節で実験した。

● 評価

j・北京の健闘が目立つ。ただ、「一片」「英法联军」は上手く訳せていない。最後の一センテンスも「为了」のかかる範囲を誤認するとともに使役が適切に訳されていない。

蓬萊は、使役文を平叙文に変換して訳出しているのが面白い。「湖已经修复」も受け身で訳せている。ただ、その後のフレーズの連動文は上手く認識できていない。

■ 試験 1

● 原文

圓明園像一片廢墟。这儿的建筑在鸦片战争的时候被英法联军破坏了。许多文物也被偷走了。现在湖已经修复，建筑当作遗址保存，为了使人们不忘记历史。

● 訳文例

圓明園は一面、廢墟のようだ。ここの建物はアヘン戦争のときに英仏連合軍によって破壊された。多くの文化財も盗み去られた。現在、湖は既に修復されているが、建物は遺跡として保存されている。人々に歴史を忘れさせないために。

● 蓬萊

圓明園は一面の廢墟のようです。こちらの閣はアヘン戦争に英法連合軍が破壊した。多くの文化財は同様に掠めて行った。現在の湖はすでに改修されて、遺跡とみなして保存することを建築して、人々を使うため歴史を忘れない。

● j・北京、Excite

圓明園は 1 片の廢墟に似ている。ここの建物はアヘン戦争時に英の仏連合軍に破壊された。多くの文化財も盗んでいかれた。今湖はすでに修復して、建物は遺跡の保存と見なして、人々を使用するために歴史を忘れない。

● infoseek

圓明園は一面廢墟の似です。この建築は阿片戦争の時にイギリスの方法の連合軍に破壊された。たくさん文物も盗みとられた。現在湖はもう修復する、建築が遺跡をとみなすのは保存する、人々が歴史を忘れなくてことために。

最後のフレーズでは、j・北京同様の誤りを犯している。

infoseek は、前の二者と比べると、明らかにシステムの性能が落ちる。出力する日本語が「一面廃墟の似です」「人々が歴史を忘れてくことために」のよう

に壊れている箇所すらある。ただ、最後のフレーズは、「为了」が文末までかかることを、一応正確に認識しているようであるし、その他のフレーズも文法構造はそれなりに把握できているようである。

■ 試験 2

● 原文

①据悉，欧盟执行委员会周三首次表示，微软依然在滥用其市场垄断地位，欧盟将就此采取相应措施。②欧盟执行委员会表示，欧盟将给微软最后一个机会，否则将强制命令微软这个全球的领导者与竞争对手在服务器市场共享关键的接口信息，并且要求微软提供媒体播放器领域更加公平的竞争环境。③欧盟在声明中表示，“执委会的初步结论是，微软依然在滥用‘市场垄断地位’。”

● 翻訳例

①聞くとところによると、EU 執行委員会は、マイクロソフトが依然として市場独占的地位を濫用しており、EU はそれに対して相応の措置をとるであろうと、水曜日にはじめて表明した。②EU 執行委員会は、EU がマイクロソフトに最後の機会を与えるか、そうでなければ地球規模の指導者であるマイクロソフトと競争相手とが、サーバ市場における核心的なインターフェース情報を共有するように強制的に命令するとともに、マイクロソフトがメディアプレーヤー分野においてさらに公平な競争環境を提供するように要求するだろうと表明した。③EU は声明の中で次のように表明している。「執行委員会の初歩的な結論は、マイクロソフトは依然として『市場独占的地位』を濫用しているということだ。」

● 蓬萊

①聞くとところによれば、ヨーロッパ連合執行委員会周三に初回示して、マイクロソフトは依然としてその市場が地位を独占することを乱用して、ヨーロッパ連合はすぐにこれで銘々の措置を取る。②ヨーロッパ連合執行委員会は示して、ヨーロッパ連合はすぐにマイクロソフトに与える最後に 1 個の機会、でなければすぐに強行して命令するマイクロソフトはこの世界的な幹部と競争の相手はサーバー市場に共に享受して肝心であり他人の話を受けて話して情報は、しかもマイクロソフトに媒体放送器領域一層公平な競争環境を提供することを要求する。③ヨーロッパ連合は声明に示して、“委初歩の結論が、マイクロソフトが依然として‘市場は地位を独占する’を乱用していることであることを執る。”

● j・北京、Excite

①聞くとところによると、EU の執行委員会は、マイクロソフトは依然としてその市場を濫用して地位を独占して、EU はここで相応の措置をとると水曜日に初めて表している。②EU の執行委員会は、EU はマイクロソフトに最後に 1 つの機会をあげて、さもなくばマイクロソフトのこの全世界の指導者と競争相手にサーバーの市場で肝心な点のインターフェースの情報を分かち合うように命令する（ことを）強制して、しかもマイクロソフトにメディアに器の領域の更に公平な競争の環境を放送することを提供するように求めると表している。③EU は声明の中で表して、“執行委員会の初歩的な結論は、マイクロソフトは依然として‘市場が地位を独占する’を濫用している。”

● infoseek

①聞かように、ヨーロッパ盟の執行委員会ひと回りする三は初めて表す、(小さくて軟らかくて依然に(で)濫用する其れの市場独占する地位)、ヨーロッパの盟はこれで措置をとり相応する。②ヨーロッパの盟の執行の委員会は表す、ヨーロッパの盟は小さくて軟らかい最後のに一つ機会にあげる、さもないと命令の小さくて軟らかいこの全世界の指導者と競争を相手がサービス器の市場に(で)肝要のインターフェースの情報を共に享受する強制する、しかも小さくて軟らかく媒体を提供する器具の領域を放送する。更に公平な競争が環境を要求する。③ヨーロッパの盟は声明の中に表す、“執の委員会の初歩の結論はです”、(小さくて軟らかくて依然に(で)濫用する”市場は地位を独占する”)。

■ 試験 2 新聞記事

次は、IT 関連の新聞記事である。なお便宜上、各文に番号を付してある。

● 評価

まず、名詞の訳からみていこう。

原語	蓬萊	j・北京	infoseek
欧盟	ヨーロッパ連合	EU	ヨーロッパ盟
微软	マイクロソフト	マイクロソフト	小さくて軟らかい
服务器	サーバー	サーバー	サービス器
媒体播放器	媒体放送器	——	——
执委会	——	執行委員会	執行(の)委員会

蓬萊と j・北京は「欧盟」「微软」「服务器」を正確に訳しているが、infoseek はいずれも単語として認識していない。このほか、蓬萊と infoseek は「周三」（水曜日）を訳せていない。これらは、現在のニュースを読む上で基本的な語彙であるといえ、今後、辞書の収録語彙を増やして欲しい。

次に文法事項に移ろう。①「在濫用」の「在」を進行として訳せているのは、蓬萊（乱用して）のみである。また、「市場壟断地位」というフレーズは、いずれも正確に把握できていない。多少文言めかした言い回しであるためであろうか。この文については、蓬萊が構文を正確に把握しており、訳語も自然である。

②は、初めのフレーズの「表明」が文末までかかる、という構造は、おおむね把握できているようだ。「否則」以下の部分は、j・北京が的確に構文を把握して訳している。蓬萊では、訳が乱れているが、「关键的接口信息」の「接口」の、「インターフェース」という名詞的用法が辞書に登録されておらず、「共享」の目的語として把握できなかったためであろう。infoseek は構文はほぼ正確に把握できているが、日本語が不自然なのは相変わらずである。

個別に見ていくと、「将给」の「将」は副詞として、漢文の再読文字の「まさに～せんとす」の意味で使われている。蓬萊は「すぐに」と訳せているが、j・北京は処置式の構文に訳している。「微软这个全球的领导者」は、蓬萊「マイクロソフトは」j・北京「マイ

クロソフトの」と訳している。これは、近頃新聞などで増えている英語直訳調の少タイレギュラーな文であるので致し方ないだろう。

③では、j・北京の訳語がもっとも的確である。ただ、「市場壟断地位」はここでも上手く訳せていない。蓬萊は「执委会」という略称を動詞フレーズに誤認したために、括弧内の構文が上手くとれなかった。省略語ではあるが政治や企業の分野でしばしば見られる語なので対応してほしかった。infoseek は、括弧の付き方がおかしく、訳文も日本語になっていない。

■ まとめ

蓬萊、j・北京はいずれ劣らぬ成績であるといえよう。訳文の自然さでは、訳語の選定が的確な j・北京に軍配が上がる。一方、個々の構文の把握能力では蓬萊にも優れた点が認められる。下訳に使えるレベルにまでは達していると言えよう。両者ともに文言的言い回しは苦手としている。文言処理の困難は理解するが、近年、新聞その他で文言的言い回しが多用される傾向があるので、是非とも対応を進めていただきたいところだ。

infoseek は、訳文に日本語としておかしい箇所が多く、まだまだ改善の余地が大きい。構文自体はそれなりに把握できているだけに、辞書の語彙を増やし、自然な日本語が生成されるようにブラッシュアップしていくなど、今後の改善に期待したい。

□ 日→中翻訳

次に、日中翻訳の性能について、検証する。例文は、漢情研メールマガジン第 6 号掲載の平林氏のコラムから引用した。これも便宜上、文番号を付した。

● 評価

蓬萊では、認識できなかった日本語がそのまま出力されている。「ホコリ」というカタカナ書き換え語、「ほしくないんじゃないか」という口語表現はともかく、「ひょっとしたら」「ずいぶん」などを訳せなかったのは残念である。それらの部分を除くと、⑧⑨は比較的良好に訳せている。①は、「数年前から」の掛かる語を誤認している。②③⑦⑩は、副詞の位置・用法、主述構造などがおかしい。

j・北京は、今回試みた中では、もっとも上手く訳せている。語彙では「ドライブ」をスポーツ球技の用

■ 日中翻訳試験

● 原文

①数年前から消しゴムの使えない教室で授業をしている。②この教室には土足で入ってはいけなし、もちろん飲食などもってのほか。③窓を開けることもどうやら快く思われていないらしい。④ひょっとしたら学生にもあまり入ってきてほしくないんじゃないか。⑤ホコリの嫌いなコンピュータは、このように思わせるほど教室ではずいぶんと丁寧に扱われる。

⑥ホコリだけではない。⑦ドライブやキーボードを盗むやつがいるので、休み時間中はしっかり施錠される。⑧学生達は教師が来て鍵を開けてくれるまで、暗い廊下で待たなければならない。⑨しかも太陽光線を避けて日の当たらない場所に作られているので、教室の中もやっぱり薄暗い。⑩内装は総防音壁、声が反響せず、中にいると何だか陰鬱な気持ちになる。

● 蓬萊

①在橡皮从几年以前不能用的教室里讲课。②不能毫无顾忌地进入这间教室不用当然饮食期期艾艾地除此以外。③开窗，也怎么不好像以为是愉快。④是不是没想余进入了也ひょっとしたら学生？⑤ホコリの讨厌の电脑这样让以为是在教室里对はずいぶん敬重地被处理。

⑥ホコリ只不是。⑦继续偷取驱动器和键盘因为在，所以在整休息时间坚固地被锁上。⑧您教师来开钥匙为止学生们必须在暗的走廊上等一会儿。⑨因为而且，避开太阳光线被天不中的地方做，所以教室的里头也仍然昏暗。⑩全防音墙声音不反响了，要是在里头室内装饰就成为忧郁的什么心情。

● j・北京、Excite

①从数年前在橡皮不能使用的教室做着授课。②不可用穿着鞋进入这个教室，当然饮食等荒谬。③开窗的事也好歹没高兴地好象被看作。④也许不太不是也请进入学生吗？⑤尘埃讨厌的计算机，这样使人想到的那样教室相当被郑重其事对待。

⑥不仅仅是尘埃。⑦因为偷旋转球和键盘的东西在，休息时间内好好地被上锁。⑧学生们必须到教师来给开钥匙，在暗的走廊等。⑨因为在而且避开太阳光线日不合适的地方被做（制作），教室中还是也发暗。⑩内部装饰总防音墙，声音不反响，变成在中的话总觉得忧郁的心情。

● infoseek

①再三年从前在教室讲课橡皮的能用有。②这教室不脱鞋进入不许、不用说饮食盛满别的。③窗户打开好歹愉快想到没有。④突然作学生太进入要没有员大蛇内科。⑤餐叉利讨厌计算机、这想到教室太坏郑重处理。

⑥不是仅仅餐叉利。⑦因为有偷驾驶和键盘的家伙，休息时间中央作坚固上锁。⑧学生们教师来在打开钥匙的之前，在黑暗的走廊必须等待。⑨而且避开太阳光线因为在日的不碰的地点制成，教室的中央也仍然发暗。⑩内部设备总防音墙壁、声音反响作、中央有是什么阴沉心情变成。

語と誤認している他は、おおむね問題なく訳している。前半の段落では、②～⑤で日本語の曖昧表現を置き換えた副詞に、例えば④「也许不太不是也」のように不自然さが目立つ。後半は比較的良好に訳しているが、訳語の副詞にとくに不自然な箇所が目立つ。また、⑩では「気持ちになる」の「なる」を、この文全体のように変化する、の意味に取り違えている。

infoseek であるが、文法構造の把握が不十分で、日本語の単語を逐語訳し並べているだけ、といった印象を受ける。

□ 終わりに

■ 確実な性能向上

第2号ではじめて機械翻訳ツールをレビューしたとき、お世辞にも役に立つと思えるものはなかった。しかし今回、二年間で機械翻訳の性能がかなり向上していることが確認できた。下訳に使えるレベルにまで達したと言って良からう。

中日・日中翻訳の成績を総合すると、j・北京が最

も好成績をあげている。附属ツールが充実し、価格も比較的安いので、機械翻訳ソフトを購入する際の有力選択肢となろう。購入を考えている方は、Excite 翻訳であらかじめ性能を検証してみていただきたい。

蓬莱も、中国語の構文の把握など、優れた部分がある。今回利用しなかった対話翻訳機能を利用すれば、より自然な訳文の生成が可能になるものと思われる。

infoseek は、上記二者に比べて、格段に完成度が落ちる。辞書・文法解析アルゴリズムともに、さらなるブラッシュアップが必要である。

■ 応用の可能性

下訳に使えるレベルにまで機械翻訳の能力が向上することで、これらのソフトの実務・教育での応用がいろ

いろと考えられるようになった。

ビジネス・技術文書などを大量に翻訳する必要がある翻訳者は、これらのソフトに習熟することで作業を効率化することができよう。

その一方、中国語教育の現場では、初級レベルの単純な単文翻訳の宿題などについて、学生が無償の Web 翻訳結果をそのまま利用してくるケースが今後爆発的に増加していくことだろう。しかしその際、教員には、そういったシステムをただただ敵視するのではなく、むしろ積極的に利用しつつ教育に取り込む姿勢が求められよう。例えば機械翻訳の結果をいかにブラッシュアップするかという実践的な中国語翻訳の訓練への応用など、あらたな中国語授業の可能性が開けているのだから。

❖ ネットワーク

WWW ブラウザ

秋山 陽一郎

□ 概要 (2002-2003)

この一年は、特に Internet Explorer (以下 IE) 以外のブラウザに特筆すべき変化が相次いだ。まずは今年 7 月に OneStat.com が発表したブラウザシェアの調査結果を昨年のデータとつぎあわせてみよう。^[1]

WWW ブラウザ	2002/6	2003/7
1. Microsoft IE 6.0	46.4%	66.3%
2. Microsoft IE 5.5	23.9%	14.5%
3. Microsoft IE 5.0	23.6%	12.7%
4. Mozilla	0.4%	1.6%
5. Microsoft IE 4.0	1.2%	0.8%
6. Netscape Navigator 4	1.4%	0.6%
7. Opera 6.0	0.6%	0.6%

全体の 95.4% という、独占状態が依然として続いている IE は、昨年 4 月時点でのシェアが 96.6% だったことを考えれば、むしろ他のブラウザの健闘と評価することもできるが、中でも昨年 6 月に v1.0 を正式にリリースした、Mozilla は前年より 1.2 ポイントもの伸びを見せており、後述する Safari (0.14%) も、今後 MacOS X におけるシェアが IE と逆転していくことが見込まれることから、今後の動向が注目される。

以下、2002 下半期～2003 上半期にかけての主要なブラウザについての動向を紹介しつつ、特に昨今、OS レベルでの対応が着実に進んでいる Unicode 3.1 CJK Extension (Ext.A=6,582 字、Ext.B=42,711 字) の対応状況を検証してみたい^[2]。結論からいってしまえば、Ext. A はすでに実用段階にあり、Ext. B はフォント待ちの状態にある。

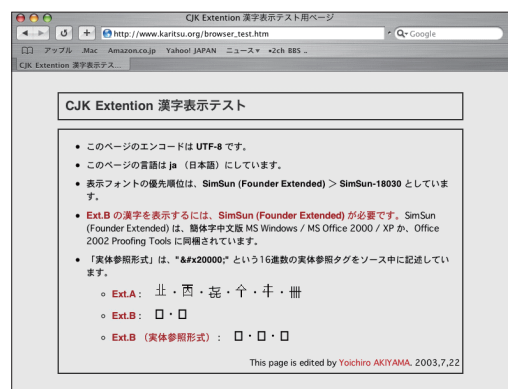
■ Safari (Mac OS X 用)

この一年間のブラウザ業界は、ある意味、Mac OS

X 用の新ブラウザ “Safari” を中心に動いたといっても過言ではないだろう。

今年 1 月に Apple 社が発表し、6 月 24 日に正式に公開された Safari は、その速さが話題を集めた。この Safari は、UNIX の統合デスクトップ環境の KDE (K Desktop Environment) でおなじみの Web ブラウザ Konqueror に採用されている KHTML レンダリングエンジンに基づいており、Apple 社のベンチマークによれば、起動時間で Mac OS 用 IE5 より約 20%、またページ読み込み速度にいたっては 3 倍も速いという^[3]。Microsoft は今年 6 月 13 日に Mac 版 IE の開発停止を表明しており^[4]、すでに Power Mac G5 から、Safari がデフォルトブラウザとして採用されている。なお Safari は当初英語版のみのリリースであったが、すでに Public Beta 2 から日本語インターフェースにも対応している。

さて、この Safari における多言語対応状況だが、残念ながら筆者の手元に Mac 環境がないため、細部にわたるテストはできなかったが、すくなくとも Mac OS X 10.2 以降で Ext. A 漢字が特別な設定を要することなく表示できることが確認できた (上図参照)。また Ext. B の漢字についても、これまでに Microsoft の中文 Windows や 中文 Office、あるいは Office



2002 Proofing Tools などに同梱されている Simsun (Founder Extended) によって表示可能であることが報告されている^[5]。要するに「フォント待ち」の状態にある。

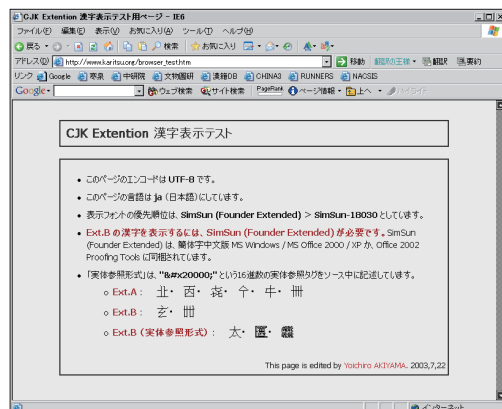
■ Internet Explorer

何かと「緊急」・「重要」なセキュリティパッチのリリースばかりが目立つ一年だった IE だが、大きな動きと言えば、やはり IE 6 Service Pack (以下、SP) 1 を最後にスタンドアローン版を廃止することになった (つまり、今後 IE 7 や IE 6 SP 2 がリリースされることはない) というニュースだろう^[6]。昨年 9 月に、Microsoft 独占禁止法訴訟の和解案に準拠した Windows XP SP 1 を公開して、Windows 上のデフォルト・ブラウザを IE 以外のブラウザに設定できるようになったものの (注: IE のアンインストールはできない)^[7]、次期 Windows (コードネーム: Longhorn) では、IE は OS の一部に完全に組み込まれてしまうことになるという。

IE の動向でもうひとつ注目すべきことは、今年 6 月に発表された Mac OS 用 IE の開発停止である。これまで Windows 版 IE と Mac 版 IE との間のレイアウト差や機能差に対して余分なコストを費やしていた Web サイト制作者は少なくなかったことと思うが、ついにその問題が解決されることがないまま、IE は Mac OS から撤退してしまうことになる。なお Microsoft によれば、旧来の Mac 版 IE 5 のサポートについては当分継続するとのこと^[8]。

多言語面においては、すでに本誌第 2 号にレビューされている通り、Windows 2000/XP + IE 6 上で Ext. B 漢字を表示することは可能だが、Microsoft のサイトから Windows 2000 Surrogate (超大文字集) をダウンロードして導入する必要がある点はいまだに変わらない^[9]。これが未導入だと Ext. A の漢字は表示できるものの、Ext. B 漢字は Windows XP においても文字化けを起こしてしまう。評者の環境 (Windows XP Professional SP 1 + IE 6 SP 1) では、Windows 2000 Surrogate (surreg.exe) を実行するだけで Ext. B 漢字を IE 上で表示できるようになったが、もしこれでも表示されない時は Windows のレジストリエディタ (regedit.exe) を使って、

```
[HKEY_CURRENT_USER\Software\
```



```
Microsoft\Internet Explorer\
International\Scripts\42]
```

というキーの [IEFixedFontName] と [IEPropFontName] という値のデータを [Simsun(Founder Extended)] に設定する。ただし、レジストリの編集はリスクが大きいので、初心者にはおすすりできない。変更は自己責任で慎重に行ってほしい。

■ Netscape / Mozilla (+和ジラ)

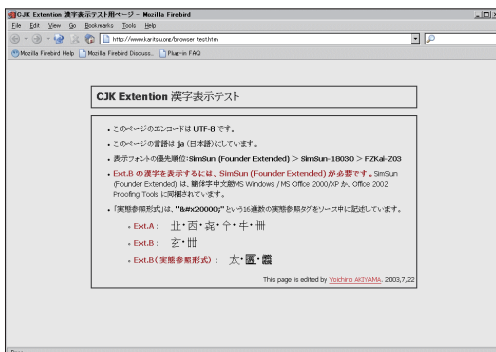
昨年 6 月に Mozilla 1.0 が、また 8 月にはタブ・ブラウジング機能などを新たに搭載した Netscape 7 がそれぞれリリースされた。翌 9 月にリリースされた Windows XP SP 1 によって、Windows から IE を非表示にして、Mozilla や Netscape をデフォルト・ブラウザに設定できるようになったものの、結局、両ブラウザに対する追い風になることはなかった。Mozilla 1.0 の開発が遅れたことに加え、今年に入ってから Safari や Opera 7 の登場によって高速・快適というイメージからすっかり離れてしまった Mozilla は、4 月に発表した新しい開発ロードマップで、今後は軽量化・高速化を図って Phoenix ベースのブラウザへ転換すると発表し^[10]、さらに 7 月には、AOL から Mozilla Foundation が独立するに至った^[11]。

漢字処理面において注目すべきは、和ジラ (Wazilla) の登場である。和ジラは通常の Mozilla (洋ジラ) で放置されている日本語処理関連のバグを修正・補強しているのが特徴で、「gb2312, euc-kr に切り替えるとユーザー一覧が出ない」のようなバグにも対応している。(Mozilla の Ext. B 漢字表示テストの結果は、次の Mozilla Firebird に準じる。)

■ Mozilla Firebird (旧 Phoenix)

昨年9月に公開されたPhoenixは、MozillaやNetscapeよりも高速軽量化をはかるというコンセプトで開発されたブラウザで、Mozillaのユーザーインターフェースを記述するために開発されたXUL(XML-based User Interface Language)というクロスプラットフォーム(OSに依存しない)で拡張やカスタマイズにも柔軟な言語をベースにしている。環境にも左右されるので一概にはいえないが、筆者の手元の環境ではIEよりもわずかに起動が速い。このPhoenixは、その後、その名称がBIOSを製作している米国企業の商標権に抵触するとの理由から、今年5月にMozilla Firebirdと改名された^[12]。

現在のところ、OSではWindows・Linux・Mac OS Xに加えて、非公式ながらSolaris・OS/2・BeOS・AIXにも対応している。ただし、中国語(簡体字・繁体字)や韓国語のローカライズ版が用意されていないが、未だに公式な日本語版はなく^[13]、またMozilla Firebirdのバージョン自体も、8月上旬の時点でまだ0.6の段階なので、修正すべきバグも少なくない。さらにバージョンアップするたびに古いMozilla Firebirdのディレクトリを削除して、プロファイルもその都度新しく作成しなおさなくてはならないため、日本の一般的なユーザーの間に広まるには、まだまだ敷居が高いという感もある。ただ多言語面では、まずGB18030への対応と、Unicode 3.1 Ext. B 漢字の表示はともにフォントさえあれば、特別な設定を施すことなく行え、この手軽さの点では明らかにIEより優れている。また少なくともWindows 2000/XP上では、Mozilla Firebirdの中国語版を日本語ロケールのまま使用できることが確認できている。



■ Opera

最後に、これも今年1月末にv7.00が公開されて一層の高速化を図ったタブ型ブラウザのOperaだが、これはExt. Aは難なく表示できるものの、残念ながら、今のところExt. Bは表示できない(文字化けを起こしてしまう)。アドレスバーからの検索がJIS外字にも対応しているだけに、今後の発展に期待したい^[14]。

注

- [1] http://www.onestat.com/html/aboutus_pressbox23.html
- [2] CJK Extension 漢字表示テストページ
http://www.karitsu.org/browser_test.htm
なおExt. B 漢字の表示には、Simsun (Founder Extended) をインストールしておく必要がある。
- [3] <http://www.apple.co.jp/safari/>
また以下のURLに独自のベンチマーク検証結果を含む、より詳細なレポートがある。
<http://pcweb.mycom.co.jp/news/2003/01/14/33.html>
<http://pcweb.mycom.co.jp/news/2003/01/14/34.html>
- [4] <http://www.zdnet.co.jp/products/0306/14/11.html>
- [5] <http://nika01.hum.ibaraki.ac.jp/~nikaido/kawara005.html>
- [6] <http://japan.cnet.com/news/media/story/0,2000047715,20054861,00.htm>
- [7] <http://www.karitsu.org/news/20020831.htm>
- [8] http://www.zdnet.co.jp/news/0306/14/nebt_11.html
- [9] <http://www.microsoft.com/downloads/release.aspx?ReleaseID=31114>
- [10] <http://jt.mozilla.gr.jp/roadmap.html> (邦訳文)
http://www.zdnet.co.jp/news/0301/15/ne00_mozilla.html
- [11] <http://jt.mozilla.gr.jp/press/mozilla-foundation.html>
- [12] <http://jt.mozilla.gr.jp/projects/firebird/release-notes.html>
<http://slashdot.jp/article.pl?sid=02/11/28/235239&mode=thread>
(師茂樹氏の教示による。)
- [13] 非公式ながら日本語パックが公開されている。
<http://www.geocities.co.jp/SiliconValley-SanJose/9076/>
- [14] 同様の検索機能を持つ、Mozilla系ブラウザは残念ながらJIS外漢字には未対応。

その他

千田 大介

□ メールソフト

■ 増加する多言語対応メールソフト

Windows 環境で使える多言語対応メールソフトの定番といえば、お手軽な Outlook Express、フリーでセキュアな Netscape^[1]、機能を詳細に設定でき通好みの Becky!^[2] の三者であるが、それぞれに特色があり一定数の愛用者がいる。このほか、Outlook・akira22+^[3] などとも一定の支持を得ており、また多言語対応機能も安定している。

しかし、ここ一二年でいくつかの多言語対応メールソフトがデビューしている。以下では、それらのうち2種をとりあげ、多言語対応を中心に検証する。

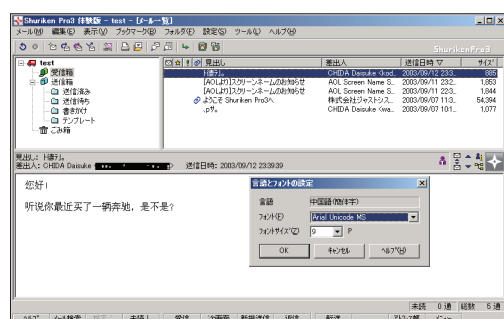
● Shuriken Pro 3

Just System のメールソフト。一太郎のおまけの Shuriken ではない、単体パッケージソフト (4,000 円) である。Just System 社のサイト^[4] で体験版 (30 日間) を配布している。以下では、その体験版を試用した。

Shuriken Pro は、バージョン 3 で多言語に対応した。対応言語は、

日本語、中国語 (簡・繁)、欧州諸語、Unicode (UTF-7、UTF-8)

簡体字中国語は GB2312 と HZ のみの対応。GB18030 や GBK のメールは稀ではあるが、できることならば、今後の対応をお願いしたい。



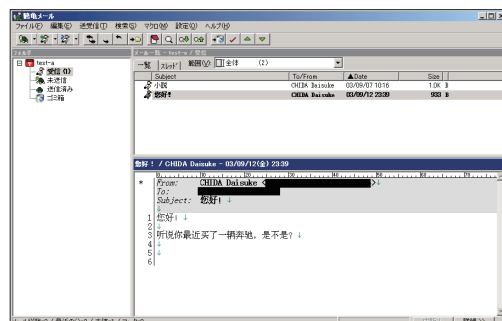
Shuriken Pro 3 の多言語設定方法は独特である。言語とフォントの関連づけが設定されていない言語のメールを開こうとすると、フォント設定のダイアログが開くので、そこで対応フォントを設定する。Outlook Express のような自動設定ではないものの、初心者にもわかりやすいなかなかに考えられた設計である。ただ、メール作成画面では、フォントを設定しなくても文字コードを切り替えるだけで、日本語以外のメールも作成できた。エディタとビューワとが別個の設計となっているのであろうが、設定は共通化してもらいたい。

問題点としては、メールのサブジェクトが多言語未対応であることが挙げられる。このため、中国語のサブジェクトは文字化けを解消できないし、サブジェクトに日本語以外を入力する場合には、メール編集画面に一度文字を入力してカット＆ペーストすることになる。これは非常に使いづらいので、今後の改善に期待したい。

全体としては、ユーザーインターフェースもこなれており、高いレベルでまとまったメールソフトである。今後、以上に紹介した問題点が改善さて、有力な非 Outlook 系多言語メールソフトへと成長することが期待される。

● 鶴亀メール^[5]

鶴亀メールは、日本語テキストエディタとして広い支持を集める秀丸エディタの兄弟ソフトにあたり、エディタには秀丸がそのまま使われている。シェアウェア (2,000 円 + 税) だが、秀丸エディタ登録ユーザーは無料で利用できる。執筆時点の最新バージョンは



3.06。

鶴亀メールは日本語の他、以下の言語・文字コードにも対応している。

欧文、簡体字中国語、繁体字中国語、韓国語、ロシア語

言語とフォントの関連づけ設定の必要が無いことは、鶴亀メールの大きな特徴である。あらかじめ Windows に各言語のフォントが組み込まれていれば、メールソフトがエンコードとフォントの関連づけを自動で設定してくれる。もっとも、返信メールが元メールの言語を引き継ぐ設定はデフォルトでは選択されていない点、注意が必要である。

気になるのは、UTF-8 などの Unicode 系エンコーディングに対応していない点である。多漢字・多言語混在メールの送受信や、Outlook Express などのように、UTF-8 を自動選択してメール送信する機能を持つメールソフトがあることを考えると、是非とも対応してもらいたいところである。また、HZ や GB18030 にも対応していない。

総じて、多少の不満はあるものの十分に実用的な多機能・多言語メールソフトに仕上がっている。特に秀丸エディタ愛用者にとっては、操作フィールが近くマクロも使えるなど、お勧めのメールソフトであると言える。

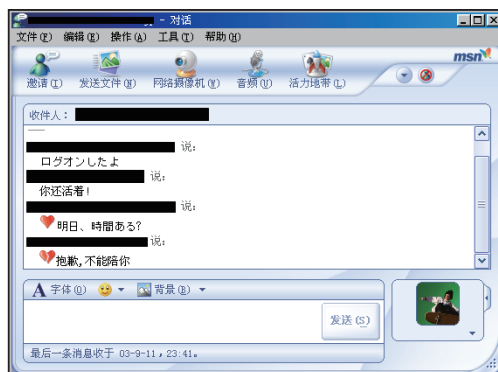
□ インスタントメッセンジャー

■ 主要製品の多言語対応

インスタントメッセンジャー（以下 IM）については、第 2 号にレビューが掲載されているが、その後状況は大きく変化している。アジア地域のブロードバンドユーザーは大幅に増加し、それにともない主要 IM は日本語・中国語などのアジアの各言語版を提供するようになったのだ。

たとえば、IM の草分け ICQ^[6] は、以前は英語版しかなく、日本語化パッチを当てるなどして利用されていたが、現在では日本語・中国語（簡・繁）などの各国語バージョンが提供されるようになっており、アジア言語環境でも手軽に利用できる。

ICQ の他の有力な IM、Windows Messenger / MSN



MSN Messenger 6.0 簡体字中国語版

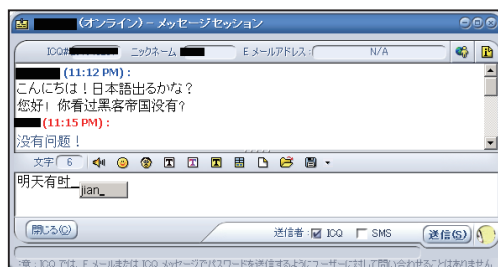
Messenger^[7]、AOL Messenger^[8] などは、いずれも各ローカル言語に対応したバージョンが出ている。いずれも無償配布されている。

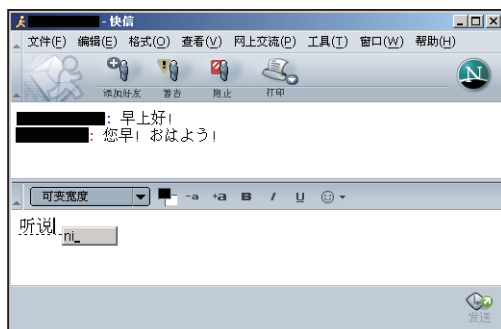
以上三種の IM について、日本語版 Windows XP Pro. と簡体字中国語版 Windows 98 (IE 日本語サポート適用済み) との間でチャットをテストしてみた。テストにあたっては、それぞれの言語版の以下のバージョンの IM を利用した。

		ICQ Lite	AOL
日本語	Windows Messenger 4.7	Build 1105	AIM 5.1.3036 / Netscape 7.1
中国語	MSN Messenger 6.0	Build 1102	Netscape 7.0

以上のうち、AIM 以外の各ソフトは、日中いずれの Windows でも問題なく日本語や中国語、あるいは両者が混在したショートメッセージのやりとりができた。AIM 日本語版は、内部処理が Unicode 化されていないようで、中国語を入力すると「?」化けしてしまい多言語メッセージは扱えなかった。しかし、Netscape 7.x が AOL と ICQ とに対応した IM を提供しているので、それを利用すれば AOL のサービスを

ICQ Lite 日本語版





Netscape 7.0 簡体字中国語版で AOL Instant Messenger を利用

利用して、問題なく多言語メッセージのやり取りができる。

多言語機能は大差ないものの、細かい使用感には若干の相違が見られた。例えば、ICQ だけは Windows 98 環境での Global IME による入力に対応していない。もっとも、市販・フリーの IME もいろいろとあるので、それらを利用すれば問題なろう。

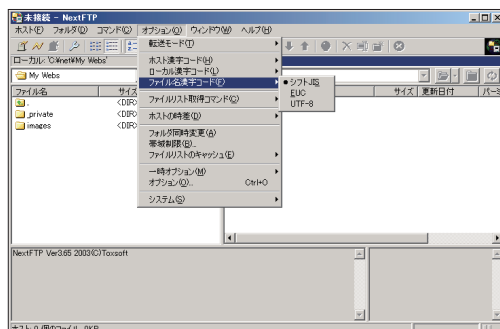
また、多言語混在メッセージの表示は、Windows Messenger / MSN Messenger では TrueType フォントで表示され、表示文字サイズがやや小さめであることもあり、若干見づらく感じるがあったが、Netscape・ICQ は比較的大きな字のビットマップで表示され、見やすかった。いずれも、日中混在メッセージの表示には、GBK 系統のフォントが自動選択されていた。

これらの点は、実用上はさして問題ならない微差であり、上記三種のサービスとそれぞれの IM の多言語対応は、甲乙付けがたい高いレベルにあるといえる。むしろ、ビデオ・音声チャットやファイル転送などの付加機能、知人がどのサービスを利用しているか、などの視点から IM を選ぶべきであろう。

□ FTP ソフト

■ FTP ソフトの多言語対応

従来、Web では英数ファイル名を利用するのが原則であったこともあり、FTP ソフトの多言語対応は進んでいなかった。しかし、中国の FTP サーバから GB コードファイル名のファイルを落とす、というようなニーズもあるし、近年の Unicode 環境の普及により、



Next FTP

例えば UTF-8 で命名した漢字ファイルを Perl で操作するというようなこれまでになかった使い方が生まれてきている。そもそも Mac OS X サーバではファイル名は全て UTF-8 である。

従って、Windows 環境の FTP クライアントでも多言語対応の必要とニーズとは高まってきているのではなかろうか。

● Next FTP^[9]

使いやすい GUI への評価が高い Next FTP は、バージョン 3.5 以降で、従来の日本語に加えて UTF-8 など Unicode 系の漢字コードにも対応した。アップロード・ダウンロード時に、Shift JIS・EUC・UTF-8 などのテキストを自動でコードコンバートすることができるし、ファイル名もそれらの文字コードに対応している。

他の FTP クライアントについては情報が見あたらないが、是非とも Unicode と多言語への対応を進めて頂きたい。

注

- [1] <http://www.netscape.co.jp/>
- [2] <http://www.rimarts.co.jp/>。最新バージョンでは、サブジェクトへの中国語等の直接入力が可能になっている。
- [3] <http://www.tg.rim.or.jp/~khf07113/>
- [4] <http://www.justsystem.co.jp/>
- [5] <http://hide.maruo.co.jp/>
- [6] <http://www.icq.com/>
- [7] <http://www.microsoft.com/>
- [8] <http://www.aol.com/>
- [9] <http://www.toxsoft.com/nextftp/>

❖その他

テキストエディタ

山田 崇仁

□ Unicode 対応から見た一押しエディタ

本誌第二号に師茂樹氏による「エディタの Unicode 対応」という記事が掲載されている。それからまる二年が経過したわけが、インターネット上では日々新たなエディタが公開され、また既存のエディタソフトも更なる改善をとげてきている。

エディタ（ここではテキストエディタと同義語）はそもそも文字コードのみで構築されたテキストデータを専門に取り扱うソフトであり、エディタの基本的枠組みは文字コードにより規定される。そのため、エディタと文字コードとは切っても切れない関係にあるといえるだろう。

現在の国際的的文字コードの主流が Unicode である事は紛れもない事実であり^[1]、これは今後も変わる事がないだろう。そこで、ここでは多漢字を中心とした Unicode 対応度という面からいくつかのエディタを比較してみることにした。

まず評価用のテキストデータは自作の『論語』全文テキストデータを使用した。『論語』は Unicode の CJK 統合漢字収録文字のみで入力可能だが、その中には「八佾篇第三」のように JIS 第一・第二水準に含まれない文字も多く存在する^[2]。一口に Unicode といっ

ても、実際には複数の符号化方法があるので、ここではよく使われている UCS-2 と UTF-8 の二つの符号化方法で保存テキストファイルを対象とした。

エディタのチェックポイントとして、以下の二つを設定した。

- ① Unicode で保存されたテキストファイルの読み書き+表示+入力に対応しているか
- ② JIS 第一・第二水準に含まれない文字の検索・置換・grep に対応しているか

これは評者のエディタに対する利用目的を念頭に設定したものである。無論読者の中には他の利用目的からのチェックを主張される向きもあるだろうが、百花繚乱の観のあるエディタの中から、自分好みのそれを選ぶ際の参考にいただければ幸いである。

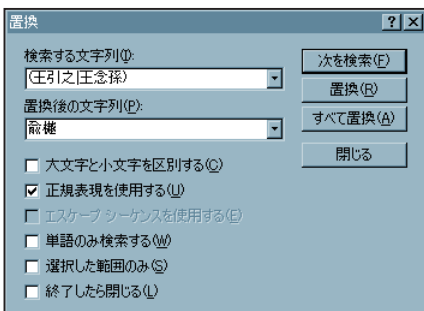
■ Unicode データの読み書きについて—— JIS 相当部分しか対応していないのがほとんど——

まず一つめのポイントである。Unicode で保存されたテキストの読み書きに対応するエディタは少なからず存在する。著名な秀丸・EmEditor・xyzzы・Aprotool TM Editor・MM エディタ・akira22++・Qx エディタなどは、皆 Unicode ファイルの読み書きに対応している。

しかし、実際には「Unicode のうち JIS と重複する部分のみの対応」しかしていないものが大部分であり、「JIS 第一・第二水準に含まれない文字を文字化けせずに表示・入力可能」なエディタとなると、とたんに数が少なくなる。おそらくインターネット上で公開されている Unicode 対応のエディタのほとんどがこのチェックに引っかかってしまうだろう。

評者のように日常的に JIS 第一・第二水準に含まれない文字を扱う者にとって、このチェックに通らな

図 2：EmEditor の置換ダイアログボックス



くてはいくら魅力的なエディタであっても評価の対象にすらならない。縦書き表示が売りの MM エディタや MS-DOS の頃から根強い愛用者を持つ Qx エディタはこの時点でアウトになってしまう。何れもエディタとしては独自の魅力的な機能を搭載しているだけに残念である。秀丸・EmEditor・xyzy・Aprotool TM Editor・akira22++ はファイルの読み書き・入力共に問題なく実行できた。

■ JIS 第一・第二水準に含まれない文字の検索・置換・grep に対応しているか

多くの文字列を効率的に扱う上で、検索・置換・grep の機能は必須ともいってよいだろう。そこで第一のチェックポイントに合格したエディタを対象に、第二のチェックポイントについて検証してみた。

残念ながら、秀丸・xyzy については検索・置換の文字指定に、JIS 第一・第二水準以外の漢字を入力すると「？」に置き換えられてしまった^[3]。EmEditor については、文字指定は問題なく行えるものの、UTF-8 のファイルを対象に grep (EmEditor では「ファイルから検索」と呼ぶ) を実行できない。Akira22 と Aprotool TM Editor については、文字の指定・UTF-8 のファイルを対象にした grep 共に問題なく実行可能だった。

□ チェックの結果

今回はエディタを評価するために二つのチェックポイントを設定してみたが、はじめのチェックポイントでかなり多くのエディタが引っかかってしまった。二つのチェックポイント両方に合格したのは akira22++ と Aprotool TM Editor のみであり、EmEditor が次点、xyzy・秀丸が更にその下という結果だった。

エディタ名	チェック 1	チェック 2
Akira22+	○	○
Aprotool TM Editor	○	○
EmEditor	○	△ (UTF-8 未対応)
秀丸エディタ	○	×
xyzy	○	×
その他 (MM エディタ・Qx 等)	×	×

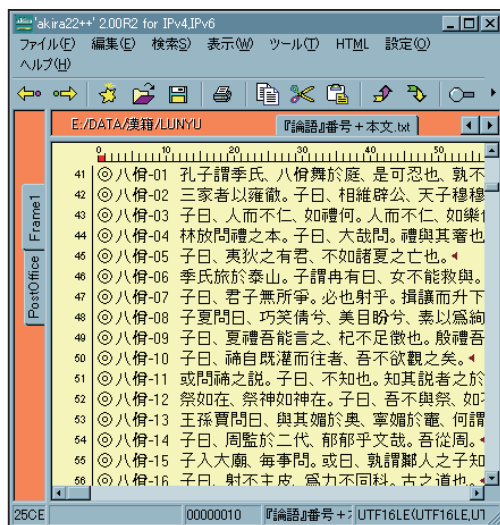


図 1 : akira22++ のエディタ編集画面

評者の現状での一押しは akira22++ である。Aprotool TM Editor も優れたエディタだが、何分更新が止まって三年以上たつし、また操作方法にもかなり癖があるので、初心者には余り勧められない。従って、より標準的な操作で動作可能な akira22++ を一押しエディタとした。

□ 個々のエディタの紹介

● akira22++

<http://www.tg.rim.or.jp/~khf07113/>

綾の明氏作成のテキストエディタ+メーラー (シェアウェア。US\$25)。上記のチェックを全てクリアした一押しエディタである。Lisp での機能拡張ややりのマウスジェスチャまでサポートしているという贅沢ぶりである。

akira22+ は画面のレイアウトや配色が独特なので、初めは戸惑うかもしれない。しかし自分好みのデザインにする事も可能なので問題ないだろう。拡張領域 B については文字表示は問題ないものの、該当文字の後ろに一文字分の空白が生じてしまう。また、検索・置換のキーワードとして設定することはできなかった。

● Aprotool TM Editor

<http://hp.vector.co.jp/authors/VA002891/>

TMEDIT/index.htm

前寺正彦氏作成のソフト (シェアウェア。2,000 円)。Unicode や各ローカル言語への対応のほか、文字鏡番

号や Shift-Mojikyo という独自の手法を使って文字鏡を文字コードとして利用可能。また Unicode に対応したコンソールウィンドウなど、通好みの機能が満載のエディタ。動作がやや不安定なのと更新が三年以上停止状態なのが残念である。

● EmEditor

<http://www.emeditor.com/jp/>

株式会社エムソフト作成のソフト（シェアウェア。3,000 円）。Unicode を中心とした多言語環境への対応と、各種プラグインによる機能拡張を特徴とする。本誌読者層には繁体字簡体字との相互変換・Unicode 文字と実体参照文字との相互変換プラグインなどが有用だろう。

拡張領域 B については、文字表示は問題ないものの、文字が重なって表示されてしまった。

現在はバージョン 4 の開発に中心が注がれており、正規表現の充実などが実現されるようだ。

● xzyzy

<http://www.jsdlab.co.jp/~kamei/>

亀井哲弥作成のソフト（フリーウェア）。Emacs ライクな操作性に加えて Common Lisp を核に柔軟な拡張が可能であり、更にはエディタにとどまらずメーラー・ブラウザなど、これ一つあれば大抵の事ができてしまうという奥の深いソフトウェア。

その余りの奥深さ故に、入り口の段階で躊躇する場合も多いと思われる。インターネット上では、xzyzy の初心者向けのページがいくつか公開されているので、興味のある方はそちらを参照していただきたい。

● 秀丸

<http://hide.maruo.co.jp/>

日本のオンラインソフトの代名詞といってもよいソフト（シェアウェア。4,000 円）。

昨今では中国語への対応など、日本語以外の文字編集にも対応を強化しつつあるが、中国語や多漢字文書については、現状では文字の入力と簡単な編集程度しか実行できないのが痛い。しかし豊富なマクロなどの蓄積はまだまだ侮れない。

原稿執筆現在では、次期バージョン 4 のベータ版が公開されており、日々改良と不具合解消の繰り返されている最中である。

□ おわりに——日本のエディタ界の Unicode 対応について——

以上、日本のエディタ界における Unicode 対応状況についての簡易調査の結果を中心に書いてみた。

師氏の論考から二年をへたが、エディタの Unicode 対応度は師氏の当時からほとんど何も変わっていない。変わった点と言えば、綾の明氏作成のエディタが akira22++ にアップした程度なのである。三年もバージョンアップしていない Aprotool TM Editor の Unicode 対応状況が未だにトップクラスというのは、このソフトウェアの優秀さを示すと共に、日本のエディタ界における Unicode への関心や取り組みが一部の作者のみ突出していて、他の作者は余り関心がないというのを如実に示しているといえよう。

ここから見えるのは、各エディタの作成者及び利用者の多くが Unicode の多漢字環境に対する関心が低いという点ではないだろうか。

彼らが日常エディタで読み書きするデータは、文字集合としては JIS X 0208 の範囲で事足りるのであり、たまたま符号化方式が Unicode の複数の方法に対応しているに過ぎないのだろう。彼らの認識としては「英数字+日本語（＝ JIS X 0208）」なのであり、日本語以外の言語は別言語の枠組みとして捉えられているのだ。

そのため Unicode の普及により多言語に対応するソフトウェアが増えてきたものの、あくまでそれらは日本語と別個の枠組みでしかなく、「複数言語を同時に使用する」あるいは「言語の枠組みを取り扱う」という意味での Unicode の利用には興味関心が及ばないのである。別にこれに関しては日本に限った話ではない。そもそも Unicode ベースのはずの日本語 Windows 上で動かない海外ソフトなど未だに珍しくもないし、Unicode の符号化方式の一つである UTF-8 自体 ASCII コードを従来通り使用したために生まれた方式なのである。

しかし、Unicode の漢字領域を見ると、各国の文字コードを寄せ集めて作った CJK 統合漢字領域の枠を超え、最新の Unicode4.0 では『大漢和辞典』を凌駕する漢字が登録されている。中国を核とする東アジア共通の文字であった漢字が、それぞれの国の枠を超えてまた新たな共通文字としての生命を復活させつつあ

る現在、ローカルな枠を飛び出すエディタが増えて欲しいと願う評者なのであった。

追加：

校正締め切り直前に、DOS時代の定番エディタであるVZエディターの流れをくむ、WZエディターバージョン5のベータ版が公開された^[4]。

そこで、WZにも二つのチェックポイントを試したところ、全て合格という結果が出た。拡張領域Bについては表示が出来ないものの、JIS内のみのUnicode対応だったバージョン4時代から見れば格段の進歩である。他のオンラインソフトに比べればやや高いも

の、十分お薦めできるエディタである。

注

- [1] Unicodeの最新情報については、本誌の該当ページを参照されたし。
- [2] 拡張漢字領域Bへの対応については、個々のエディタの紹介文の中で触れる。
- [3] xyzyについて、設定さえ突き詰めればうまくいくのかもしれないが、それを調査する時間的余裕もなかったので今回は割愛した。
- [4] <http://www.villagecenter.co.jp/soft/wz50/download.html>

その他

千田 大介

❑ Adobe Reader / Acrobat

■ 言語パックの扱いが変わる

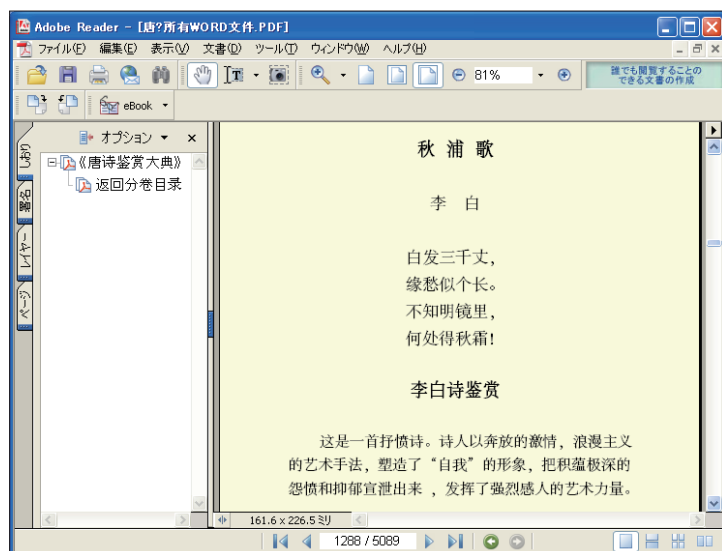
レイアウトデザインされたドキュメントを共有するためのツール、Adobe Acrobatがバージョン6にメジャーアップデートした。新バージョンから、フリーのビューワ、Acrobat ReaderがAdobe Readerに名称変更した。

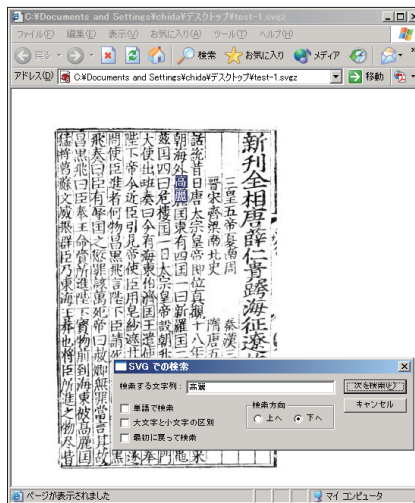
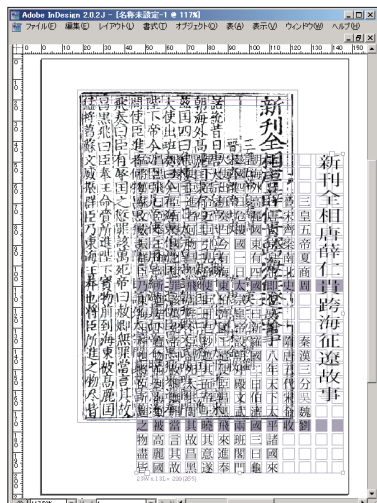
Acrobat 6の新機能については、さまざまなコンピュータ雑誌などで紹介されているのでここでは言及せず、多言語機能に関する変更について見ておく。

Acrobat 5.xまでは、中国語・ハングルなどの使われたPDFを閲覧するためには、あらかじめアジア言語パックをAdobe社のサイトからダウンロードして組み込んでおく必要があった。しかし、この言語パックのダウンロード場所がわかりにくいこともあり、広く知られてはいなかった。

新バージョンでは、中国語などのPDFファイルを開こうとすると、プログラムが自動的にAdobeサイトに接続し、当該言語のフォントをダウンロードするように変更された。これで、中国語などのPDFファイルへのアクセシビリティが向上するものと思われる。

しかし、これはブロードバンド常時接続環境を前提としているのであろうが、中国にノートPCを持参するようなケースを考えると、事前に各言語フォントを組み込んでおけた方が便利な面もある。自動アップデートによる組み込みとあわせて、ファイルのダウンロードサービスも提供してもらいたいものだ。





なお、中国語フォント名が指定された PDF は、やはり中文環境でないと開くことができないし、簡体字ファイルを開くこともできない。これらの点は、今後の改善が望まれる。

□ InDesign 2.0

■ InDesign の SVG 書き出し機能

SVG (Scalable Vector Graphics) は XML ベースの言語で、Web 上でベクタデータの描画が可能になるなど、さまざまな利点を持つことから注目されている^[1]。

SVG 書き出しに対応したソフトとしては、Adobe Illustrator が知られるが、実は InDesign も SVG 形式による書き出しに対応している。SVG と圧縮 SVG

(SVGZ) とに対応しており、また文字コードも UTF-8、UTF-16 などが選択できる。ただし、SVG のインポート機能は備わっていない。これを利用すると、例えば版本画像とテキストを重ねて検索可能な SVG ファイルを作成することも比較的簡単に実現できる^[2]。

テキスト主体とした SVG を書き出すためのツールとしては、現在、InDesign がもっとも使いやすく、安定したツールであると言える。

注

- [1] SVG については、<http://www.adobe.co.jp/svg/> 参照。
- [2] この技法については、安岡孝一「透明テキスト付き画像へのいざない」(『東洋学へのコンピュータ利用第 14 回研究セミナー』、2003.3) 参照。

中国のソフトとネットの動向

千田 大介

□ 政府ソフトの正規版化をめぐる

■ 国外ソフト締め出し発言

中国の政府機関で正規版ソフト採用が推進されていることは、前号のレビューでも言及した。その後、

2003 年元旦より《中华人民共和国政府采购法》(中華人民共和国政府購買法)が施行されて、政府の資材調達への入札制の導入が義務づけられた。このため、ソフトウェアの調達も公開入札で実施されているが、オフィススイート分野では WPS Office の連勝、OS でも紅旗 Linux の導入が目立っている。

そんななか、2003 年 8 月 18 に「国家部委 WPS

Office 2003 換装大会」(中央省庁 WPS Office 2003 換装大会)が開催されたが、翌日の複数の新聞紙上で、大会席上で国务院購買センター高志剛所長の、以下のような発言が報道された。

政府部門がハードウェアを購入する場合には、正規版ソフトをバンドルで購入しなくてはならず、もしも国産ソフトである場合には手続きは非常に簡略になる、という強制的な規定を政府購買法に盛り込む準備をしている。

発言では、この措置が「明らかに外国産ソフトに対する公然の差別である」ことを認めていた。WTO で問題になる可能性も自覚されているようで、「個人的な見通しでは、2010 年には政府が国産ソフトを傾斜的に購買することは制限を受けると思う」とも述べたという。

■ マイクロソフトの苦境

この発言の背後に見え隠れするのは、中国政府のマイクロソフト不信である。中国は、Windows にバックドアが仕掛けられており、政府の機密情報が漏れるのを本気で心配しており、そのために Linux への傾斜を強めているとの説がある^[1]。この説の真偽のほどは定かではないが、中国政府はマイクロソフトの Windows ソースコード開示を受けたとされ、機密漏洩問題を重視しているのは確かなようだ。

この点、2003 年春のビル・ゲイツの中国訪問は大きな話題になったものの、中国市場におけるマイクロソフトの苦境を救う決定打とはなかったようである。そればかりか、マイクロソフト中国の正規版普及活動が、裏目に出た事例もある。

2003 年 5 月には上海市教委に対して、学校内で利用されている海賊版 Office のライセンス料支払いを請求した。それに対して教育関係者は一斉に反発、同 8 月、WPS Office 2003 への全面リプレイスを発表した。それでも Windows の不正利用の状況は変わらないのだが、関係者はマイクロソフトが Windows に対してもライセンス料を請求してきたら、Linux に乗り換えるまでだ、と語っている(南方日報^[2])。

このように、海賊版を利用していながら盗っ人猛々しいとしか言いようのない対応を、いやしくも教職にあるものがとっているのが、中国における知的財産権

意識の現状である。マスコミの報道もマイクロソフトそしてアメリカへの反感が基底にあるのか、ライセンス料支払いを請求された側に好意的である。

ともあれ、ライセンス料請求が自らへの反発に結びつき、開拓された市場は金山公司など安価な製品を擁する国内ソフトハウスに奪われるのであるから、マイクロソフトにはお気の毒というほかない。こうなると、中国語版の価格を大幅に引き下げるか、あるいはアメリカ議会でのロビー活動を通じて元切り上げ圧力をかけるくらいしかないのではないか。

一方、このような知的財産権保護意識と一向に進まない海賊版取り締まりとが、中国におけるソフトウェア産業の成長を妨げる最大の障害になっていることは疑いを容れない。このような状況下では、政府機関の国産正規版ソフト傾斜調達によって一部の国産ソフトハウスは保護されるであろうが、しかしそれが中国ソフトウェア産業全体の発展に結びつくかという、はなはだ疑問である。それよりも、ソフトウェア産業が正当に収益を上げられる環境の整備に努めるのが、困難な道ではあるが、唯一の方法ではないのか。

また、2002 年から 2003 年にかけては、大規模なウイルス禍が相次いだ。中国でウイルスが猖獗を極めた原因の多くは、海賊版 Windows を利用しているために、セキュリティパッチが適用されていなかったことにあるようだ。ウイルス問題ともなると、中国国内だけの問題では済まない。その意味からも、より徹底した海賊版対策と、知的財産権意識の啓蒙とが求められる。

■ WPS と金山公司の動向

政府購買のおかげで、DOS 時代の威風を取り戻しつつある金山公司^[3]であるが、2003 年 8 月末に WPS Office の新バージョン、2003 を発表した。

WPS Office は毎年のようにアップグレードを重ねているが、これは WPS を導入した政府機関から送られてくるバグレポートや要望に一々応えることで、ブラッシュアップと機能強化とを進めていることが背景にある。とはいえ、バージョンアップ頻度が高いのは確かであるが、2003 では正規ユーザーに対する無償アップグレードキャンペーンを展開している。近頃的好業績を反映して、開発人員も増強され、余裕が出てきたようである。

また、前号で言及した WPS 蒙文版であるが、同

社サイトに正式情報が掲載されている。GB18030 と Unicode に対応しているとのこと。入手でき次第レビューしてみたい。

■ 永中 Office

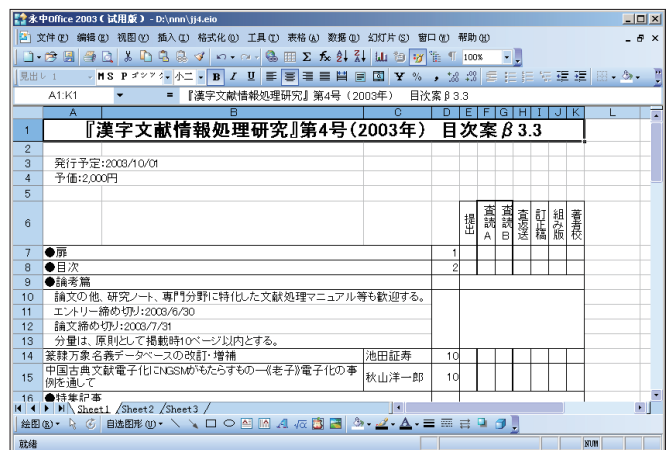
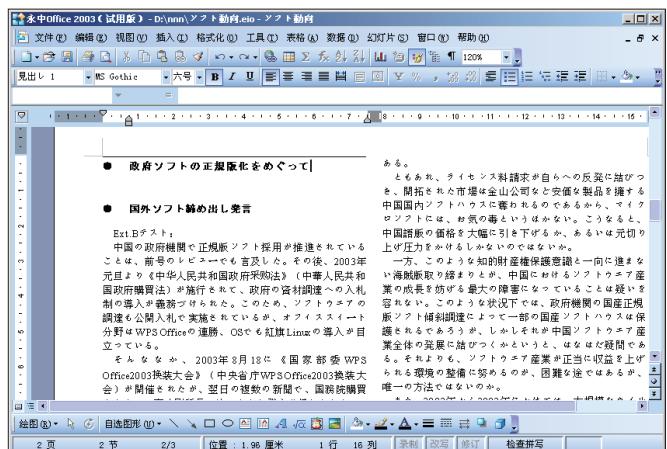
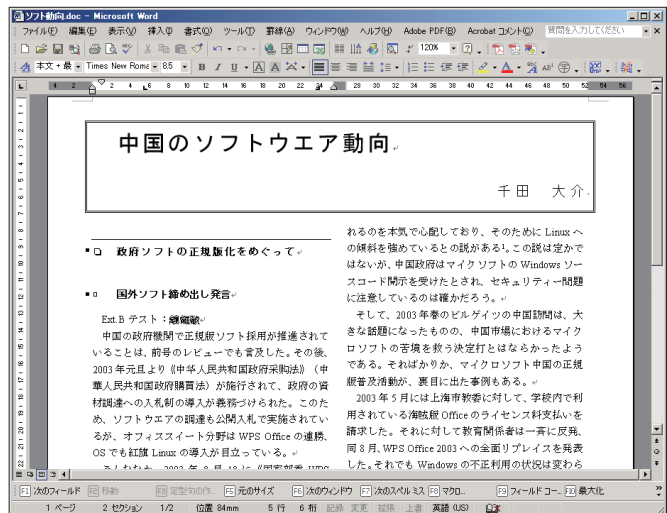
2002 年、マイクロソフト、金山公司に続く第三の中文 Office、永中 Office がデビューした。2003 年 9 月はじめには、新バージョン「2003」が発売されており、Web サイト^[4]から試用版が入手できる。以下、試用版を Windows XP 多言語版（簡体字中国語モード）で試してみた。

永中 Office は、JAVA ベースのソフトで、動作には Sun JAVA VM が必要。Windows 版と Linux 版があり、どちらも大差がないようである。電子表格（スプレッドシート）、文字処理（ワープロ）、簡報制作（プレゼンテーション）の三種が、共通のソースコードの上で動作する。

試用してみたところでは、Office アプリケーションとしての基本的な機能は一通り備わっているようだ。日本語の Word・Excel ファイルも読み込むことができた。Word ファイル読み込みみでは、段組み・スタイル情報などは保持されたが、セクション区切りを改ページと誤認し、フォントがデフォルトフォントに置き換わってしまった。また、Unicode の Ext.B の文字は読み込めなかった。なお、WPS では日本語フォントが選択できないという問題があるが、永中では MS Mincho、MS Gothic などの日本語フォントも使うことができた。Excel では、セルの結合もフォントも保持されているが、縦書きに設定した見出しは、横書きに変わっている。

気になる点としては、JAVA ソフトの宿命として、全般に動作が重いことがある。最近のマシンは高性能化しているので、それでも十分に高速なのだが、他のソフトと併用していると、若干の反応の遅れが気になってくる。また、インライン入力に対応していない。

既にかなり完成されたソフトであるが、個人版 1,998 元（約三万円）と WPS（プロ版 1,298 元）に



比べてかなり高額であることが普及のネックになるだろう。あるいは、正規版ソフト導入の動きが広がる中で、OS の Windows 離れが発生することがあれば、そのときにこそ Linux 環境で動作するオフィススイートとして大いに飛躍するかもしれない。今後の動向に注目していきたい。

□ Google 封鎖事件をめぐる

■ 事件の概要と背景

2002 年 8 月 31 日午前、突然、中国国内から <http://www.google.com/> にアクセスしようすると中国国内の別のサーチエンジンに飛ばされるようになり、中国のネットワークカーたちに動揺が広がった。

実は、中国のインターネットには海外サイトへのアクセスに対するフィルタリング、通称「グレートファイヤーウォール」（電子版万里の長城）が存在する。グレートファイヤーウォールの構築は、1997 年にさかのぼる。当初から CNN には接続できないが Play Boy サイトにはつながるといふ、あからさまな政治的目的をもったフィルタリングであった。ちなみに、日本の有名マスコミで封鎖されているところがないのは、当たり障りのないニュースしか流れないと舐められているからか、あるいはそれらのマスコミを飼いつけているという中国側の自負であろうか。

Google が問題とされたのは、そのキャッシュ機能にある。北米を中心に、いくつかの在外華人団体が中国共産党政権に批判的な立場のオンラインニュースサイトを立ち上げており、中国国内からそれらのサイトへのアクセスは以前から制限されていた。国際的な人権保護団体、チベット問題などのサイトも同様である。しかし、それらのサイトの内容が、Google キャッシュに登録されており、中国国内からも閲覧可能な状況にあったのだ。

中国側のフィルタリングもあまり完璧ではなく、Google のアメリカ以外のサーバの多くにはアクセスが可能であったし、また Google 側でも積極的にミラーを立ち上げるなどの対策がとられた。中国国内でも批判や苦情が大きくなったためか、Google へのアクセス制限そのものは、9 月 12 日に解除されたが、キャッシュへのアクセスは制限されたままであった。まだ、9 月 9 日には Alta Vista へのアクセスも封鎖されてい

る^[5]。

これは、SARS への対応を巡って大問題を引き起こしたのと同じ、共産党政府の伝統的マスコミ政策に起因している。共産党の愚昧な人民を教え導く、という一党独裁の統治原則は、人民を情報統制下に置き情報不足の状態に置くことに眼目がある。そのために、中国ではあらゆるメディアは共産党に奉仕するものと定義され、かつあらゆるマスメディアは中央・地方政府もしくは省庁の管轄下に置かれており、知る権利も報道する権利も確立されてはいない。

インターネット時代になり、英語などの語学能力さえあれば容易に海外サイトにアクセスして情報を入手できるようになったし、電子メールその他の手段で情報を交換することもできるようになったため、共産党のメディア政策は大いに揺らいでいる。しかし、さまざまな論議を引き起こした SARS 隠し事件も、結局は中国のメディア政策を根本から変えるには至っていない。我々が中国と情報をやり取りする場合にも、Web であれメールであれ、中国当局の閲覧の可能性を常に考慮して、慎重に対応する必要がある。

万里の長城が遊牧民族の侵入を防ぎ得なかったように、電子版万里の長城をしてもすさまじいスピードで変化するインターネットの全情報をフィルタリングするのは不可能であると思われる。中国では、経済発展に伴い沿海地域・北京などの都市市民層が形成されつつあり、健全な世論育成の必要性が言われ始めているが、中国が人民を信じて情報をオープンにすることが果たしてできるのか、中国経済の行方とともに注意深く見守る必要がある。

■ 中国のサーチエンジン

中国国内向けポータルサイトとしては、新浪^[6]・網易^[7]・搜狐^[8]の所謂「三大門戸」が知られている。最近のポータルサイトは、より多くのユーザーに訪問してもらう必要から提供するサービスの多様化が進んでおり、多くの機能を他社のソリューションを導入して提供するようになっている。たとえば、Yahoo! Japan のロボット型サーチエンジンは Google であるし、路線情報は「駅すばあと」である。

中国三大門戸もこの例に漏れず、ディレクトリは独自のものを構築しているが、ロボット型サーチエンジンは他社のソリューションを利用している。

現在、中国トップシェアを誇るロボット型サーチ



エンジン、百度^[9]である。搜狐、Yahoo! 中国などで採用されており、中国国内ネットワークの90%以上が利用しているとされる。百度トップページは、Googleを模倣したデザインだが、検索効率もなかなかにはすばらしい。

この百度を急速に追い上げているのが、慧聪^[10]である。2002年8月に「Googleを超える第三世代知能サーチエンジン」のキャッチコピーで市場に参入するや、ライセンス料支払い遅延問題で百度との関係が悪化していた新浪のサーチエンジンに正式採用され、今や中国国内の有力ポータルサイトの大半に、何らかの検索機能を提供するまでに急成長している。

これらのロボット型サーチエンジンは、中国国内向けサービスだけに、グレートファイヤーウォールによって遮断されているサイトにはヒットしない。しかし、例えばそんなサイトの名称の一つ「大参考」で検索すると、百度はヒット無し、慧聪は「股市大参考」などの中国国内サイトにヒットするから、慧聪の方がデータベースへのフィルタリング機能がより洗練されていることがわかる。

ともあれ、ポータルサイトの分野でも、グレートファイヤーウォールという非関税障壁によって、中国国外の企業に不利な状況が作りだれていることになる。WTOに加盟したとはいえ、IT分野での中国市場開放は実質的にはなされない、ということであろうか。

注

- [1] http://www.zdnet.co.jp/news/0209/26/cead_cooper.html
- [2] <http://www.nanfangdaily.com.cn/>
- [3] <http://www.kingsoft.com/>
- [4] <http://www.evermoresw.com.cn/>
- [5] http://www.zdnet.co.jp/news/0209/10/nebt_14.html
- [6] <http://www.sina.com/>
- [7] <http://www.163.com/>
- [8] <http://www.sohu.com/>
- [9] <http://www.baidu.com/>
- [10] <http://www.isearch.sinonet.com/>

附記

締め切り直前に、WPS Office 2003 試用版の配布が始まった。試用した結果、「金山文字」(ワープロ)で日本語フォントが選択できないという問題が修正されているのを確認した。

勝村哲也先生への追悼文

石川 正敏

勝村哲也先生には京都府立医科大学付属病院に先月来病気で入院中のところ、去る9月10日に御歳六十六で逝去されました。ここに謹んで哀悼の意を表します。

先生は、この数ヶ月、体調を崩され、ご実家のある京都にて入退院を繰り返しておられましたが大病ではないと伺っており安心していただけに、突然の訃報を聴き、私自身、その時は驚きのあまり言葉がありませんでした。

勝村先生は、平成12年に京都大学人文科学研究所を退官され、同年に島根県で開学されました島根県立大学総合政策学部教授に着任されました。以後、逝去されるまでの約3年半の間、メディアセンター長、北東アジア地域研究センター長、同大学大学院の北東アジア研究科長の要職を歴任され、開校間もない島根県立大学の運営、大学院の設立、さらに様々な方の研究の相談に乗られるなど多方面でご尽力されました。

私は、島根県立大学の開学時に助手として採用していただき、勝村先生の研究や講義等のお手伝いをすると同時に、私の研究や教育方法などについて師事することになりました。ただ、勝村先生は、大変ご多忙で常に島根県立大学と京都のご自宅、さらに東京や果ては海外まで各地を飛び回っておられたため、私が直接先生にお会いできた時間も限られていました。ですので、私からみた勝村先生といえば、いつもキャスター付きの鞆をガラガラと引き連れ笑顔を見せながら突然現れて、あれこれ話をされたかと思ったらもう大学には

いないと言うように突風という感じでした。雑用などの仕事に関しては、先生は思い出された時に昼夜を問わず電話で指示をだされ、逆にこちらから電話するとつかまらないということも幾度もあり、苦勞することもありました。ですが、私以上に先生は仕事を抱えておられましたので、私は少々時間的に厳しい仕事でも期待に応えなければと思っていました。仕事以外では、様々なご配慮をいただいたほかに、先生からお聞かせ頂いたご指導や雑談を含む様々な話は、情報科学が専門の私にとっては新鮮なものも多く大変勉強になりました。勝村先生にとっては決して優秀な助手だったとは言えなかったと思いますが、多少なりとも先生のお役に立てたことを誇りに思います。

島根県立大学は開学4年目となり半年後には一期生の卒業を迎えようとしています。また、先ほども書きましたが勝村先生のご尽力もあり、本年度からは大学院も開設されました。このように大学の研究体制の充実するなか、私は勝村先生からさらに様々なご指導がいただけたと思っていただけに、先生のご逝去は本当に残念でなりません。そして現在、私は、勝村先生が残されたeKanjiを含む北東アジア関係資料に関するデータベース構築プロジェクトに参加しており、先生のご期待に添えるよう、できる限りの努力をし、完成させたいと思っております。

最後に、改めて勝村先生のご冥福を心よりお祈り致します。

学術リソース レビュー

本レビューは、各レビュー執筆者が本会の掲示板上で日々行われている情報交換をベースとして、それぞれ独自にインターネットリソースやソフトウェアの検証作業をすすめた末に執筆されている。

ようやく日本でも国立国会図書館のデジタルライブラリーや東洋学文献類目・漢籍データベースを初めとするデジタルデータベースなど、学術目的に利用可能なインターネットリソースが充実してきたが、中国のそれに比べれば未だ道遠しと言わざるを得ない。特に日本の中国関係学会からの情報発信が少なすぎる事は問題だろう。本レビューで紹介する Web サイトも圧倒的に中国発のそれが多い事が端的に示している。Web サイトは作るより維持管理する方が遙かに重要である。決して「作成＝完成」ではない。

今号も国学宝典・基本典籍庫・中国基本古籍庫などの中国発の大規模テキストデータベースや、超星やそのミラー版とも言える普来得のレビューに紙幅を割いている。これらのソフトウェアに興味をお持ちの方は、是非ともご一読頂きたい。

Contents

学術サイト	図書館と OPAC	小島 浩之 138
	東アジアの言語に関連したサイト	山崎 直樹 142
	中国語教育サイト.....小川 利康 145	中国古代史.....山田 崇仁 149
	中国近現代史.....佐藤 仁史 152	中国古典文学.....千田 大介 158
	中国現代文学.....小川 利康 160	仏教.....師 茂樹 162
学術ソフト・製品	国学の CD-ROM.....千田 大介 165	中国の文献データ製品...千田 大介 169
	電子辞書 2 機種.....田邊 鉄 173	字通.....山田 崇仁 174

❖学術サイト

図書館とOPAC

小島 浩之

□ はじめに

筆者は昨年の本誌特集1において「NIIにしてもNDLにしても今秋を一つの区切りとして事業に取り組んでいるようなので、現段階では不透明な部分も多い。従って本書の出版時には相当な変化がみられる可能性もある」^[1]とまとめた。結局、昨年の本誌出版直前に、どちらも大幅なリニューアルがなされ、Webサイトのレビューというものの難しさを改めて感じた。今年度のレビューは昨年度の拙稿の訂正、補足を兼ね、NDLとNIIから採り上げていこう。

□ 1. 国立国会図書館 (NDL)

NDLは昨年10月に関西館が開館し、東京本館と関西館の2館体制となった。この結果本会の活動に関係のあるアジア関係資料とデジタルアーカイブズ関係は、ともに関西館の所轄となり新たなスタートが切られた。NDLについて述べるべきことは多くあるのだが紙幅の関係もあるので、本稿ではNDL-OPAC、アジア言語OPAC、近代デジタルライブラリーに絞って言及する。

● NDL-OPAC

<http://opac.ndl.go.jp/index.html>

これまでに比べて情報量も格段に多くなった。特に和書の書誌データは明治期以降の約275万件(2003.6.30現在)が検索可能で、昨年までの「1948年以降約200万件」に比べて75万件ものデータが増加している。検索についても高度な条件設定が可能となっている。また利用者登録すればNDL-OPACを通じて複写サービスなども受けられ、ソフト、ハードの両面を通じて、国立図書館として遜色無いものになっている。さらに待望の雑誌記事索引もNDL-OPACの一機能として公開された。公開直後は最新データのみ

だったが、約一年を経た現在では1983年以降の雑誌記事を検索することが可能となっている。

● アジア言語 OPAC

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

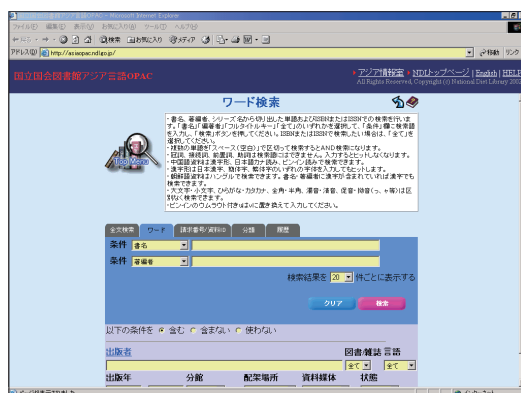
NDLでは多言語対応OPACも昨秋にお目見えした。UTF-8対応で現在のところ日本語、中国語(簡体字、繁体字)、ハングルに対応している。仕様はこれまでに先行しているNIIを中心とする各大学図書館のものに準じており、システムは昨年紹介した獨協大学と同メーカーのものである。またNIIの漢字統合インデックスを利用しているという。このため漢字で検索値を入力する際、字体を区別する必要はない。

● 近代デジタルライブラリー

<http://kindai.ndl.go.jp/index.html>

NDLが所蔵する明治期刊行図書のうち、著作権保護期間の満了を確認した約4万7千冊を画像化したもの。中国学関係の書籍も多く含まれている。公開画像はLindraおよびGIFの2種類の形式が用意されている。Lindraは高圧縮で軽いのだが、OS環境がWindows 98以上もしくはMac OS 9以上に限定され、かつプラグインのダウンロードが必要となる^[2]。本

アジア言語 OPAC



サイトについては公開当初より、様々な紹介・批評が種々の媒体に掲載され、大きな反響を巻き起こした^[3]。確かに、日本における大規模デジタルライブラリの嚆矢として、質・量ともに高く評価されるべきものだろう。

なお公開と同時に、著作権切れかどうか不明な書籍について公開調査が行われた。公開調査は本年 8 月 1 日で一旦終了したが、現在も広く情報の提供を呼びかけているので、情報をお持ちの方は報告していただければと思う^[4]。

□ 2. 国立情報学研究所 (NII)

NII は昨年来、学術コンテンツのポータルサイト GeNii の構築に力を入れている。GeNii のコンテンツとして新たに加わった二つを紹介する。

● Webcat Plus

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

Webcat Plus は、NII によって提案された次世代 OPAC と言えるだろう。従来の NACSIS-CAT 参加図書館による書誌・所蔵情報データベースに加え、商業用の図書情報データベースである「Book」データベースをも基礎データとしている。これにより図書館所蔵のものだけでなく、流通現場の最新の図書情報をも併せて知ることができるようになった。検索には連想検索という新たな手法が用いられたため、検索値に関連するキーワードを含む図書へも導いてくれる。残念ながら現在は日本語のみの対応のため、今後の拡張に期待したい。

● 大学 Web サイト資源検索 (JuNii)

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

全国の大学や研究機関等が Web 公開している研究成果等にメタデータを付与し、データベース化したものの。現在試験運用中だが、大学等が公開する Web サイト上の研究成果等 (約 6,000 件) のポータルサイトとなっている。

□ 3. 国内漢籍目録など

● 全国漢籍データベース

<http://www.kanji.zinbun.kyoto-u.ac.jp/kanseki/>

早いもので、このデータベースも発足から 3 年目となった。2003 年 9 月 1 日現在の参加館は 17 機関で、

大学図書館に加え公共図書館や三康、宮内庁書陵部といった研究機関も加わり充実した内容となっている。しかし、17 機関のうち自己のサイトに漢籍データベースへのリンクが貼ってあるのは、半数の 8 機関にすぎない。その 8 機関も、サイト中のリンク集のページに申し訳程度にリンクされている場合が多い。これでは何のために漢籍データベースに参加しているのか解らないではないか。自館の所蔵目録の一部なのだからもっと堂々とリンクを貼り、漢籍総合目録としての意義を世間に訴えて欲しい。また参加機関の収録漢籍の範囲(特殊文庫ならばその文庫名、既存冊子体のデータならばその書誌データなど)を表示するページを設けていただくとありがたい。検索すれば解るものもあるのだが、最初に一覧できるとなお良いように思われる。

● 大原漢籍文庫目録

[http://www.rib.okayama-u.ac.jp/library/](http://www.rib.okayama-u.ac.jp/library/kanseki.htm)

[kanseki.htm](http://www.rib.okayama-u.ac.jp/library/kanseki.htm)

岡山大学附属図書館資源生物科学研究所分館が所蔵する明・清時代の農書コレクション 4,838 冊。書誌情報だけでなく簡単な解題も付されている。研究所の前身たる大原農業研究所が、大正 12 年に蒐集したものだという。

● 関西大学図書館廣瀬文庫

[http://www.kansai-u.ac.jp/Library/annai/](http://www.kansai-u.ac.jp/Library/annai/collection/hirose_int.htm)

[collection/hirose_int.htm](http://www.kansai-u.ac.jp/Library/annai/collection/hirose_int.htm)

元関西大学学長廣瀬捨三の蔵書を受け入れた際のリストを公開したもの。国書、漢籍、洋書の別に PDF ファイルになっている。漢籍は 89 点 (1,137 冊) で一見したところ江戸期の和刻本がほとんどを占める。詳細な目録の作成が待たれる。

● 京都大学附属図書館所蔵谷村文庫目録

[http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/tani/](http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/tani/index.html)

[index.html](http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/tani/index.html)

● 京都大学附属図書館所蔵清家文庫目録

[http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/](http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/seike/index.html)

[seike/index.html](http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/seike/index.html)

谷村文庫については本誌第 2 号のレビューでも採り上げている。ただ当時は“谷村文庫書名一覧”で準貴重書のみ公開であった。ここに全容が公開され、順次全体のデジタル化に向かっていることは歓迎したい。清家文庫は儒学の名家清原家の蔵書。和刻や鈔本の漢籍を多く含んでおり、779 点中 160 点が全文で

ジタル化されている。

●貴重資料画像&文庫目録 統合検索

<http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/kichosearch/search.html>

京都大学附属図書館でデジタル公開されている貴重書および特殊文庫を統合検索できる。上述のような特殊文庫以外の、一般貴重書として公開されている漢籍を検索するのに便利である。デジタル化されている一般貴重書の中には永楽大典（巻910-913）などが含まれている。

●中国清代民国公私文書コレクション

<http://ddb.libnet.kulib.kyoto-u.ac.jp/exhibit/b65/b65cont.html>

京大法学部所蔵の中国公私文書から295件の画像データを収録したもの。残念ながらコレクションの由来は昭和18年に購入されたという以外、詳細は一切不明とのこと。

●駒澤大学図書館蔵禅籍善本図録

<http://www.komazawa-u.ac.jp/~toshokan/zenseki/>

2000年3月に発行された『禅籍善本図録』をデジタル化したもの。中国関係は明・清版を中心に宋元版をも含んでいる。UTF-8による公開。

●流通経済大学天野元之助文庫

<http://www.rku.ac.jp/~hara/amanotop.htm>

中国農業史研究者の天野元之助の蔵書目録。流通経済大学所蔵のものと天野家所蔵のものがPDFで公開されている。

●学塾三餘堂関係資料目録

在野の儒学者藍澤南城が越後に開いた学塾“三餘堂”、およびその流れを汲む“藍澤義塾”伝来資料の目録。柏崎市立図書館のWebサイト（<http://lib.city.kashiwazaki.niigata.jp/>）中、“トップページ”郷土のことを調べる＞藍澤南城と三餘堂”で公開されている。資料は柏崎市立図書館所蔵のものと新潟県立図書館所蔵のものに分かれているが、どちらも漢籍（和刻中心）を多数含む。

●洲本市立図書館古書目録

<http://www.library.city.sumoto.hyogo.jp/kosyo.html>

洲本市立図書館に創設以来保管されてきた古書の目録。漢籍は和刻本を中心とした200点余りが含まれる。寛政の三博士と称された儒者柴野栗山の蔵書を多

く含む点に特徴がある。文字コードがS-JISのため、全ての文字を表示できないことが難点。

●福井県立図書館松平文庫目録【電子版】

http://www.library.pref.fukui.jp/index_da1.html

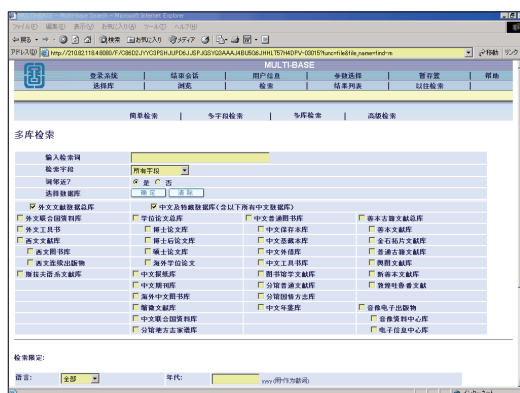
越前松平家および藩校伝来の漢籍が検索できる。検索可能な漢籍は1968年に出版された『松平文庫目録』国書漢籍の部所載のもの。

□ 4. 中国国家图书馆

国家图书馆（<http://www.nlc.gov.cn/>）のOPACは2002年11月28日付で新たなシステムに更新されている。システムの文字セットとしてUTF-8を採用しているため、日本語資料の書誌の表記は、簡体字ではなく日本の漢字でなされている。異字体テーブルも用意され、異なる字体間において検索上の互換性を保持している。例えば漢書、汉书どちらの入力でも同じ検索結果を得ることができるので、検索値の入力に簡体字用IMEへ切り替える手間が省けるようになった。

トップページより文献検索へ進むと、IDとパスワードを求められるので一瞬戸惑うが、匿名登録ボタンをクリックすれば検索画面に入ることができる。検索画面には簡単検索、多字段検索、多庫検索、高級検索の4種があり（デフォルトは多庫検索）、利用者インターフェースも充実したものになっている。ただし統合検索である多庫検索以外は、検索範囲のデフォルトが中文及特蔵データベース内の図書になっているので注意が必要である。多庫検索以外で外国文献を検索したい場合や、より絞り込んだ範囲で中国書を検索したいならば、最初に検索画面左上の“选择库”をクリックし、必要な範囲を指定せねばならない。

中国国家图书馆



□ 5. 臺灣國家圖書館

臺灣國家圖書館 (<http://www.ncl.edu.tw/>) の OPAC (<http://www.ncl.edu.tw/n2Frame.htm>) も一新された。古籍から出版予定図書まで、一般流通資料から灰色文献まで種々の OPAC が取り揃えられている。このように、資料の種別によって検索画面が異なり細分化されているというのは、特定の資料を検索するのに威力を発揮する。反面、資料を特定せずブラウジングするには非常に使い勝手が悪い。是非とも統合検索用のインターフェースも用意して欲しい。

また台湾地区の総合目録たる横断検索の NBINet (<http://nbinet.ncl.edu.tw>) もインターフェースが更新されている。ただ一度に複数の異なる書誌条件で検索できないのが難点。まずタイトルならタイトルだけで検索し、そこから別の書誌条件で絞り込む必要がある。つまりタイトルに経済という語句を含み、出版者が文史哲出版、出版年が 2003 年の図書を検索したい場合、3 回の検索により順次絞り込んでいかねばならない。

□ おわりに

本誌のこのレビューも今年で三回目を迎えた。最初に依頼を受けた際、果たしていつまでネタが続くであろうかと心配した。しかしそれも杞憂だったようで、書くべき事が追いつかず、自己の勉強不足が恥ずかしい限りである。

この 1 年は OPAC のインターフェース改良、そして新たな漢籍目録の公開が際だっていた。相互貸借や複写依頼など他のサービスと OPAC の連結や、より利用者のニーズに合った形での検索結果の提示など様々な工夫が随所に見られるようになった。Web インターフェースの OPAC が登場してから 10 年を経ないうちに、次世代 OPAC へと展開し始めているのは大変な驚きである。また漢籍目録の Web 公開が大学図書館だけでなく、公立図書館にも広がりをみせていることは喜ぶべきことであろう。今回紹介した柏崎、洲本、福井の例でも明らかなように、日本には江戸期の儒学者、

藩校の蔵した漢籍が多く残っている。唐本のみならず、これらと和刻漢籍の体系的な整理・公開も進展することを期待したい。

さらに NDL の近代デジタルライブラリーに見られるような、大規模な電子図書館の登場は、喜びの声をもって迎えられると同時に、様々な問題も投げかけることとなった。システム、画像形式といったハードの問題から著作権などの法整備の問題まで、デジタル図書館の前に立ちはだかる壁は、まだまだ高い。

注

- [1] 拙稿「国内図書館における学術研究情報発信の現状」(『漢字文献情報処理研究』第 3 号, 2002.10)
- [2] 高圧縮の画像形式の場合、OS 環境が限定されたり、プラグインの導入が必要となる場合が多い。従って利用者への便宜を図るため、公開者側は複数の画像形式を準備する無駄を強いられている。こういった例は国会図書館以外にもアジア歴史資料館に見られる(アジア歴史資料館の詳細は本誌の佐藤氏のレビュー参照)。しかも高圧縮の画像は企業が種々の形式を競い合っている。従って今後、公開者の選択した画像形式が淘汰されるということもあり得る。これは公開者、利用者の双方にとって由々しき問題だろう。これら高圧縮の画像形式の国際標準化が今急がれているのである。
- [3] なかでもメールマガジン“ACADEMIC RESOURCE GUIDE”(<http://www.ne.jp/asahi/coffee/house/ARG/index.html>) の 7 回にわたる特集記事は一読の価値がある。
- [4] 一般的な現代中国人名辞典に出ている人物が、不明扱いされている例がままあり、近現代の中国人に関する事前の調査不足が否めない。またリストが EUC のため表示できない文字が多いのも難点。せっかく UTF-8 の OPAC を公開したのに、これでは竜頭蛇尾でないか。

東アジアの言語に関連したサイト

山崎 直樹

□ はじめに

この文章では、東アジアの諸言語に関係したサイトをいくつかと、言語情報処理に関するサイトをいくつか紹介する。

□ 東アジアの言語に関連するサイト

● 国内言語学関連研究機関 WWW ページリスト

<http://www.sal.tohoku.ac.jp/~gothit/kanren.html>

後藤齊（東北大学大学院文学研究科言語学研究室）氏の手になるサイト。下に掲げるトップページの「目次」を見れば、いかに網羅的であるかがわかる。

第一部 学会編

東日本編

北海道／東北／関東／中部

西日本編

近畿／中国／四国／九州／外国

第二部 プロジェクト・研究会／リンク集／電子

辞書／検索等／多言語処理／出版、書籍流通／資料展示等／非研究機関、未整理／研究者／終了した催し

対象言語別編

日本語編／東アジア言語編／英語編／ドイツ語編／ロマンス語編／諸言語編

試みに「東アジア言語編」を見ると、中国語、朝鮮語は、もちろんのこと、ウィグル語、チベット語、モンゴル語、ツングース諸語などのサイトも紹介されている。

● 日本中国学会

<http://www.soc.nii.ac.jp/ssj3/index.html>

日本中国学会のサイト。機関誌である『日本中国学会報』の掲載論文目録があるので、調べものに便利だが、複写依頼の便を考えて、個々の論文の掲載ページ数も載せてほしいものである。

● Journal of East Asian Linguistics

<http://www.kluweronline.com/issn/0925-8558>

Kluwer Academic Publishers から出版されている標題の学術誌のサイト。既刊号の掲載論文の摘要が閲覧できる。

摘要を表示したところ

Journal of East Asian Linguistics		
Previous Abstract	Full Text PDF (186 K)	Next Abstract
Journal of East Asian Linguistics 12 (3) 259-311, July 2003 Copyright © 2003 Kluwer Academic Publishers All rights reserved		
Temporal Reference in Mandarin Chinese		
Jo-wang Lin Department of Foreign Languages and Literatures, National Chiao Tung University, Hsinchu 30050, Taiwan E-mail: jowang@cc.nctu.edu.tw		
Abstract		
This paper discusses how Chinese, a so-called tenseless language, determines its temporal reference. For simplex sentences without time adverb or aspectual marker, I show that temporal reference is correlated with aktionsart or grammatical viewpoint. For sentences with an aspectual marker, I discuss the temporal semantics of <i>le</i> and <i>guo</i> in detail, showing how their tense/aspectual meanings contribute to temporal reference. I propose to analyze <i>le</i> as an event realization operator and <i>guo</i> as an anteriority operator. For subordinate clauses, I show that temporal reference of complement clauses of verbs is		

論文の全文も電子化されている（PDF）が、“Pay per View” で提供されている。

□ ディレクトリ

「○○語に関するサイトには、どのようなものがあるか」を調べようと思うなら、まず、著名な検索サイトのディレクトリを見るのがよい（これは常識であって、言うまでもない）。以下に、今後、調べようと思う人のために、どの階層の下を見ればよいかを記しておく。

● Google Directory

<http://directory.google.com/>

“Science > Social Sciences > Language and Linguistics > Natural Languages” と辿ると、個別言語（及び語族）のリストが得られる。

● Yahoo! Asia

<http://asia.yahoo.com/>

“Home > Social Science > Linguistics and Human Languages > Languages > Specific Languages” と辿ると、個別言語のリストが得られる。

言語学は、日本では「人文科学」という範疇に収められることが多いが、上記のサイトでは、どちらも“Social Science(s)”の下にあることに注意。

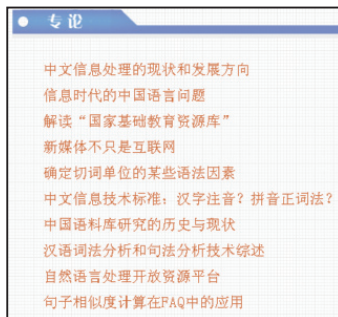
□ 中国語に関するサイト

● 中国语言文字网

<http://www.china-language.gov.cn/>

中国の言語計画を司る国家組織であるところの“国家语言文字工作委员会”が主催している。よって、このサイトは「中国の諸言語の公式サイト」といった趣がある。

内容は、非常に豊富である。「中国語情報処理」のコーナーなどは一見に値するだろう。



中国语言文字网

「中国語情報処理」のコンテンツの一部

また、2000年に制定され、「すわ、言語統制か？」と話題を呼んだ「全国共通言語文字法」（国家通用语言文字法）に関する情報も手に入る。

ただし、どのページにも小細工が施してあるらしく、メジャーなブラウザの最新版でないと、うまく（＝製作者の意図どおりに）表示されないかもしれない。

このような現象はこのサイトだけに限ったことではないが、公的なサイトであるにもかかわらず、限られた環境でしか閲覧できないWWWリソースの作り方をするのは、いかがなものかと思う。“accessibility”という言葉を知らないのかと言いたくもなる。

□ その他の言語に関するサイト

● 羌族语言与文化网站 The Qiang Nationality Language and Culture Web Site

<http://victoria.linguistlist.org/~lapolla/qiang/qiangmain.html>

中国の少数民族・羌の言語と文化に関するサイト。研究文献目録と関連リンク集は、情報量が多く、一見の価値がある。リンク集を特に取り出して以下に掲げる。

● Links to other sites related to Tibeto-Burman languages, endangered languages, typology, and other topics in linguistics

<http://victoria.linguistlist.org/~lapolla/qiang/links.html>

チベット・ビルマ諸語関連のサイト、周辺の諸言語のサイト、言語情報処理ソフトや関係する言語の文字フォントが得られるサイト、関係する地域の地図が得られるサイト、一般言語学のサイトへのリンクが、豊富にある。

□ 漢字圏全体を見渡したサイト

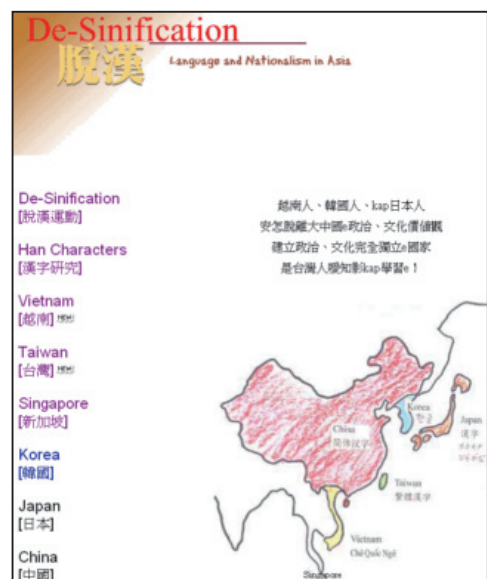
● 脱漢 De-Sinification ---- Language and Nationalism in Asia

<http://www.de-han.org/>

トップページのアオリ文句を以下に掲げる。

越南人、韓國人、kap 日本人安怎脫離大中國 e 政治、文化價值觀建立政治、文化完全獨立 e 國家是台灣人噯知影 kap 學習 e !

“脱漢”のトップページ



これは、「漢羅」と呼ばれる「漢字ローマ字混じり文」で、台湾語を記した文章である（漢字だけ拾って読んでも概要はわかる）。

このサイトの文書の多くは、この“漢羅”による台湾語で書かれている（文字コードは、Big5）ので、筆者には解読不能だが、言及されている研究や文献を見ると、学術的なものが多い。単なる扇動サイトではない。漢字圏における言語とナショナリズムの関係や、漢字圏の言語政策を、地域横断的に考えたい人は、参考書目のコーナーを見るだけでも、得るものが多いであろう（挙げられている文献は、英語か標準華語の文献である）。

扱っている地域は、ベトナム、韓国、シンガポール、日本、中国。

□ 言語情報処理に関連するサイト

● 言語情報処理ポータル

http://www.kc.t.u-tokyo.ac.jp/NLP_Portal/

（社）電子情報技術産業協会（JEITA）知識情報処理技術委員会によって運営されているサイト。

サイトの中には、「アジアの言語資源カタログ」という検索システムもある。もっとも、“Chinese”では何もヒットしなかったが。

言語情報処理ポータル
検索画面の一部

Field Name	Keywords(SEE THE FOOTNOTE)
Price	FREE
Subject language	CHINESE
Type	TEXT
Identifier	
Identifier	
Identifier	
Identifier	
Identifier	
Identifier	
Identifier	

● 中央研究院語言学研究所

<http://www.ling.sinica.edu.tw/>

“文献語料庫研究室”には、さまざまなコーパスへのリンクがある。

● 中華民國計算語言學學會

<http://ckip.iis.sinica.edu.tw/ROCLING/>

“語料庫 / 詞庫”に、内外のさまざまなコーパスへのリンクがある。また、“中文句結構樹”には、文法構造や文中での機能に関する情報を埋め込んだタグ付

きコーパスがある。

どのような情報が埋め込んであるかという、例えば、“小蜜蜂吸完花蜜”という文であれば、

```
S ( agent : NP ( property : VH13
: 小 | Head : Nab : 蜜蜂 ) | Head :
VC31 : 吸完 | theme : NP ( Head : Naa
: 花蜜 ) )
```

という記述になる。これらの情報を指定して検索することが可能である。

検索画面

設定毎頁筆数: 50

關鍵字: 詞: 語義角色: 特徵:

句型:

關鍵詞示例:

型式1. Q: 如果我要找詞類為 VH 開頭的。
A: 請在詞類欄輸入 VH%

型式2. Q: 如果我要找角色為 manner 但詞類不是 VH 開頭的。
A: 請在角色欄輸入 manner・詞類欄輸入 ~VH%

註 % 萬用字元符號
~ 非的項目

句型示例:

型式1. Q: 詞組水平結構表示
A: RolesCat:Word < RolesCat:Word < ...

型式2. Q: 上下層結構表示
A: Cat(RolesCat:Word)

執行 清除

● 北京大学計算語言学研究所

<http://www.icl.pku.edu.cn/>

“科研信息>学术成果”とたどると、“现代汉语分词和标注系统”というオンラインで動作するアプリケーションソフトに辿りつける。

これは、入力された中国語のテキストを、自動的に単語単位に分節し、品詞のラベルを付してくれるものである。オプションで音注も施してくれる。

北京大学計算語言学研究所

“现代汉语分词和标注系统”の分析結果

汉语文本自动分词与词性标注测试结果

输入的句子为:

当我们在这里住一宿时, 差点儿上了当

切分与标注的结果为:

当/p{dang1} 我们/r{wo3men5} 在/p{zai4} 这里/r{zhe4li3} 住/v{zhu4} 一/m{yi1} 宿/q{xiu3} 时/Dg{sh12}, /w{,} 差点儿/d{cha4dianr3} 上/v{shang4} 了/u{le5} 当/v{dang1}

例えば、“当我们在这里住一宿时、差点儿上了当”は、次のようになった。

当 / p{dang1} 我 们 / r{wo3men5} 在
/ p{zai4} 这 里 / r{zhe4li3} 住 /

v {zhu4} 一 / m{yi1} 宿 / q{xiu3}
时 / Dg{shi2} , / w{.} 差 点 儿 /
d{cha4dianr3} 上 /v {shang4} 了 /
u{le5} 当 /v{dang1}

中国語教育サイト

小川 利康

□ 定着するか？ネット学習サイト

近年のトレンドを見れば中国語教育サイトでも「Web コンテンツの大規模化」及び「遠隔教育の普及」が徐々に本格化し、すでに定着した感がある。三年前にも記したように^[1]、今や運営の主体は企業に移り、コンテンツも質量ともに向上した。とはいえ、個人の善意で Web コンテンツを提供しているのではない以上、営利企業として収益が上げられなければ成功とは言えない。

日本でも既に複数の企業が中国語学習サイト運営に乗り出している。とはいえ、世に言う e ラーニングのように、学習管理システム^[2] (LMS: Learning Management System) まで備えたものは少ないようだ。依然として語学教材をウェブベース化 (WBT: Web Based Training) したものが主流で、大まかな分類として、

- ・遠隔授業型
- ・オンライン学習型

の二つに分けられるが、その両者の混淆も含まれる。注意して置かねばならないのは、この二つの形態いずれにしても、現状に於いては常に教室における対面型授業の擬似的な再現が意図されているため、現状においては「いつでも、どこでも」を実現しているほかは、どちらも対面型授業の長所、短所ともに引き継いでいる。仮想現実を描くゲームが現実世界の比喻を乗り越えたところに存在価値を見いだしたように、対面型授業の枠組みから自由にならねば常にその代替品として扱われることになる。その運命を甘受するのがネット学習サイトなのかもしれないが。

第一の遠隔授業型とは、Web カメラとマイクを利用して、多地点を結んで「どこからでも」参加できる対面授業を実現する。少人数教育を旨として教師一名に対して四名以下の学生が指導を受ける形がほとんどで、学生教員間のリアルタイムコミュニケーションである点で、対面型授業に最も近いオンライン教育といえる。テレビ電話など専用機器を利用する形でいち早く普及し、最も支持を集めている形態といえる。しかしながら、動画音声を安定的且つ明瞭に提供できるかどうかは不安要素がつきまとう上、少人数教育だけに教員の資質がサービスの評価に直結するので人材の確保、教育プログラムも不可欠となる。運営企業側からすれば、講師を「どこからでも」採用でき、現に中国から任用している例がある。ネットワーク回線の利用により、通信費削減だけでなく、中国の低廉な人件費のメリットも享受できるわけだ (画面は Live! China の例)。

もう一方のオンライン学習型はネットワークを経由して教材を提供するもので、WBT の基本となる。教材内容は、例えば授業内容録画ビデオ + PowerPoint 教材 (授業内容の忠実な再現)、マルチメディア要素を付加した HTML, Flash^[3] (教科書のデジタル化) に



至るまで様々だが、遠隔授業型とは異なり、リアルタイムコミュニケーションも必要としないので、「いつでも、どこでも」が可能となる。また、デジタル教材作成に一定のコストがかかるとしても、一定の受講生数が確保できれば、遠隔教育と異なって人的コストの負担は軽く、収益を確保しやすいと予測される。もっとも、この形態での学習にはコミュニケーションの要素が欠落するので、多くの場合、遠隔授業型と併用するか、メール・BBSによる対応が必要になってくる。

以上の二つの形態に代表される学習サイトが日本で定着できるか否かを占うには幾つかのポイントがあるだろう。

まずユーザ側から見た学習効果の問題が挙げられる。学習効果とは明確な到達目標とその目標に基づくテストニングがなければ測定不能であるが、それが可能な学習コースの設計がなされ、学習到達度を判定するシステムを備えているかどうか。有料サイトの全教材を試すのは不可能であるが、管見の限りでは、雑然と教材ファイルを断片的に並べるだけの中国語教育サイトがほとんどで、文法体系も、習得目標もほとんど考慮されていない。教材もオンライン用に最適化されておらず、紙媒体の教科書をそのまま移植したものである。

次にコストの問題。少々酷な話だが、大学や高校を選ぶのと違って、ユーザ側からすれば、ネット学習に求めるのは学歴でも、友人関係でもない。通常の学校教育を選ばずに、ネット学習を選択するのは、言うまでもなく時間・費用の節約のためだ。だとすれば、ネット学習ほど、費用に対する学習効果がシビアに測られる世界はないだろう。もしも任意性の高いネット学習サイトが定着できるとしたら、ひとえにこの一点にかかっていると思われる。以下では、教育形態から遠隔教育、オンライン学習の二つに分けた上で、海外の華人向け漢語教育（うち、学歴、非学歴双方を含む）に関連するものを簡単に紹介する。なお、昨年の春の拙文と内容上同じサイトを紹介している部分がある。今回改めて巡回してレビュー内容もアップデートする必要と考えたためなので、ご了解を乞いたい。

□ 遠隔教育サイト

● ECC WEB Lesson 中国語

<http://www.eccweblesson.com/china/index.html>

● しゃべるねっと Live China!

<http://www.shaberunet.com/livechina/>

2003年6月にECCが、8月に「しゃべるねっと Live China!」がそれぞれ中国語 Web 講座をスタートさせた。ブロードバンドの浸透と中国語ブームの相乗効果によるものであろう。いずれも中国語入門からの学習コースで、ネイティブスピーカーによる個別指導を前面に押し出している点では共通する。相違点は、ECCが事前10分をオンライン教材の学習に当て、続く20分がWebカメラを介した個別学習となるのに対し、Live China!は25分間全て個別学習に当て、教材は別途紙媒体の教材を配している点である。ECCは音声教材CDとパソコン対応教材も別途配布されているので、違いは対面学習時間の差だけで、両者とも授業形式は似ている。なお、Live China!は講師（全員女性）の3分の2が中国在住者である点が特色である。中国からの授業もネットワーク環境さえあれば確かに可能なので、リスクもあろうが、新たな可能性を感じさせる方法である。

□ オンライン学習サイト

eラーニングの主流は資格教育である。とりわけコンピュータ関連の資格教育分野は相当数に上る。資格教育は特定試験合格を目的としているだけに、明確な到達目標を設定しやすく、学習効果も測りやすく、オンライン学習に適しているといえる。中国語学習サイトでも資格試験合格やスコア向上を目的とするサイトがある。

● 「楽々スクール中国語」

<http://www.omronsoft.co.jp/SP/win/school/toku.html>

この学習サイトの特徴は中国語入力ソフトも含めた統合的環境をユーザに提供する点であろう。運営主体であるオムロンソフトウェアは中国語入力 IME である cWnn の開発販売を行ってきた企業であり、その経験と実績を生かして昨年末 12 月より「楽々スクール」をスタートさせた。10 日間試用期間が設けられていて、オンライン学習の実験を試すことが出来るが、特色として気づいた点を挙げると 1) 具体的目標として HSK (漢語水平考試) のスコア向上が掲げられ、実際に HSK のオンライン模擬試験を受けて実力を確かめられる。2) 北京語言大学の編集になる信頼のおける教材を提供、は他の学習サイトと比しても優れた点であることは間違いない。ただ、一年間で全ての学習を終えるには少々負担が大きいと感じられるのも事実である。発音判定機能も含め、様々な工夫が凝らされているだけに、価格 (非公開: 学校団体と個人で異なる) 設定も含め、もう少し利用しやすいように、入門編、初級編、中級編といった形で細かくレベルを設定してはどうだろうか。

●北京語言大学 (日本語)

<http://www.eblcu.net/newjapan/index.html>

北京語言大学は学部四年制の一部科目をオンライン化して、ネット上での学習でも取得可能となっている。現状では一部分だが、ほとんどの科目がオンライン提供されるようだ。一部卒業論文など北京での研修が不可欠の科目もあるが、学部四年間のうち大部分の期間を日本で履修し、一部を北京で履修さえすれば、大卒の資格認定を受けられる。これは相当衝撃的である。だが、履修時間数を見れば、履修時間の負担量も半端ではない。例えば「初級中国語総合」は 22 単位で想定時間数が 880 時間とある。これだけの時間を一人モニタに向かい中国語学習できるか相当疑問である。



現状での受講生のレベルについては全く知るすがないが、既習者扱いで一部科目の履修を免除された中級レベル以上の学習者でない限り、受講は現実的とは言えないだろう。北京語言大学で提供する教材のうち、初歩の教材だけならば搜狐 (<http://www.sohu.com/>) が運営する下記のサイトで利用できる (ユーザ登録は必要)。

●Just Talk (日本語)

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

このほかにも有料学習サイトは英語圏向けを中心に多数存在し、無料コンテンツも含まれているので、一見の価値はあるが、上記の圧倒的なコンテンツと比べると見劣りするのはやむを得ないだろう。

●学中文 (GB)

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

国家对外汉语教学领导小组办公室 (国家汉办) がコンテンツを提供している。英語圏向けの中国語学習サイトで、漢字の書き方、基礎会話、旅行会話内容など、初歩的な内容が中心。無料コンテンツでも登録が必要。ほかにも類似したサイトとして、次のサイトが挙げられる。これだけ類似サイトが乱立すると、もはや細かく論評するゆとりがない。

●中国上网 (GB)

<http://www.chineseon.net/main.php>

●空中汉语 (GB)

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

最後に、北京師範大学作成になるコンテンツを提供する学習サイトを紹介する。

●“你好, 汉语”教学网站 (GB)

<http://www.hello-han.com/index.php>

Flash で作成された教材は一部未完成の部分もあるが、全て無料で利用できる (ただし、登録が必要)。音声もクリアで十分実用に耐える内容で、中級レベル



の学習者ならば活用できるだろう。将来的には北京語言大学に追従して学歴教育を狙っているのかもしれない。

□ 華人向け漢語教育サイト

ここまで紹介してきた学習サイトは基本的に日本人向け、もしくは外国人向けであるが、最近までオンライン学習サイトといえば、華人子弟向け学習サイトが中心であった。その意味では老舗とも言えるサイトが多いが、新規発足のサイトもあり、この分野も需要旺盛のようだ。

● 全球華文網絡教育中心 (Big5)

<http://edu.ocac.gov.tw/>

これまで華僑向けの国語教育を担ってきた僑務委員會がオンライン教育のために設置したサイトであり、これまで蓄積してきた教材を Web 化したもの。Flash Movie を多用した教材は全て無料公開され、ユーザ登録なしで利用できる。最近、高校向けの国語教材も追加された。教科書研究の素材としても貴重。海外で教材が入手困難な地域からの要望を反映したものだそう。



● 大華風采 (GB)

<http://www.better-c.com.cn/happy.asp>

幼児向け北京語教材。香港と北京に拠点を置き、海外や香港での幼児向け北京語教育教材を提供する。香港では今や北京語学習は必須となりつつあるが、どうせ習うならば幼いうちからの方が良いだろう。「香港回歸」の影響が子供の教育にも及んでいる。無料教材も一部あるものの、ほとんどは有料教材。サンプルを



見た範囲では、歌あり、アニメありで、なるほど子供向けと思わせる教材が多い。

● 讲华语运动 (GB)

<http://mandarin.org.sg/>

推廣華語理事會による讲华语运动。この一年で内容は急速に非常に充実してきており、台湾と双璧をなしている。台湾と違って、会話や文法だけでなく、笑い話あり、学習ゲームあり、中国文化紹介ありのコンテンツは「楽しく学ぼう」というスタンスが貫かれている。このところ新しい教材の追加は見られないようだが、内容は豊富。

● 对外汉语教学信息 (GB)

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

外国人向け中国語教育の情報を提供するサイト。この原稿執筆でも随分と参照させて頂いた。このサイトの運営主体が前掲の国家对外汉语教学领导小组办公室 (国家汉办) である。1987 年から存在していた組織だが、実質的に動き出したのは 1998 年であり、近年活発に活動しているようだ。2001 年からは对外汉语教师资格考试を実施^[4]、2002 年 1 月には「高等学校外国留学生汉语语言专业教学大纲」を制定して、大学ごとに異なるカリキュラムに最低履修条件を提示するなど、留学生教育の質の確保に力を入れている。このほか海外世界各地で「汉语桥」の名前で弁論大会を開催しているのも国家汉办である。妙な盟主意識を持ちやしないか、少々心配ではある。

注

- [1] 本誌創刊号「語学教育・CAI 関連サイト」及び「インターネット学習サイト」(『日本の中国語教育——その現状と課題・2002』好文出版所収)で言及してきた。

- [2] 一般に LMS といった場合、企業内教育で人事管理部門がシステムティックに個々の社員の学習進度の管理及び人事考課への反映を行うためのシステムであり、そのようなシステムをストレートに持ち込むことまでは意図していない。
- [3] 現在は Flash への移行が進んでいる。HTML よりも動画、音声をより最適化した形で提供し、著作権保護の

ため素材をパッケージ化しやすいためである。

- [4] 当初は受験資格が厳しすぎたためか（教学経験必須）、今年からは条件が緩和され、大学卒業見込みでも受験可能になった。

補記：サイト選定に当たり岩野忠昭氏より貴重なアドバイスを頂戴した。ここに記して感謝する。

中国古代史

山田 崇仁

□ はじめに

『四部叢刊』『四庫全書』『基本古籍庫』『基本典籍庫』『國學宝典』…。近年中国から大規模な中国古典テキストコーパスが続々と発表・発売されるようになってきた。市販 CD-ROM 以外にも、香港中文大学の漢達文庫 (<http://www.chant.org/>) では独自校訂の古典と出土資料のテキストコーパスが有料で公開されている。この分野の老舗である漢籍電子文献や寒泉については今更言うまでもないだろう。

またテキストが電子化されていなくても、超星数字図書館 (<http://www.ssreader.com/>) のようなデジタル画像化された書籍が入手可能な場合も少なくない。研究へのコンピュータやインターネット利用を避けていては、研究を行うための材料の収集の速度と量という面で決定的な遅れを生ずるようになってきたのだ。

□ とりあえずググっとけ！ ——出土資料を例にして——

インターネットで調べものをする際、サーチエンジンを利用する事が多いだろう。その中でも Google (<http://www.google.co.jp/>) はロボット型サーチエンジンの代表格である。「Google で調べる」という意味の「ググる」という動詞？が通われるくらいだから、「インターネットで調べもの＝ググる」となっている方も多いだろう。

中国に関する情報を調べるにも「ググる」作業は欠かせない。特に現代中国については、インターネット上に公式からアンダーグラウンドに至るまで洪水のご

とく多量の情報が流れているので、日々これを押さえるだけでも大変な作業になるだろう。これら膨大な情報へのナビゲータとして、中国学にも Google は欠かせない存在となっているのである。

実はこの現象は何も現代中国情報に限った話ではない。一見無縁の観もある考古学の分野でも「ググる」事は重要な研究作業の一環なのだ。

従来、出土資料や考古学的発掘の情報を得るためには、考古学専門の新聞もしくは雑誌から入手するしかなかった。この状態は基本的に今も変化がないが、昨今では中国の Web サイトに掲載されている情報が無視し得ない存在となってきたのである。

例えば、昨年末に 1,000 枚以上の竹簡が発見された湖北省九連墩楚墓については、湖北日報の特集サイト「九連墩古墓之謎」(<http://www.cnhubei.com/9ldgm/index.htm>) がほぼリアルタイムで情報を発信していた。

図 1：九連墩古墓之謎より



また、青銅器に鑄込まれたのべ数千文字の金文で学会を驚かせた陝西省眉県出土の西周墓についての情報も、今年の初めには陝西省の地方新聞の Web サイトで多くの画像あるいは金文の釈文、更には著名な考古学者の小論などが掲載されている。

更に言えば、今日新出出土文字資料の釈文が Web スペースで真っ先に掲載される事など珍しくもない。それらを踏まえた Web 掲示板でのやりとりや Web サイトへの論文掲載も活発に行われている。前号でも紹介した簡帛研究 (<http://www.jianbo.org/>) はその代表格であり、そこでは Web サイトの特徴を利用して膨大な量の論述が掲載され、中には著名な研究者がここだけで掲載しているものまで存在するため、現在では雑誌媒体と同じくらい無視し得ない存在となってきたのだ。

Google は様々な分野の情報を探するためのナビゲータとなりえる。これからは中国学の素養の一つに「いかに上手くサーチエンジンを使いこなすか」が含まれてくるのかもしれない。

□ 日本中国学は何をしているのか？ —— Google の検索結果からの嘆き——

■ 専門家からの情報発信の少なさ

Web サイトに見られる中国古典の情報は、中華人民共和国発のそれが圧倒的である。対して日本の中国学の情報発信といえば中国学愛好家のそれが圧倒的であり、そのエネルギーは瞠目に値するものである。例えば絶版書籍の復刻を目的とする復刊ドットコム (<http://www.fukkan.com/>) には中国学関係の絶版書の復刊願いが多く掲載されているし、「無いものは作ってしまえ！」とばかりに『晋書』の翻訳に取り組む Web サイト「解体晋書」(<http://www.jin-shu.com/>) さえあるのだ。

一報「研究者」側からの情報発信は実に微々たるものである。大学の所属研究室紹介での研究業績一覧の掲載でもあればましな方で、何らかの形で研究業績公開などは滅多に見かけない^[1]。これら大学の中国学関係 Web サイトは、少数の大学院生に実質的な運営を任せているのが実情だろう。継続して Web サイトを更新し続けてサイトに厚みと奥行きをだす事は多

大のエネルギーを必要とするため、この形態を採る限り、担当者の入れ替わりで Web サイトが陳腐化する事態が避けられない。この辺りが日本中国学発の専門的情報が少ない理由だと思われる

それでも大規模な研究機関ともなると、注目すべき業績を公開しているのである。東京大学東洋文化研究所 (<http://www.ioc.u-tokyo.ac.jp/>) や京都大学人文科学研究所^[2] (<http://www.zinbun.kyoto-u.ac.jp/index.html>) の東西二大研究機関は独自のデータベースが充実している。どの分野でも独自のコンテンツとスタッフを持つところが圧倒的に有利なのだろうか。

□ Web サイトの紹介

今回は冒頭に考古学関連の話を書いたので、考古関連の Web サイトを中心に、筆者が日常利用する中国古典関係便利サイトを紹介しておこう。

■ 調べものに便利な Web サイト

● ネットで百科@ Home

<http://ds.hbi.ne.jp/netencyhome/index.html>

平凡社『世界大百科事典』の Web 版。中国学に関しては『アジア歴史事典』よりも新しい知見が反映されているのでお勧め。原則として有料だが、三分間の無料お試しサービスがある。

● Japan Knowledge

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

こちらは小学館『日本大百科全書』など、複数の辞書を串刺し検索可能な Web サイト。有料。

■ 定番データベースの状況

● 全国漢籍データベース

<http://kanji.zinbun.kyoto-u.ac.jp/kanseki/>

収録される研究機関が増えた。東大東文研・京大人文研の所収漢籍が同時に検索できるのは便利。詳しくは本誌小島氏のレビュー記事を参照のこと。

● 東洋学文献類目検索

<http://www.kanji.zinbun.kyoto-u.ac.jp/db/>

CHINA3/

収録データの拡大やデータベースの Unicode 対応を積極的に進めている。将来的には CHISE への移行も視野に入れていると思われる。上の漢籍データベースと並び、日本中国学が誇るデータベースであるとい

えよう。

●漢籍電子文献

<http://www.sinica.edu.tw/~tdbproj/handy1/>

本年当初より検索システムバージョン2が公式の検索システムとなった。バージョン1は廃止される旨アナウンスがあったが、執筆現在ではまだ廃止されていない。また、バージョン2の無料エリアでは利用人数の制限を設けてあるようで、時々アクセス過多で利用できない場合があった。

■考古学

●中華人民共和国 国家文物局 (GB)

<http://www.sach.gov.cn/>

中国考古学関連の総元締。最新の情報を入手すると言うよりも、中国の考古学の情勢を知るために利用するサイト。

●中国考古 (GB)

<http://www.kaogu.cn/>

中国社会科学院考古研究所のWebサイト。最新情報から各種データまで幅広く取り扱う。

●中国文物信息网 (GB)

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

中国考古学関連の新聞「文物報」のサイト。適宜新着の考古関連情報が掲載される。

●「中国網日本語版」の考古関連ページ

<http://www.china.org.cn/japanese/>

日本語で書かれているので斜め読みには便利。ただし、固有名詞を中心に誤訳が多いので注意。

●殷墟 YINRUINS

<http://jgw.ayinfo.ha.cn/>

殷墟および甲骨文に関するWebサイト。

●中国甲骨文獻庫 (GB)

<http://www.cn-oracle.com/>

甲骨文に関する研究論文を原載誌からスキャンした画像を掲載する。中国はもとより日本人の研究論文まで載せている。

●先秦諸子 要覧子籍 (Big5)

<http://hk.geocities.com/chinpcp/>

公開された出土資料の中から、春秋戦国～秦に至る範囲のそれを中心にデジタルテキスト化して公開する。Big5 未収録字は外字で対応する。著作権をどうクリアしているのかは不明。



図2：中国考古トップページ

■ デジタルライブラリー

●国立国会図書館 近代デジタルライブラリー

<http://kindai.ndl.go.jp/index.html>

明治期の図書を画像化して公開する。中国学関係の書籍もかなりの数が公開されている。本サイトについては、本誌小島氏のレビュー記事を参照のこと。

●超星数字図書館

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

「質より量」のデジタルライブラリーの横綱格。近代デジタルライブラリーとは好対照。しかし数十万冊の圧倒的な量が有無をいわせぬ存在感を見せつける。

筆者などは、図書館に目的の本がなかったらとりあえずここで探すのが習慣になってしまった。

この二つのデジタルライブラリーは、一点一点の質的向上を求める日本と、とりあえず量を充実させる事にターゲットを絞った中国とのお国柄の違いがよく表しているのかもしれない。

□ 最後に——インターネット上の日本中国学の厚み——

本研究会発行のメールマガジン第31号(8/1発行)掲載の二階堂氏によるコラム「早くも来てしまった「本よりネット」時代？」にこんな記載があった。

しかし、ネットの資料を見ていると、どうもその中間をつなぐ、概説的なものや翻訳については、あまり充実していない気がします。

中国古典学の分野から見てみると「専門家向けの

内容」「愛好家向けの内容」が両極端と言えるだろう。先に述べたとおり日本発の中国学向けの Web サイトは「愛好家向けの内容」が圧倒的であり、「専門家向けの内容」というのは微々たるものである。また、二階堂氏が指摘する「中間層」的な情報発信についてもまず見られないだろう。そういう視点からすると、Web サイトにおける日本中国学の「厚み」というのは非常に薄っぺらであり、特に上層部の薄さたるや深刻なほどである。

筆者は自分の専門ではない事象について調べる場合、とりあえず「ググる」というのが身に付いている。中国学の専門家による情報発信の薄さを目の当たりにすると、他分野の専門的だと思われる情報が果たして本当にそうなのかどうか疑問に思ってしまうのである。これが単にどの分野でも同じであればインターネットの特性として理解すればよいが、中国学の特有であるとすれば大問題ではないか。

これからは研究者たるもの、自らの発表した研究を公開あるいは概要を積極的に公開する時代に来ている

のだろう。これは自らへの反省でもある。今後は研究機関あるいは研究室発行雑誌などが気軽に Web 上で閲覧できるところが増えてくるのを希望したい。

専門分野で「ググってみた」ら中国人の研究しか引っかけこないのはある意味悔しい。インターネット上でも日本中国学の健在ぶり見せつけないと、いずれ世界から相手にされなくなるはずだ。

注

- [1] 研究室の紹介程度しかない Web サイトは、そのページがどこに向けて発信された情報であるかを如実に表している。それは研究者の自己アピールと言うよりも、大学入試対策の一環としての情報発信なのである。
- [2] 京大人文研は 21 世紀 COE プログラム「東アジア世界の人文情報学研究教育拠点（漢字文化の全き継承と発展のために）」の採用が決定している。今後どのような動きが見られるのか要注目だろう。

中国近現代史

佐藤 仁史

□ はじめに

本誌 2 号において筆者が本欄を執筆した際には、国内外におけるデジタル・リソースの構築は緒についたばかりであった。それ以来の 2 年間に、研究活動における実質的な利用に耐えうる本格的なデジタル・リソースが続々と公開されるようになり、研究テーマによっては研究活動のスタンスに大きな影響を与えるようになってきている。膨大な一次史料の全貌を把握することが困難な中国近現代史における研究情報のデジタル化には他の分野と異なる方法が必要であるが、その一つの解答を示したのが台湾における大規模な一次資料データベースである。最近のブロードバンドの普及や保存媒体の大容量化に伴って大量の画像データをやり取りすることが容易になったため、これらの多くは史料の画像データの方式によって構築されている。以下に紹介する主要なデータベースでは画像データ自体の対外公開は依然として限定されているもの、詳細

な書誌情報の検索が可能で、史料収集の効率が飛躍的に向上している。

これら以外にも、国内外において史料や研究文献の目録、種々の研究情報のデジタル化が確実に進んでいる。以下では、台湾における一次史料デジタル化の進展、一次史料、図像資料、史料所蔵目録、研究文献目録、研究情報・その他、という分類に即して、意欲的な作業を進めているサイトを紹介していきたい。なお、多くのサイトは本誌 2 号の拙稿において既に紹介しているが、有用だと思われるものは再び取り上げている。また、筆者の専門の関係上、近代史のものを中心に取り上げることをご寛恕頂きたい。

□ 台湾における一次史料デジタル化の進展

● 臺灣總督府公文類纂查詢系統 (Big5)

<http://db1.sinica.edu.tw/%7ETextdb/sotokufu/query.php>

國史館臺灣文獻館（旧臺灣省文獻委員會）が所蔵す

る膨大な檔案史料のうち、植民地統治期のものには「臺灣總督府公文類纂」「臺灣總督府專員局公文類纂」「臺灣拓殖株式會社文書」がある。このうち前二者を、中央研究院との協力によってデジタル化したものが本サイトである。内容は詳細な書誌情報と史料の画像データとで構成されている。日本から利用できるのが前者の機能で、機関名、西暦、日本の年号、文書の名称などから必要な資料を絞り込むことが出来る。植民地統治機構の文書という性質上、検索結果の文書名に日本語が含まれる場合があり、その部分を Unicode（実質は実体参照形式）、あるいは画像で出力することを選択できる。

また、外字については「視覚輸入」という、外字画像の一覧から選択する方式が採られている。史料の画像データは現在のところ中央研究院をはじめとする限定された機関からのみ閲覧が可能である。全面的な対外開放について現在討論が行われているという^[1]。もちろん、全面開放が実現するに越したことはないが、現在の方式でも日本において必要だと思われる情報を事前に絞り込むことで現地における収集が効率化できるので、極めて有用であることに変わりはない。

●国立故宮博物院圖書文獻館資料庫 (Big5)

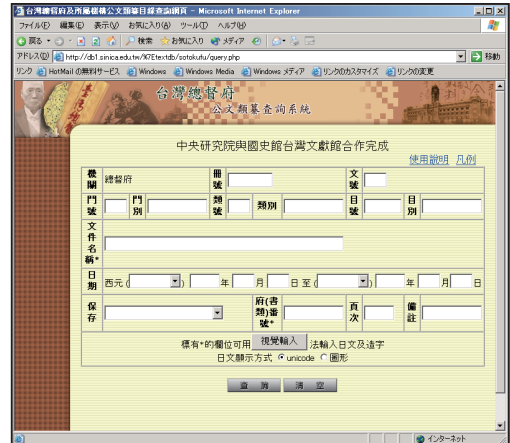
<http://www.npm.gov.tw/tts/npmmeta/dblist.htm>

同館が所蔵する種類の資料をデジタル化したデータベースである。とりわけ目を引くのが「軍機處檔案資料庫」であり、ここでは清初から光緒朝までの時期の、14 万件あまりもの漢文奏摺録副の目録を検索することができる。検索結果は国立故宮博物院圖書文獻館において、複製本を複写するか、画像ファイルからプリントアウトすることができる。同館における奏摺録副の所蔵状況や利用方法の詳細については堀池明氏による解説があるのであわせて参照されたい^[2]。

●内閣大庫檔案 (Big5)

<http://saturn.ihp.sinica.edu.tw/~mct/newpage1.htm>

中央研究院歴史語言研究所が所蔵する 31 万余件の内閣檔案の画像データベースで、13 万件余り（2002 年 9 月 1 日更新分）のデータ量を誇る。国外からは「文字目録検索」という書誌情報が利用でき、上奏文の具奏者、具奏者の官職、上奏内容などから条件に符合する檔案名を知らせてくれる。檔案の画像ファイルは同所傳斯年圖書館閲覧室より接続が許されている機関から「全文影像検索」にアクセスし、そこで検索・閲覧



臺灣總督府公文類纂目録查詢系統の検索画面

をすることができる。「全文影像検索」は台湾国内ユーザーに限定されている。

●漢籍電子文獻 (Big5)

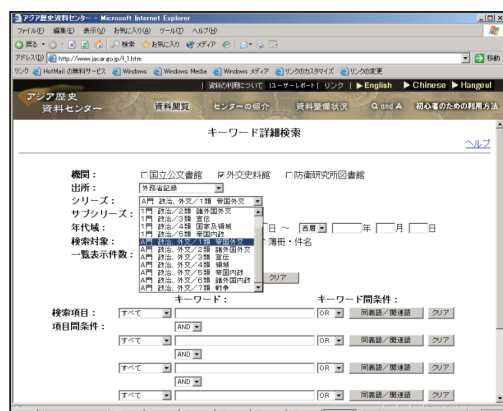
<http://www.sinica.edu.tw/ftms-bin/ftmsw3>

電子テキストの老舗サイトである。中国近現代史料の殆どが画像データであるのに対して、当該サイトは異彩を放っている。近現代史に関連する内容では、近年の文化政策や台湾アイデンティティの高まりを反映して台湾史研究に関する充実ぶりがひととき目を引く。台湾の地方志の全文テキストである「臺灣方志」、清実録、東華録、奏摺、各種文牘から台湾関係記事を抽出した「臺灣檔案」、文集を中心とする「臺灣文獻(一)～(五)」などからなり、データの追加が継続されている。

●國史館 (Big5)

<http://www.drnh.gov.tw/>

國史館所蔵の各種檔案も画像データベースの作成が始まっている。上述の「臺灣總督府公文類纂目録查詢系統」以外にも、民国 15 年から 49 年までの 7,000 巻あまりに及ぶ「國民政府檔案」の構築がほぼ完成し、現在書誌情報については試用が開始されている。また、國史館には「蔣中正總統文物」「資源委員會檔案」「臺灣省地政處檔案」などの膨大な檔案が所蔵されており、これらは民国期や台湾における政治や行政の実態に関する貴重な情報源である。このうち「蔣中正總統文物」についてはオンラインでの一部公開が予定されており、また「資源委員會檔案」「臺灣省地政處檔案」についても、画像データの取り込みが進んでいる。これらが完成した暁には近代史研究に裨益するところは計り知れない。



キーワード詳細検索の画面

□ 一次史料

● 国立公文書館アジア歴史資料センター

<http://www.jacar.go.jp/>

アジア歴史資料センターは国の機関が所蔵しているアジア関係歴史資料をウェブ上で公開する電子資料センターとして、2001年11月30日に開設された。目下公開が進んでいるのが、国立公文書館、外務省外交史料館、防衛庁防衛研究所図書館が所蔵する明治初期から太平洋戦争終結までの資料をオンラインで検索・閲覧できる「アジア歴史資料データベース」である。現在はデータの追加作業の最中であり、2003年6月27日段階における公開総数は約32万件、約370万画像である。最終的には2700万画像を超えることが予定されており、完成の暁には、アジア近現代史の分野では他に類をみない巨大なデータベースが出現する。

しかし、このデータベースにも問題がないわけではない。三品英憲氏が指摘するように^[3]、史料のデジタル化・データベース化が進められているのは国立公文書館、外務省外交史料館、防衛庁防衛研究所図書館の3機関であり、他の機関の史料がどのように扱われるかについては依然として不明である。また、仮に国の機関が多く参加したとしても不十分である。例えば、外務省外交史料館史料の「K門内外人外国在留、旅行」によって租界や植民地における邦人の状況を把握することができるが、これ以外に、邦人団体の活動を示す史料や引き揚げ時の史料といった彼らに密着した史料群も参照する必要がある。これらは密接な関係にあるにもかかわらず、片方は公文書として、片方は

そうでない私的な史料として別々に扱われてしまうのである。史料のデジタル化の進展はデジタル化が進展しやすい史料とそうでない史料との差を大きくするだけに、デジタル化に対応した史料論を検討する必要があるだろう。

なお、当該サイトは独自の方式によって構築されているので検索や閲覧には注意が必要である。目録は、文書の先頭からおおよそ300文字程度が収録されており、原文の一部も検索対象になる。また資料画像はDjVuという画像フォーマットで処理されているため、閲覧にはDjVuビューアーのプラグインをダウンロードする必要がある^[4]。

● 超星数字图书馆 (GB)

<http://www.ssreader.com.cn/pdg.html>

中国における学術情報デジタル化の注目すべき取り組みとして近年多くの注目を浴びているのが当該サイトである。中国近現代史の領域に限定しても、民国叢書や『光緒朝硃批奏摺』『汪偽政府行政院會議録』といった檔案資料、『北洋政府公報』をはじめとする公報類などの膨大な一次資料を利用することができるのは瞠目に値しよう。しかし、最近になって著作権の問題や経営の安定性をめぐって多くの問題点が露呈することになってきた。これらについては、千田大介氏が本誌3号において詳述しているので^[5]、本論では贅言をさけ地方志について言及するにとどめたい。当初より方志叢書収録の地方志が公開されていたが、後に新編地方志も追加された。日本では東洋文庫や明治大学図書館が新編地方志を意識的に収集しているものの、現地に赴かなければ閲覧できないものも少なくない。したがって、新編地方志のオンライン化は極めて有用であったのだが、最近半年ほどアクセスできない状況が続いている。有償サービスである以上、予告や説明もなく内容を変更してしまう事態は運営側のユーザ軽視ととられても仕方ない。また、有償サービスという形態をとってはいるものの、国家のデジタル図書館のプロジェクトによる位置づけもなされている。図書館が有する公共性を一定程度担う存在であり、公共性のあり方やそれをどのように担保するのかについての意識に基づいた運営が必要であろう。

● 赛尔—普来得读书学习卡网上图书馆 (GB)

<http://cn.poledu.com/cernet-poledu/main/index.jhtml>

教育部や清華大学、北京大学などが共同出資し、商

務や IT を中心にするオンライン教育の普及を目指して設立された賽尔网络が普来得との協力によって構築したサイトである。「网上图书馆」には超星数字図書館で公開されている書籍の画像データがある。超星では消えてしまった地方志のデータが利用できるのは便利である。利用には 108 元の「赛尔—普来得学习卡」の購入が必要で、1 年間有効である。

● 文史資料图书馆 (GB)

<http://www.ssreader.com.cn/zhuanti/wszl/index.html>

中国各地で編纂された文史資料を網羅的に収集し、画像データを公開しているサイトである。全国政協文史資料研究委員会によるテーマ別のもの以外に、市・県レベルの全国政協文史資料研究委員会によって地域毎に編纂された文史資料も多数収録されている。文史資料には民国期や建国初期を経験した古老の回想録や聴き取りが多数含まれており、調査された背景を慎重に踏まえれば口述資料に準ずる価値を持つものもある。これらの資料は新編地方志の編纂に際して参照されたが、地方志に採用されなかったものも多く、ここから貴重な情報を得ることも可能である。文史資料は 1980 年代には内部資料として発行されており、90 年代のものでも関係者の間にわずかに流通しただけで、現地に赴かなければ入手できないものも多く、オンラインで公開された意義は大きい。

● 中国関係電子テキスト・アーカイブ

<http://www.zinbun.kyoto-u.ac.jp/db/archive-china.html>

京都大学人文科学研究所の下位ページにおいて公開されている。梁啓超の『变法通議』『自由書』『新民説』の Unicode テキストデータが公開されている。

□ 画像資料

● 番社采風圖 (Big5)

<http://saturn.ihp.sinica.edu.tw/%7Ewenwu/taiwan/index.htm>

中央研究院歴史語言研究所の下位ページに公開されている。同所が所蔵する 18 世紀先住民の風俗を描いたもの。

● 臺灣老照片數位博物館 (Big5)

<http://www.sinica.edu.tw/photo/>

台灣國立藝術學院が収集した台湾の歴史や生活・風

俗に関連する写真を紹介するサイトである。写真を用いた学習のためのインタラクティブな工夫も施されている。

● 平埔文化資詢網 (Big5)

<http://www.sinica.edu.tw/~pingpu/>

平埔族やその文化に関する研究の成果が中央研究院のサイトの一部として公開されているものである。研究文献、契約文書、歴史地図、写真、音楽など音声資料、映像、現地調査の記録などの諸資料が豊富にある。しかし、どのような立場や視角から平埔族の文化が「構築」され、展示されているのかが気にかかる。

□ 史料所蔵目録

● 近史所檔案館藏函目彙編查詢系統 (Big5)

<http://db1.sinica.edu.tw/~textdb/archives/archives.html>

中央研究院近代史研究所檔案館所蔵のオンライン檔案目録である。内容は、従来の『外交檔案函目彙編』『經濟檔案函目彙編』に加えて、新たに『中外地圖函目彙編』が加えられた。『外交檔案函目彙編』では、総理各国事務衙門、外務部、外交部の清末から民国 27 年までの檔案を検索できる。『經濟檔案函目彙編』が対象とする檔案は商部、農商部、実業部、經濟部などをはじめとして清末民国期全般にわたり、機関によっては 1950 年代のものも含まれている。『中外地圖函目彙編』では民国期の 1 万 3 千幅の地図を、省や水路を単位として検索することができ、様々な領域での活用が期待される。

● 旧植民地関係資料総合目録検索

<http://mokuoku.biwako.shiga-u.ac.jp/eml.asp?mode=manmou>

滋賀大学経済経営研究所が所蔵する満洲・蒙古・支那・朝鮮・台湾関連の資料目録をオンライン公開したものである。現在史料の約 3 割がマイクロフィルム化されている。旧高等商業学校が収集した旧植民地関係資料は、現地史料の発掘に比べるとその注目度が高かったとはいえない。これに加えて、史料の劣化や管理体制の原因によって十分に整理・公開されていないばかりでなく、廃棄の危機にさらされている場合もある。かかる史料の整理・利用によって史料の財産としての価値を高め、広く認知されることは、研究環境を整備することにもつながる^[6]。そうならば、近現代

史研究における日本と東アジアとの関係を考えると避けては通れない課題である旧植民地研究の活性化が期待できよう。さしあたり、「旧植民地関連資料聯合目録」の構築から着手してみてもはどうだろうか^[7]。

●全国漢籍データベース

<http://www.kanji.zinbun.kyoto-u.ac.jp/kanseki/>

京都大学人文科学研究所、東京大学東洋文化研究所、国立情報学研究所が幹事機関を務める全国漢籍データベース協議会による、全国漢籍データベースである。UTF-8 ベースで構築されており、多言語表示が可能となっている。

●東洋文庫所蔵書籍オンライン検索

<http://www.toyo-bunko.or.jp/>

国内屈指の東洋学史料所蔵機関である東洋文庫の漢籍目録をオンライン化したものである。中国近現代史の分野では、方志、家譜、文集などの漢籍を検索できる「漢籍資料オンライン検索」に加えて、近代中国研究班が収集する、中国文逐次刊行物や中文図書も逐次情報が更新されていて有用である。

●中国絵画所在情報データベース

<http://cpdb.ioc.u-tokyo.ac.jp/>

東京大学東洋文化研究所東アジア美術研究室が調査してきた、中国大陸の所蔵機関・個人及び臺灣故宮博物院を除く世界中の全ての地域に存する中国絵画の所蔵目録である『中国絵画総合図録』正編 5 冊、続編 4 冊のうち、正編をオンライン化したもの。

□ 研究文献目録

●中国学术期刊网 (CNKI) (GB)

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

中国の学術情報については、清华同方光盘股份有限公司が運営する CNKI (中国知识基础建设工程) によって得ることができる。ここでは、中国学术期刊 (光盘版) 電子雑誌社が出版する各種データベースの一部をオンラインで利用することができる。「中国期刊全文数据库」では 1994 年以降の学術論文情報を無料で検索できる。論文や重要新聞の記事の全文検索・ダウンロードは機関契約や CD-ROM の購入が必要となる。

●國家圖書館中華民國期刊索引影像系統 (Big5)

http://www2.read.com.tw/cgi/ncl3/m_ncl3

1991 年以降に台湾や香港、マカオで出版された中文・欧文の期刊 3,000 種類に掲載された、30 万件に

も及ぶ論文を検索できるデータベースである。

●全國博碩士論文摘要検索系統 (Big5)

<http://192.83.186.1/theabs/01/>

台湾国内における碩士・博士論文の書誌データベースである。内容にばらつきがあるものの、一般的な書誌情報以外に摘要や目次、参考文献目録が附されており、利便性を高めている。博士論文はいうまでもなく、碩士論文の中にも数多くの創造的・開拓的な仕事が含まれており、そのような情報を海外からでも簡単に把握できる。

●法制史文献目録

http://wwwsoc.nii.ac.jp/jalha/bksrch_e.htm

法制史学会のウェブページで公開されている文献目録であり、1990 年以降の法制史関連文献 1 万 9000 件のデータが検索できる。また、「法制史研究総目次 (1951 ~ 2001 年)」「法制史学会総会・研究大会・部会報告題目一覧 (1949 ~ 2002 年)」を加えた全 2 万数千件の一括検索もできる。

●東洋学文献類目検索

<http://www.kanji.zinbun.kyoto-u.ac.jp/db/>

CHINA3/

現在は Unicode に対応した第 4.6 版である。

□ 研究情報・その他

●中央研究院歷史語言研究所生命醫療史研究室 (Big5)

<http://www.ihp.sinica.edu.tw/%7Emedicine/>

近年注目を集める医療、衛生、身体といった研究分野に関する研究情報が充実している。資料については「内閣大庫檔案中有關疾病醫療資料庫」があり、また、研究文献については、「身體史文獻書目」「《黃帝內經》研究文獻書目檢索資料庫」、医学や薬学に関する国内外の文献目録である「自建書目資料庫」がある。

●金陵俱樂部

<http://homepage2.nifty.com/z-gaot/>

明治大学文学部高田幸男氏の個人サイトである。南京や江南地方の歴史資料や史跡情報が詳細に紹介されている。特に「ZaoGao 日記」には江南各地の所蔵機関の訪問記録が含まれており貴重である^[8]。

●東洋史リンク集

<http://www.okayama-u.ac.jp/user/le/toyoshi/>

aoki/

大阪大学文学部青木敦氏のウェブページの一部である。有用なページが纏められており、また簡潔なコメントも附されている。

□ おわりに

以上見てきたように、中国近現代史の分野における研究情報のデジタル化は、史料所蔵機関や研究機関が多く資金を費やしたプロジェクト方式によって担われている。このようなプロジェクト方式によるデジタル化が進んでいるのに対して、かならずしも文章に反映されるとは限らない個人の研究活動の情報や成果のデジタル化はほとんど進められていない。例えば、現在では現地の図書館や檔案館などにおいて史料収集を行い、そこで独自に入手した史料を用いて研究をすすめることがごく当たり前の研究スタイルとなっている。省レベル以上の所蔵機関のみならず、市・県レベルの所蔵機関に赴くことも少なくない。しかしながら、これらの機関の所蔵状況やアクセスの方法、入手した史料などについての情報は十分に交換・共有されてはいないように思われる。研究情報のデジタル化への個人レベルでの寄与は、このような情報の交換、蓄積、共有をする場を意識的に形成していくことがさしあたって最も切実に必要とされているように思われる。

研究情報のデジタル化の進展によってもたらされる情報収集の効率化は、デジタル・リソースの存在を念頭に置いた上での史料方法論への注意を研究者に喚起する。情報が簡単に検索・入手できる状況においては、必要な情報を有効に取捨選択することは、研究者としての嗅覚を日ごろどのように鍛えているのかという点に左右される^[9]。また、キーワード検索によって必要な情報に容易に到達できることは必ずしも諸手を挙げて歓迎できない側面も有している。なぜなら、史料をめくることによって目に飛び込んでくる直接には「無駄」と思われる情報が、背景の理解に役立ったり、思いがけず新たな研究テーマをもたらしてくれたりすることがあるからである^[10]。これらは、徹底的に史料を読み込むことによって会得できる、ある種伝統的な方法の有効性を証明しているといえよう。

注

- [1] 中央研究院日治時期檔案整理小組編『臺灣總督府數位化檔案整理紀要』中央研究院臺灣史研究所籌備處、2002年。当該データベースの閲覧に際しては、中央研究院臺灣史研究所籌備處の林玉茹氏に便宜を図っていただいた。記して感謝したい。
- [2] 堀池明「台湾における清代檔案資料のデータベース化とその利用」『東方』265号、2003年。
- [3] 三品英憲・大沢武彦・大澤肇「アジア歴史資料センターについて——デジタル・アーカイブの可能性と問題——」『現代中国研究』第10号、2002年。
- [4] アジア歴史資料センター設立の背景や、アジア歴史資料データベースの検索方法などについては注3を参照されたい。
- [5] 千田大介「学術リソースレビュー 超星数字図書館」『漢字文献情報処理研究』第3号、2002年。
- [6] このような問題については、2002年12月に滋賀大学経済経営研究所において開催された「旧植民地関係資料をめぐるワークショップ——朝鮮・満州・中国・台湾——」においても討論された。三澤真美恵氏による参加記を参照されたい。
<http://www.soc.nii.ac.jp/jats/archive/misawa2003.htm>
- [7] オンライン化されている旧植民地関連資料の目録には次のようなものがある。
旧山口高商画像目録検索
<http://mickey.lib-c.yamaguchi-u.ac.jp/~cat/>
神戸大学デジタル版新聞記事文庫
<http://www.lib.kobe-u.ac.jp/sinbun/index.html>
- [8] 江南における地方文献の所蔵状況については、高田幸男「中国近現代文書へのアクセス」『歴史評論』第638号、2003年、でも紹介されている。
- [9] このような状況について千田大介氏は新編地方志の事例を紹介している。注5参照。
- [10] 史料のデジタル化がもたらす問題点については飯島渉「インターネットと近代中国研究」を参照されたい。また、藤谷浩悦氏は「中国近代史研究の動向と課題——日本における研究を中心に」において近年の中国近代史研究における未開拓一次史料の重視という傾向に対して既存史料の読み込みの必要性を指摘する。ともに、『歴史評論』第638号を参照のこと。

中国古典文学

千田 大介

□ 文献データ

■ 重点は「近代」文献に

中国古典文献コーパス構築は十三経や二十五史といった基本的な古籍から始まった。現在、唐代以前の文献のコーパス化はほぼ完了しており、構築の重点は中国でいうところの「近代」、宋元明清四代の文献に移行している。未開拓の部分が多い宋代以降の文学研究を重点的に推進している中国の研究政策をも反映した現象なのであろう。

● 全宋詩検索系統

<http://162.105.161.41/songpoem/index.htm>

『全唐詩』に続き、北京大学は新たに『全宋詩』のオンライン検索システムを公開した。『全宋詩』は北京大学が編纂した宋代詩の総集であり、同分野を研究する際の最も基本的な文献となっている。

インターフェースや検索方法は、『全唐詩』とほぼ同じである。『全宋詩』にとどまらず『全宋詞』と『宋史』の全文検索も可能である。このため、宋代詩歌の総合データベースとして利用することができる。

これまで、宋詩のオンラインデータベースとしては、台湾元智工学院網路展書読 (<http://cls.hs.yzu.edu.tw/>) があったが、収録範囲は著名詩人とどまっていた。それだけに『全宋詩』という網羅的な全集が公開されたことは、大きな意味を持つ。

ただ、『全唐詩』検索系統には歴代の詩話が収録されており宋代詩歌に言及する記事も多いが、それを『全宋詩』と同時に検索できないのは不便である。今後、両者の検索システムを統合するなど、より使いやすいシステムを目指して、更新をすすめてもらいたものである。

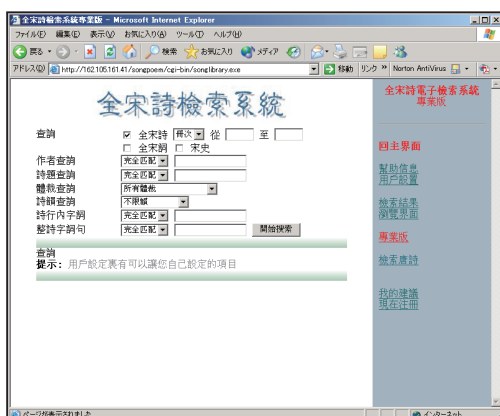
● 漢達文庫

<http://www.chant.org/>

香港中文大の華夏文庫は、南北朝以前の文献を出土文献をも含めて網羅的に収録した大規模古典文献データベースとして知られていたが、2002年に名称を「漢達文庫」に改めた。

名称改変と同時に次期プロジェクトとして『太平御覽』・『永樂大典』などの類書のデータベース化計画がアナウンスされている。南北朝以前の文献のデータベースである漢達が、宋代以降の類書をデータベース化するのは、南北朝以前の逸書の引用が多いためだという。しかし、完成の暁には、広く中国学研究者にとって有用なデータベースとなろう。

ところで、漢達は相変わらず有償データベースとして提供されており、しかも個人契約の場合でも年間250米ドルという高額である。有償サービスは超星数字図書館のように年間100中国元前後程度のものが多く、台湾中央研究院をはじめ無償のデータベースが増えている中で、漢達の価格の高さは際だっている。是非とも無料開放、もしくは値下げを検討してもらいたいものだ。



● 国立国会図書館近代デジタルライブラリ

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

2002年に公開が始まった同ライブラリの文学カテゴリ以下に中国文学のサブカテゴリが立っており、執筆時点で合計90種の書籍が登録されている。その大半は古典詩文の訓点本である。流石に、詩文解釈研究にそのまま生かすことのできるものは少数であろうが、我が国における漢詩文の受容史・研究史の資料として十分な価値を持つ。

なお、同ライブラリの詳細については、小島氏の「図書館とOPAC」レビューを参照して頂きたい。

□ 学会・研究会サイト

■ 進む公開、進め更新

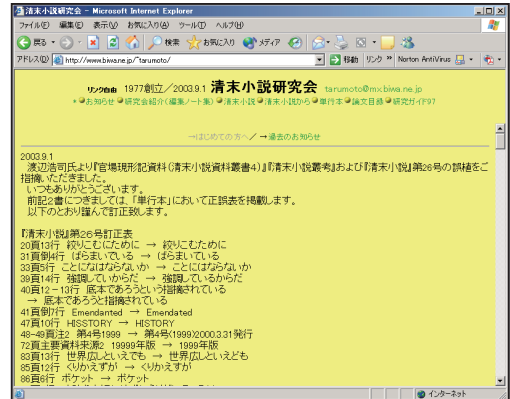
ここ数年で全国学会や全国規模の専門研究会の多くがWebサイトを開設し、もはや「ホームページを開設しました」というだけでは何ら注目を集めなくなった。このこと自体は非常に歓迎すべきことなのであるが、しかし、それらのWebサイトの中には全く更新されないまま「死んでいる」ものも少なくない。

開設当初は旗振り役、もしくは仕事を押しつけられた若手が一生懸命制作し、将来の発展が期待されるものの、旗振り役・管理責任者の離脱やWebサイト維持への意欲喪失によって、Webサイトが死んでしまうケースが多いようだ。

実際、Webサイトは開設するよりも維持する方が、遙かに手間とエネルギーを要する。その点への自覚が無いまま、Webサイト公開＝情報公開の完成であるとの誤解が一部に広まっていることも一因であろう。

また、学会・研究会サイトには、その会の目的・活動の紹介にとどまらず、学術情報の発信を求めたい。大学の中国文学セクションというのは、教育上の必要から所属教員の専門研究分野を分散させる傾向にあり、その意味で、中国文学研究を実際に担っているのはそれらの学会や研究会であるといえよう。そして、学会・研究会の多くは紙媒体の機関誌を通じて研究情報を発信している。その機関誌を、理想的には全文、最低でも目次だけはWeb上に公開していくべきであると考え。

実際、いくつかの研究会では、そのようなWebサイト活用がなされている。



● 清末小説研究会

<http://www.biwa.ne.jp/~tarumoto/>

国内の中国学研究会サイトとしては、ずば抜けた情報量を誇る。機関誌『清末小説』の全文データの他、連絡誌『清末小説から』全文、各種文献目録など、同分野を研究するために必要なデータが網羅されている。この情報公開の姿勢を、他の研究会は大いに学ぶべきであろう。

ただ残念なのは、未だにサイト全体がShift JISコードのみで記述されており、外字が一律に「■」に置き換えられている。今や、UTF-8のHTMLや、Unicode文字の実態参照による埋め込みは簡単にできるのであるから、是非とも対応して頂きたい。

● 中国古典小説研究会

<http://sasa1.misc.hit-u.ac.jp/zgxy/>

会誌『中国古典小説研究』の目次が閲覧できる。また、会報の報告も掲載されている。

● 中国都市芸能研究会

<http://wagang.econ.hc.keio.ac.jp/chengyan/>

筆者がWebサイト管理をしているだけに手前みそではあるのだが、紹介しておく。これまで研究会が発行した科研報告書、機関誌『中国都市芸能研究』の全文PDFを公開している。画像を多用しておりファイルが重いので、HTML化が今後の課題である。

それにしても、学界を代表するはずの日本中国学会サイト (<http://www.soc.nii.ac.jp/ssj3/>) さえもが、執筆時点で一年以上更新されていないというのは、流石にマズイだろう。一学会の問題にとどまらず、日本中国学という領域そのものの軽重すら問われかねないのだから。

中国現代文学

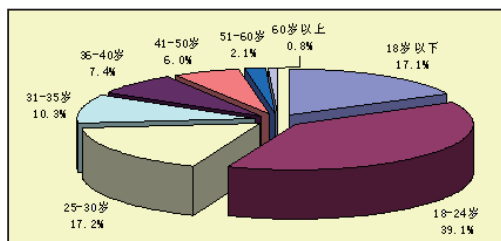
小川 利康

□ SARS 流行とネットユーザ増加

中国互联网络信息中心 (<http://www.cnnic.net.cn/develst/>) によれば、中国のネットユーザは 2003 年 6 月現在で 6800 万人に達し、半年で 890 万人もの増加を記録したという。そのうちダイアルアップユーザが依然として 4500 万人と大多数を占めているものの、ブロードバンドユーザも 660 万から 980 万に急増、自宅利用時間が増加した。このことから、中国のメディアはさかんに SARS 禍がインターネットの利用を後押ししたと見ている。まさしく「風が吹けば、桶屋が儲かる」論理で、実際は両者に因果関係はない。もし SARS の影響が認められるとしたら、利用時間数の増加ぐらいだろう（昨年 7 月 8.3 時間→今年 1 月 9.4 時間→今年 7 月 13.0 時間／全一週間あたり）。これ自体、本当に SARS に由来するのだろうか、ユーザ達自身が懐疑的（SARS による利用時間：増加した 44%、不変 41%、減少した 7.9%）であり、SARS のような一時的な現象で大勢に変動が生じることはないと考えた方が良さだろう。

むしろ特徴的なのは中国のネットユーザの若さではないか。平均的ネットユーザのプロフィールをまとめてみると、年齢は 19 才～24 才（39%）で、学歴は大卒か専門学校で、おおむね自宅からアクセスし（65%）、夜 8 時から 10 時に利用している（40% 超）。ユーザの六割強が男性、女性はやはり少数派だ。日本も若年層に偏っているが、20 代と 30 代を合わせてやっと 40%（2002 年統計による）を上回る程度で、中国での二〇代前半層への集中が如何に極端か分かる。この比率は過去六年間ほぼ不変で常に 30% 後半で推

fig. 1: 中国ネットユーザ年齢構成比 (CNNIC)



移している。

ご記憶の方も多いかもしれないが、昨年末に馬立誠（人民日報評論員）が発表した論文「対日关系新思维——中日民间之忧」はネット上で猛烈な批判を受けた。議論そのものは正常な現象だとしても、ネット上での批判は実のところ批判の名に値しない罵倒であり、暴力的な言論がネット上では現実以上に輻輳されることを今更ながら痛感した。総じて中国語圏におけるネットワークの幾つかの傾向、例えば、感情的暴走、著作権軽視といった負の側面だけでなく、BBS 隆盛に見られるコミュニケーション重視の傾向もユーザの年齢層から説明できると思われる。

その文脈からすれば（やっと本題に入れた…）、昨年号で取り上げた Web 雑誌の繁栄と営利指向のもたらす限界もまた中国のネットユーザの持つ特性に由来するものといえるだろう。その実例として、榕樹下から見て行こう。

□ 榕樹下の変貌

本誌前号で紹介したように、商業性が強いと批判を集めた榕樹下 (http://www.rongshuxia.com/rss/ld_index.rs) が今年新たな変貌を遂げつつある。

その契機となったのは、やはり BOL 中国との業務提携であろう。早くから伝えられていた買収劇に決着が付いたのは昨年末で、創業者朱威廉が榕樹下の身売りを画策と騒がれた。一日 500 万ページレビューを誇るサイトに企業が関心を寄せるのは極めて自然なことでもある。日本市場からは撤退したものの、なんとか中国では足場を強固なものにしたいと考える BOL からすれば、決して高い買い物ではなかったはずだ。

買収後の変貌を告げる第一弾として昨年末、「榕樹下充值卡」が発売された。これは従来の無料サービスに加えて、個人専用掲示板、「以文会友推薦位」（トップページから作品への直接リンクを提供）、「翰林之路」考試（試験によって経験値が上がり、出世する）の三つの機能を提供するもの。説明によると、サービスの一部有料化によって、より発展拡大を目指すというものである。今回の提携により、十分な資金を得た

はずの榕樹下の新機軸が有料サービスとは意外の感もあるが、その内容を見れば、実はポータルサイトの多くが提供するパーソナライズ機能（個人化）に類似しており、いわばアクティブユーザに特化したサービス強化が狙いと思われる。

掲示板への書き込みや投稿によって自分の持ち点を増やし、サイト上での「官位」を上げる運営形態はゲーム感覚でポータルサイトに参加できるため、若年齢層に支持されている。さらに「榕樹下充值卡」利用者向けに RPG ネットワークゲーム「流星学院」（榕樹下専版）の提供も始まった。ここにも現在のユーザの嗜好が反映されている。

もともと中国のポータルサイトのほとんどは文学作品をオンラインで提供するサイトとして始まった。現在はニュース、検索、掲示板機能などが主たるサービスへと変貌を遂げたが、逆にネットワーク文学サイトであった榕樹下がポータルサイトの機能を取り入れ、ユーザの囲い込みを図っても不思議ではない。もっとも文学作品としての質が今後も確保できるかどうかは判断が難しいところだが。

□ 文学とメディアの関係

榕樹下がマスメディアに吸収された一例とすれば橄欖樹（<http://www.wenxue.com/>）のあり方は対照的と

いえるだろう。彼らのサイトは「宗旨」で明確に営利性の排除、多元性の尊重をうたい、独立を標榜している。

この橄欖樹には南京評論（<http://www.njpinglun.com/>）の葛紅兵の名前も見えるように、それぞれのサイトが独立を保ちながらも、連携して作品を生み出しているようだ。昨年の段階から存在したサイトもあれば、この一年で誕生したらしきサイトもあり、その数は増加傾向にあるように思われる。ただし、特定メンバーの作品だけを掲載するサイトが多い。従って、一定のポリシーに沿った雑誌作りが行われていると見て良いだろう。

例えば、「詩生活」（<http://www.poemlife.com/>）は中国だけでなく海外のメンバーも編集に加わりながら運営され、「詩生活月刊」の編集発行に加えて、「詩人專欄」「評論家專欄」も設けられ、メンバー限定の作品アーカイブとなっている。メンバーが限られた Web 雑誌の場合、営利性を追求する必要もないけれど、読み手不在（批評の欠落）によって閉鎖的になる危険性もある。そのためにも Web 雑誌相互の交流は今後一層重要になってくるだろう。

今回、関連サイトをブラウズして発見したのは中国での Weblog の流行である。書き込みがしやすく、複数のメンバーでの管理が容易なサイトとして利用されているようだ。投稿サイトに飽きたらぬユーザの「俺サイト」願望は日本以上に強いかもしれない。

fig. 2: 「詩生活」のトップページ



仏教

師 茂樹

□ 定番サイトの動向

● INBUDS: URL 変更

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

一ヶ月ほど利用不可能になった後、inbuds.org から変更となった。データベースセンターの手続きミスとプロバイダの不手際とのこと。このようなミスは今後起こらないようにしていただきたいところであるが、一方、利用不可能になったことによって INBUDS がいかに大切なツールであるかが再認識されたのではなかろうか。今後、学界全体でのサポートが望まれるだろう。

● SAT: 日本撰述部の公開開始

<http://www.l.u-tokyo.ac.jp/~sat/>

大正新脩大藏経 56～84 巻に相当する日本撰述部の公開が始まった。日本の仏典に関してはこれまで非常に手薄だったので価値が大きい。

● CBETA: 続蔵経年内公開

<http://www.cbeta.org/>

杜正民氏は H-Buddhism への投稿において、新纂大日本統蔵経の電子化の公開スケジュールを発表した^[1]。新纂日本統蔵経は所謂旧統蔵経を昭和の終わりの十余年をかけて再編したもので、国書刊行会によって出版された。CBETA による電子化事業も国書刊行会の許可の下に進められているとのことである。公開時期は、(1) 禅籍のうち、78～87 巻を 2003 年末までに、63～73 巻を 2004 年 7 月までに公開。(2) 54～88 巻が完成するのは 2005 年初めの予定。(3) 1～53 巻がその後 2 年以内に完成させたい、とのこと。

新纂日本統蔵経自体は、国内外でおそらくほとんど使われていないだろうと思う。その大元は日露戦争の戦勝を記念して編纂された大日本統蔵経であり、和装本であるため国内では編・套・冊・頁・段（二段組）によってページ指定をする場合が多い。しかし、現在は入手困難であることから、台湾の新文豊から出版されている洋装の影印本が海外を中心に国内でも広く用いられている。上述の通り、新纂日本統蔵経はこれに若干の増補・校訂を加えたものであるが、二段組を三段組に変えテキストの並び順も大幅に変更してしまっ

たため、旧版との連絡が難しく普及しなかったのである。

CBETA によるデータベースには台湾版の頁表記が追加される模様であるが、これを機にこれまで不当に無視されてきた新纂日本統蔵経に注目が集まる可能性もある。それと同時に、日本の文化遺産がまたもや海外において電子化されることになったことに、忸怩たる思いを持ってしまふ。

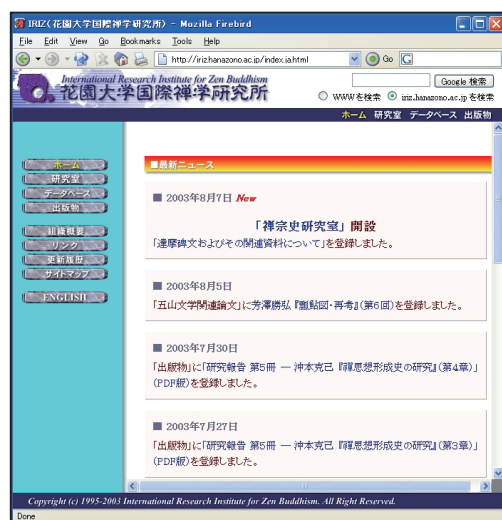
● IRIZ リニューアル

<http://iriz.hanazono.ac.jp/>

ZenBase CD 1 など、日本において仏教学情報化の先駆的業績を次々に発表してきた花園大学国際禅学研究所 (IRIZ) が、新しい URL に移行したことをきっかけにリニューアルした。文部科学省学術フロンティア推進事業を採択された花園大学国際禅学研究機構の下、これまで以上の活発な情報発信を行っている。

例を挙げれば、書籍並みのボリュームを持つ『研究報告』を（著者の同意が得られたもののみ）PDF で公開したのははじめ、仏教学だけでなく日本文化全般への影響が大きい『五山文学全集』（思文閣出版）のテキスト・データの公開（XML 版も近日公開予定）、一時話題になった達摩碑文の拓本画像とその翻刻の公開、所員・研究員による論文の電子化・公開等々である。

ちなみに評者もこのプロジェクトの末席にあって、



サーバをひとつ管理している。ここでは、漢字文献データベースを構築するのに必要な基礎データベースを公開するほか、テキスト・データベースのオープンソースの開発が可能かを実験すべく subversion というバージョン管理システム（共同開発支援システム）を用意し、研究者の方々に開放する方向で準備中である。

● H-Buddhism: URL 変更

<http://www.h-net.org/~buddhism/>

世界中から仏教学者が多数参加しているメーリングリスト H-Buddhism の URL が、その母体である H-Net の URL 変更に伴い変更された。メーリングリストへの投稿アドレスは従来のままである (h-buddism@h-net.msu.edu)。

□ 画像の威力

● 龍谷大学電子図書館「貴重書画像データベース」

<http://opac.lib.ryukoku.ac.jp/kicho/exhibi/top.html>

同図書館が所蔵する貴重書を全ページ、フルカラー画像 (JPEG) で公開している。各写本の状況によって読みやすさが変わってくるが、拡大すればほぼ問題なく閲覧が可能である。「真宗」「仏教」のほか、「文学」「自然科学」「芸能」「哲学・宗教」「歴史」などのカテゴリが立てられているが、同図書館の他の目録データベースなどと有機的に組み合わせればもっと便利になるのではないと思う。

● 駒澤大学図書館「禪籍善本目録」

<http://www.komazawa-u.ac.jp/~toshokan/zenseki/main.html>

同図書館が所蔵する膨大な禅籍の中から厳選された善本 202 点の一部を画像で公開している。書嗣情報のほか、一部のテキストに関しては解説もついている。

● 近代デジタルライブラリー

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

国立国会図書館が蔵する明治期に刊行された図書を画像として提供しているこのデータベースには、近代仏教 (学) 史を振り返る上で極めて貴重なものである。「仏教」というキーワードで検索すれば、村上专精、境野黄洋著、姉崎正治、常磐大定、井上円了らによる著作がたちどころに閲覧可能となる。詳しくは本号の小島浩之氏によるレビューを参照されたい。

● 小林写真工業株式会社製 CD-ROM

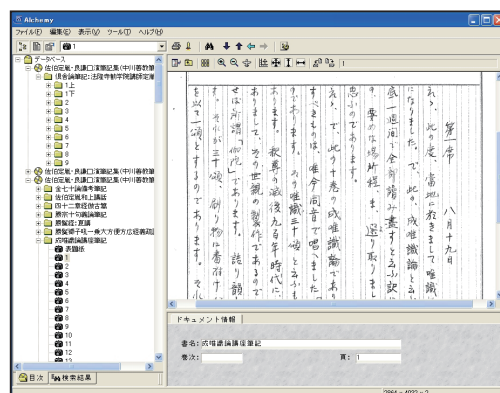
小林写真工業 (<http://www.ksk-jp.com/>) では高野山大学関連の貴重書をデジタル化し、目録検索機能付ビューア Alchemy Search^[2] をつけて、数万〜数十万円で販売している。仏教学関係では以下の通り：

- ・『高野大師行状図絵』十巻本 CD-ROM 版
- ・高野山大学附属高野山図書館所蔵『デルゲ版西藏大蔵経』(CD-ROM 版)^[3]
- ・高野山大学附属高野山図書館所蔵『高野山講式集』CD-ROM 版
- ・高野山親王院所蔵『法隆寺勸学院講師 佐伯定胤・良謙口演筆記集 (中川善教筆記)』CD-ROM 版
- ・勸修寺開山老千百年記念奉賛出版『勸修寺善本影印集成』善本影印版

このほか、『紀伊民報』、『南都拍姓楽家 東友弘家文書』、『図書館界』、『朝日新聞』、談山神社所蔵『談山神社古文書集成』、高野山大学附属高野山図書館監修『明教新誌』など。

Alchemy Search は Windows 専用で、Windows 3.1、95、98、NT 4.0 に対応 (バージョンによって異なるようである)、2000、XP についても動作確認が取れているとのことである。

評者が購入した高野山親王院所蔵『法隆寺勸学院講師 佐伯定胤・良謙口演筆記集 (中川善教筆記)』には Alchemy Search Release 6 が同梱されており、Windows XP Professional に問題なくインストールができた。基本的に CD 上のデータを閲覧することが前提となっているようで、データそのものをハードディスクにインストールしてくれるわけではないようである。



る。しかし、CD-ROM から手動でコピーしても問題なく動作するようだ。

また、Alchemy Search は Windows 専用のソフトウェアであり、かつデータ形式も独自なもののなので、そのままでは他の OS では使えない。Web など、特定の環境に依存しない公開方法が主流になってきている現在、この仕様が普及の妨げになっているかも知れず、内容が貴重なだけに残念である。しかし、画像データや目録データなどをすべて書き出すことができるので、他のプラットフォームで同等の閲覧環境を構築することも難しくはないであろう^[4]。

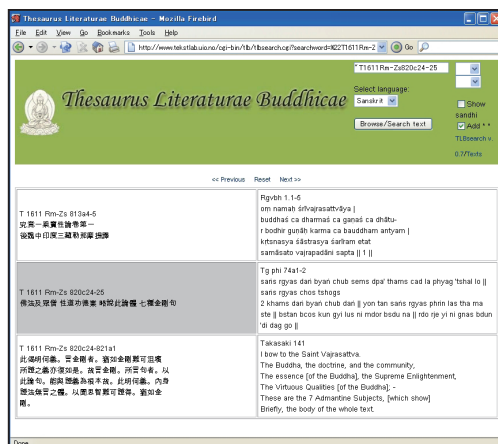
□ 新しい会本

● Thesaurus Literaturae Buddhicae (TLB)

<http://folk.uio.no/braarvig/tlb/>

オスロ大学のプロジェクトが、初期段階の成果物として、『究竟一乘宝性論』『廻諍論』等を梵蔵漢英四言語対照で検索・閲覧可能なデータベースを公開した。

テクニカル・ノートにもあるように UTF-8 ベースで作られており、ダイアクリティカル・マークなどの入力にも工夫が見られる。今後、このような形の、複数言語対照、異訳対照^[5]、本文と注釈との対照などをコンピュータ技術を使って視覚化する新しい会本が、ますます盛んになってくるのではなかろうか。



注

- [1] <http://h-net.msu.edu/cgi-bin/logbrowse.pl?trx=vx&list=H-Buddhism&month=0308&week=b&msg=jsvvFc6HhTH%2bo9Y6PuKpkQ&user=&pw=>
- [2] Information Management Research 社 (<http://www.imrgold.com/>)
- [3] チベット大蔵経は Tibetan Buddhist Resource Center (<http://www.tbrc.org/>) から販売されている。
- [4] ただし、そのような活用がライセンス的に問題ないか、メーカーに確認しなければならないだろう。
- [5] 釈惠敏氏は『瑜伽師地論』各漢訳を対照するデータベースを開発されている。“A Study on Creation and Application of Electronic Chinese Buddhist Text With the Yogācārabhūmi as a Case Study” (『印度学仏教学研究』100) 等を参照。

❖学術ソフト・製品

国学のCD-ROM

千田 大介

□ 国学とは

国学網 (<http://www.guoxue.com/>) は、膨大なテキストデータアーカイブを擁し、かつ学界動向に関する最新情報が素早く掲載される中国人文科学学術ポータルサイトとして知られている。もとは、わずか三人の愛好家が始めたサイトであったが、内容が豊富であり、かつ人民日報などの新聞にも取り上げられたこともあって、広く注目を集めるに至った。

その国学網を運営する北京国学時代文化伝播有限公司では、ここ一二年、次々に文献データ CD-ROM を発売している。以下では、それらの CD-ROM の内容・使い勝手などを、筆者が二階堂善弘氏とともに同社を訪問した際に、同公司総裁の尹小林氏からうかがった内容をふまえつつ、レビューしたい。

□ 国学の CD-ROM

国学の製品は、国学網上で盛んに宣伝されている割

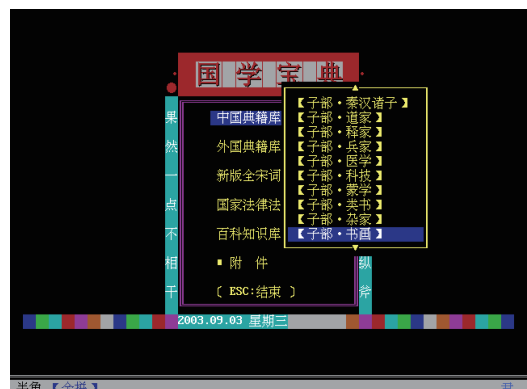
に、中国のソフト販売店や書店では全く目にしない。それもそのはず、国学の CD-ROM は在庫管理等のリスクを避けるため、一般の流通ルートにはまったく乗せておらず、もっぱら通信販売でのみ販売しているのである。しかも、発注があったらその時点で CD-R に焼きラベルを印刷してパッケージするという、ある意味家内制手工業的なオンデマンド生産体制をとっている。

このため、日本国内からの購入は中国書籍店を通るのが無難であるし、現地で購入する場合も、短期間の滞在であれば首都師範大学北校区総合楼三階の同社に直接買いに行くしかない。

以下、同社の主な文献データ CD-ROM 製品を見ていく。

● 国学宝典

十三経、正史から明清小説や西洋古典の翻訳まで、約五億字の規模を誇る、文献データ CD-ROM。開発は 1990 年代後半、そのため動作環境も MS-DOS + 中文環境エミュレータ UC DOS と、きわめてクラシカルだ。UCDOS は既に発売中止になっているが、国学宝典 CD-ROM にライセンス版が収録されている。Windows 環境では、所謂 DOS 窓で動作させることに



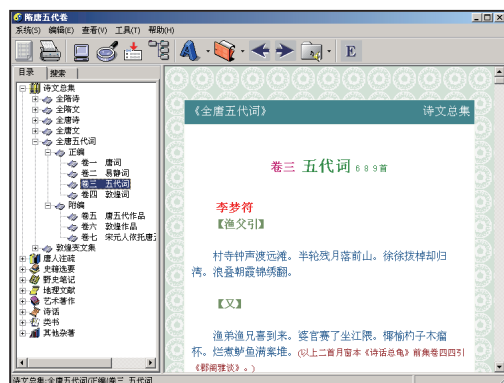


なるが、日本語版 Windows XP でも US コマンドで DOS を英語モードに切り替えることで動作した。

文献は、「中国典籍庫」「外国典籍庫」「新版全宋詩」などに分類され、「中国典籍庫」は更に「経部・十三經」「史部・正史」など 30 ほどのサブカテゴリに分かれる。

本文の閲覧の他、検索機能もある。四つまでキーワードを指定して、結果をファイルに保存するというもので、カードデータベースを自動生成するような感覚である。便利ではあるが、しかし、検索の最大単位は各分類カテゴリで、「中国典籍庫」全文検索ができないのは不便である。

五億字を手がけるに検索できるのはよいが、しかし、国学宝典の購入は、正直お勧めできない。その理由の第一は、著作権の問題である。たとえば、清代小説『説唐』の底本には解放後に大幅に書き換えられた修訂本が使われているし、正史はいずれも中華書局本を使っている。中国で著作権法の啓蒙活動が本格化する前に作られたものであるだけに仕方のない面もあるが、問題は問題である。



もう一つは、やはり動作環境の問題である。MS-DOS の基本操作法やコマンドを覚えていないと起動させることもできないし、さらにはマニュアル無しで UC DOS を使いこなさなければならない。パソコンは Windows からはじめた、というような人は、手を出さない方が無難だ。

国学側も国学宝典には問題の多いことを認めており、そのため現在では、たつての希望が無い限り、販売していないとのことである。定価は 4,500 元。国学が次々に文献 CD-ROM を出すのは、国学宝典データの再整理作業としての側面もあるようだ。

●基本典籍庫 隋唐五代卷

昨年発売された CD-ROM。本誌前号のレビューでは、これと基本古籍庫とを混同して紹介してしまったが、両者は全くの別物である。謹んで訂正したい。

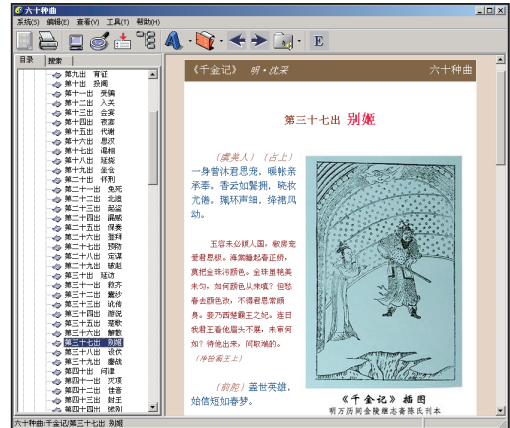
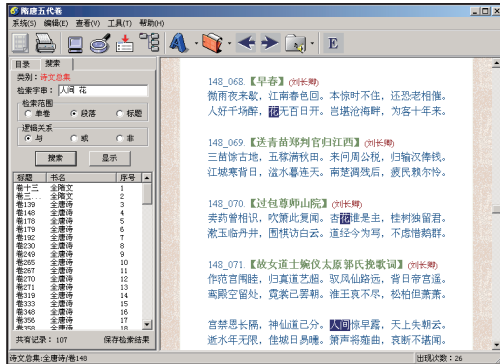
さて、隋唐五代巻は基本典籍庫の 2 であるが、1 の先秦漢魏南北朝巻はまだ出版されていないので、同シリーズの第一弾ということになる。定価は 4,150 元、輸入販売価格は 80,000 円前後である。凶器にも使えるのではないかなと思えるほど重い百科事典サイズの本製ケースに、マニュアルと CD-ROM 一枚が入っており、その中に『全隋詩』『全唐詩』『全隋文』『全唐文』などの詩文断代総集から『新旧唐書』『資治通鑑』『唐会要』などの史書、筆記小説、さらには『太平御覽』『太平広記』などの類書に至るまで、合計 136 種の文献が収録されている。

このように主立ったものを並べてみると、意外に宋代に編纂された書物が多いのに気が付く。その一方で『晋書』のような唐代に編纂された史書が落ちている。おそらく隋唐五代に関連する文献を収録するというコンセプトなのであろうが、しかし、唐代の語彙や文体を研究する参考としては使いにくいし、そもそも説明書に見える「851 年から 960 年までの間に現存した書籍」という定義と齟齬している。この点、収録書籍の選定はもう少し慎重に行って欲しかった。

さて、基本典籍庫の動作には、中文 Windows 環境が必須である。中文版 Windows はもとより、日本語版 Windows 2000/XP の言語を中国語に切り替えても動作する。

画面は左右に分かれ、左フレームにはタブ切り替えで目録と検索画面が、右フレームには本文が表示される。

検索では、まず検索する文献を目録タブで選択する。



個々の書籍の他、カテゴリを選ぶこともできるが、収録文献全てを対象とした検索ができない点は改善が望まれる。キーワードは、半角スペース区切りで複数指定でき、条件は and・or・nor から選択できる。語句の出現範囲は、巻・段落・標題に限定することができるが、最低単位が段落というのは時として広すぎるので、文単位、詩歌であれば作品単位の検索に対応してもらいたい。検索を実行すると、ヒット箇所の一覧が左フレームに表示されるので、それをクリックして右フレームに本文を表示させる。

本文は、簡体字横組み新式標点である。コピーには200字までの制限がつく。読者諸賢には繁体字版データを熱望している向きも多いと思われるが、後で触れるように簡体字横組みこそが国学のコンセプトがあるので、その点はあきらめるしかない。

基本典籍庫の収録文献には、台湾中研院や『四部叢刊』などと重なるものも多いが、『全唐文』などを他に先駆けて電子化した文献も多い。その意味で、中国古典文献を扱うものにとって、非常に価値が高い CD-ROM であるといえよう。

●六十種曲

『六十種曲』は明末に毛晋汲古閣から刊行された南戲傳奇(戯曲)の選集である。CD-ROM には『盛明雜劇』も収録されており、両者をあわせれば明代の代表的戯曲作品がほぼ網羅される。

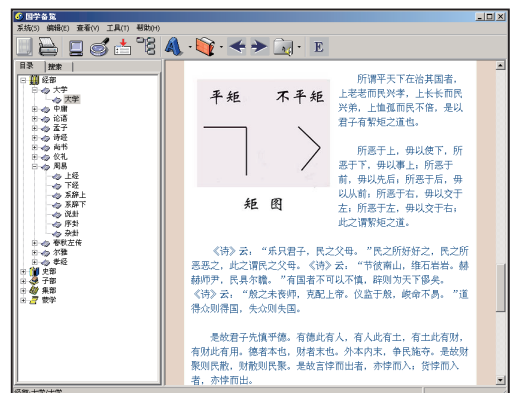
動作環境・インターフェース等は基本典籍庫とほぼ同等、本文はやはり簡体字横組みである。ただ、セットアップの必要がなく、CD-ROM をセットするだけで自動でソフトが起動して利用できるのは便利である。

難を言えば、戯曲という文体を考えると、曲牌検索・歌詞とセリフとを区別した検索などの機能も盛り込んで欲しかった。

●国学備覧

8cm CD-ROM 一枚に、儒教経典、『国語』『戦国策』『史記』などの史書、諸子百家の著作、『三国志』『水滸伝』『紅樓夢』などの明清小説、『楚辞』『唐詩三百種』などの古典詩文アンソロジー、童蒙書など 81 種の文献を収めたもの。価格はわずか 28 元、国内でも千数百円で入手できる。

これも、動作環境・インターフェースなどは前の二製品とほぼ同じで、CD-ROM をセットすると自動起動する。本文には他の CD-ROM よりも一層多くの解説図・挿絵が挿入されている。内容を理解する上で便利なのは確かであるが、しかし、それらの図の来源が明記されていないので、信頼性がどれほどのものか疑問が残る。もっとも、国学網の宣伝によれば、これは中国の高校生・大学生の教養古典ニーズを当て込んだ製品であるようだが、学術的な厳密性を求めるのは、お門違いであるのかもしれない。



□ 国学の理想と限界

■ 陸続と出版される CD-ROM

国学では、今後も更に多くの文献 CD-ROM の出版を計画している。基本典籍庫の先秦漢魏南北朝巻は、資金問題が解決次第、年末もしくは来年はじめに出版の予定だという。ほかにも、『十三經注疏』『全上古三代秦漢三國六朝文』『全秦漢魏晉南北朝詩』、さらには歴代叢書の目録を網羅的に収録した CD-ROM の出版も計画されているという。

この入力スピードは、ひとえに尹氏の個人的な能力に負うところが多いようだ。尹氏は一日で数十万字の書籍を入力することができると豪語しており、著作権問題から一般に販売できない書籍も多数テキストデータ化して所有しているとのことであった。

ともあれ、今後も国学の動向からは、目が離せそうもない。

■ 希薄なユーザーサポート意識

国学が少人数の愛好家が作った Web サイトが成長したものであることは前に述べた。しかし、この素人による運営から出発していることは、とりもなおさず学術ソフトウェアベンダとして国学を見た際の不満に結びつく。

たとえば、国学ではデータに誤りが見つくと直ちに訂正し、次回オンデマンド生産時にそれが反映されるという。しかし、いかなる錯誤があったのか、それまでに購入したユーザーに伝達する方法については全く考慮されていない。理想的には修正パッチの配布が望ましいし、それが困難な場合でも Web 等を通じた告知は必要であろう。流通形態の問題も同様だ。一般のソフト流通ルートに乗せていない旨、Web 上に明記すべきであろう。

企業を運営し製品を販売する以上は、無償で提供される Web サイトとは異なるのであるから、しっかりとしたユーザーサポート意識を持ってもらわないと困る。今後の改善に期待したい。

■ 国学データは使えるか？

国学の製品を使うとき、常について回るのがデータの信頼性の問題である。国学のデータには、版本が全

く明示されていないし、校注も付されていない。このため、傍目には著作権上の問題を抱えているかのように見える。また、中国古典を扱う研究者には繁体字本の需要が強いが、国学のデータは全て簡体字である。

この点について、尹氏の答えは明確であった。「国学のデータは、諸本を校勘して作った整理本である。また、研究で必要とされているのは簡体字の整理本だけであり、いちいち版本を見る必要性はない。」

たとえば、一部のデータでは、簡体字の大規模叢書『伝世蔵書』を出版社の許可を得て底本にしているが、また、他のテキストにもよって校勘している。このため、完成したデータは、国学オリジナルのものになるのだという。そうであるならば、著作権上の問題はあまり大きくはなさそうである。しかし、尹氏は同時に「古典の句読は誰がふっても同じだから、著作権は発生し得ない」とも語っており、全般に著作権問題を都合よく解釈する傾向が強いように思える。同社の CD-ROM の説明書には、問題発生時には国学が責任をとる旨書いてあるので、ユーザーとしてはひとまず安心ではあるのだが。

1990 年代後半以降、中国では経済発展に比例して大学進学者数が急激に増えており 2002 年度の進学率は 14% にも達している。エリート教育と大衆教育の分水嶺とされる 15% を超えるのは時間の問題であるし、北京・上海などの都市部の進学率が更に高いことを考えれば、中国の大学は既に大衆教育機関に変貌しているともいえる。中国の出版界で、日本で言えば新書に相当する知識読み物の出版がブームになっているのも、これと並行した動きとして位置づけられる。

国学とは、この知の大衆化とも言うべき動きを同時期に進展した IT 化の流れの中で体现したものであると言えよう。閲覧画面中に、人物や器物などの解説図を多く挿入して理解の助けとする一方、底本・諸版本の校勘・整理やテキストクリティックといった、研究指導を受けた経験のある研究者であれば当然身につけているべき手続きに思い至らないのは、そのような文脈のなかで納得される。

我々もこの点を割り切って、問題の存在を念頭に置いた上で、国学の製品を出典・用例探しの参考用途に限って利用するしかあるまい。なにしろ、『四庫全書』や『四部叢刊』にも匹敵する膨大なデータが比較的安価に入手できることは、他には代え難い魅力であるのだから。

中国の文献データ製品

千田 大介

□ 中国基本古籍庫

■ 発売はされたが…

『四庫全書』の三倍、約 20 億字の巨大データベース、中国基本典籍庫がいよいよ発売になった。一時消滅していた公式サイト、中国典籍網 (<http://www.cn-classics.com/>) も復活している。前号では、この典籍網という名称から、基本古籍庫と国学の基本典籍庫とを混同してレビューしてしまったが、両者は別物であった。謹んで訂正する。

中国基本古籍庫は全体を十輯に分けて順次発売になるが、今のところ第 2 輯までが発売されている。典籍網の収録書籍目録によれば、各輯それぞれ 1,000 部の書籍を収録する。いずれも、様々な分野の書籍が混在しているので、分類に関わりなく完成した部分から順次発売しているようだ。

価格は以下の通り。

- 普通型スタンドアロン版(テキストデータのみ)
 - 中国国内: RMB400,000
 - 中国国外: US\$50,000
- 研究型スタンドアロン版(テキストデータと版本画像)
 - 中国国内: RMB670,000
 - 中国国外: US\$85,000

15% 前後の値引き販売もあるようだが、最低価格のものを選んででも日本円で 500 万円を超える金額になる。研究上の必要性から言えば、版本画像が収録される研究型が望ましいのは言うまでもないが、昨今、大学の図書予算削減が厳しいとの話を至るところで耳にするから、普通型の購入すらままならない大学が多いのではなかろうか。

■ 試用版の評価

発売間もないので、筆者は執筆時点で正規

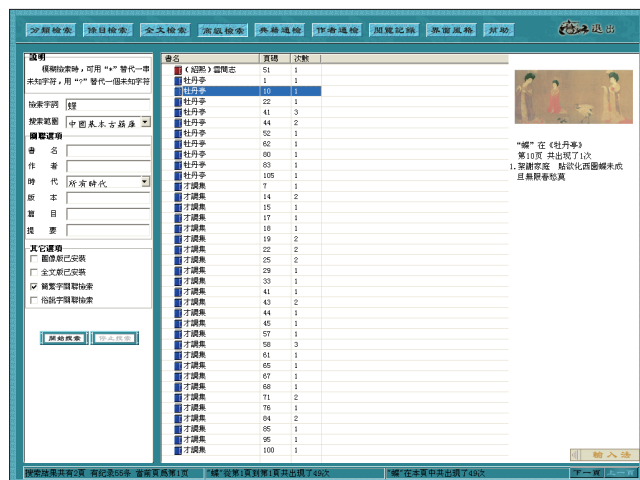
版を試用する機会に未だ恵まれていない。そこで、基本古籍庫試用版に基づいて機能を検証する。製品版とは機能が異なる可能性がある点、ご承知おきたい。

基本古籍庫は世界各言語版の Windows 2000/XP で動作するとされている。しかし、試用版を日本語版にセットアップすると、基本古籍庫そのものは動作するものの、Windows の言語設定が書き換えられてしまい、多くの日本語ソフトが使えなくなるという障害が発生する。もとの環境にもどすのは非常に困難で、最悪の場合、OS の再セットアップが必要になるので注意されたい。正規版を運用する際には、専用マシンを割り当てた方がよいだろう。

書同文公司 (<http://www.unihan.com.cn/>) のソリューションで作成されたものであるだけに、基本的な機能は『四庫全書』『四部叢刊』などの製品とほぼ同等である。ただ、インターフェースなどには大幅に手が入っている。特に、複数収録される版本画像の提示方法は工夫されており、複数の版本画像を並べて表示したり、テキストデータと画像を並べて表示したりすることもできる。ただし、本文と画像をリンクさせてスクロールすることはできない。また、本文の簡体字・繁体字切り替え、縦組み・横組み切り替えも可能である。

フォントは、「方正楷体」(FZKai-Z03) が付属する。バージョンは『四庫全書』全文検索版所収のものとおなじ 1.0 だが、タイムスタンプが異なりサイズも若干





小さくなっている、私用領域に多少の文字の出入りがあるものと思われる。

また、テキストデータをそのまま附属の RTF エディタにエクスポートして編集・印刷・保存することができる。ただし、エクスポートされるデータは、画面に表示されている部分のみである。

検索では、「*」（任意の文字列）「?」（任意の一文字）が利用できるが、『四庫全書』などの「A%nnnB」検索はできなかった。漢字関連検索機能も簡略化されており、「簡繁関連検索」「俗字関連検索」の選択肢しかない。しかも、「説」と「說」の検索結果は一致するが「説」と「說」とではヒット数・ヒット箇所ともに異なる、「觀」ではヒットするが「観」では一例もヒットしないなど、関連づけは不十分であるし、日本語環境での常用漢字体の利用についても注意が払われていない。製品版には『四庫全書』『四庫全書』並みのテーブルが装備されることを願う。

以上のように、基本古籍庫試用版には、動作環境や検索機能などに問題点があるものの、テキストデータ＋版本文画という特色を生かす設計になっており、文献データベースとしては十分な研究上の実用性を備えている。何より、そのテキストデータ・版本文画の収録量は、他の何にも変えられない魅力である。

❑ 超星数字図書館

■ 検索システムの更新

数十万冊の書籍を画像データで提供する超星数字図

書館 (<http://www.ssreader.com/>) は、自宅に居ながらにして高価な叢書や国内では閲覧が困難な書籍を簡単に読めるサービスとして、国内でも徐々に人気が高まっているようだ。

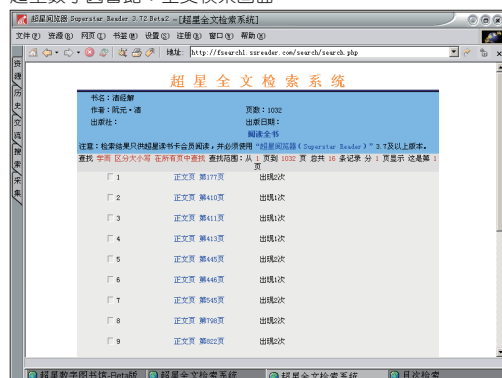
超星のここ一年の大きな変化としては、各種検索サービスの充実が挙げられる。とりわけ便利なのが、書籍の全文検索機能だ。書籍情報画面に「全文検索」ボタンが表示されている場合は、それをクリックして全文検索画面を開く。キーワードを指定して当該書籍の全文を検索し、検索結果一覧から当該ページの画像にジャンプできる。繁体字の文献であっても、検索用データベースは簡体字に変換して構成されているようで、キーワードは簡体字でしか指定できない。検索範囲も一冊限りである。

文献検索といえば、台湾中研院漢籍電子文献や、四庫全書・四庫全書など数多いが、著作権問題から現代の研究者の注釈などは提供されていない。しかし超星では、そのような注釈本や研究書なども検索できる。

ところが実際に検索してみると、あるはずの字句にヒットできないことがままある。全文検索機能は、書籍画像に自動で OCR をかけて生成したテキストと画像ファイルとを連携させているものと思われるので、ページや行を跨ぐ語句の処理が不完全であること、誤認識が多くデータの校正作業も不十分であることなどがこの原因として考えられる。今後、対応書籍を増やすとともに、質の向上にも取り組んでもらいたい。ユーザー側も、このような問題があることを認識した上で、用例探しなどに限定して活用するのがよからう。

また、書籍の検索機能でも、ここ数年間利用でき

超星数字図書館：全文検索画面





超星数字図書館：目次検索結果

なかった書籍の目次検索が復活した。超星の膨大な書籍を効率的に検索する上で非常に有用である。これもデータの校正が不十分で錯誤が多い点に注意が必要である。

■ 普来得

在米華人の陳洪星氏が運営する「美国出版在線」(Publishing Online, Inc)が提供する、主に在外華人をターゲットにしたオンライン教育サイト。語学・ITから中華料理に至るまで、数百のオンライン講座をそろえるとともに、超星数字図書館との提携によるオンライン図書館サービスをも提供している。

システムや提供される書籍データは、基本的に超星数字図書館と同一である。プリペイドカード「普来得学習卡」(定価 158 元だが現在は販促価格 108 元)を購入し登録することで、一年間、図書館と各種講座を利用することができる。

ただ注意が必要なのが URL で、カードに記され



ている URL (<http://www.poledu.com/>) が執筆時点ではアクセス不能になっており、ミラーサイト (<http://cn.poledu.com/>) から利用しなくてはならない。

利用にあたって普来得閲覧器のセットアップを求められるが、これは超星の SSReader に他ならず、既に超星を使っているのであれば、重ねてセットアップする必要はない。

超星では、あるカテゴリが突然アクセス不能になる、といった事態がしばしば発生し、サービスの安定性に問題があるが、そういった場合のバックアップとして普来得は役に立つ。例えば、2003 年前半に不安定であった北大図書館古籍、削除されている地方志などは、普来得では一貫して問題なく閲覧・ダウンロードできた。ただし、目次検索機能などは提供されていない。

超星数字図書館コンテンツを安定的に使いたい人は、バックアップ用に購入しておいても損はないだろう。

■ 著作権問題ははじけない？

前号の超星数字図書館レビューで、筆者は中国国家図書館の中国数字図書館が著作権法違反で訴えられ損害賠償を命じられた事件^[1]を紹介したが、その影響で中国数字図書館はサービス停止に追い込まれている。

前号でも触れたように、超星が書籍を著作権者に無断でデジタル化・公開しているのは事実であるが、いまのところサービスは提供され続けている。事情通によれば、超星の著作権処理には問題が多く細かな事件はしばしば発生しているようだが、中国国内の著作権法の運用がまださほど厳密ではないため致命的な事態にまでは至っていない、とのことである。事後承諾的ではあるが、まがりなりにも著作権使用料配布のしくみを作っていることがひとまず奏功しているようだ。しかし、これも今後の著作権法の運用次第ではどうなるかわからない。

恒久的にサービスを提供し続けられるように、超星数字図書館には、著作権問題とシステムそのもの、この二つの安定性の向上に一層努力してもらいたいものである。

□ 書同文とその周辺

■ 『四部叢刊』日本語版

書同文公司 (<http://www.unihan.com.cn/>) では、こ

こ一年、あらたな文献データ製品のリリースは無い。昨年来アナウンスされている『歴代石刻史料』は、今年中に発売予定とのことである。

同社関係の動きとしては、『四部叢刊』電子版の日本語版の発売がある。これは、日本の西岡漢字情報工学研究所の出資・委託によって、書同文が『四部叢刊』電子版のインターフェースを日本語化したものであり、検索・閲覧機能は中国語版と同等である。日本学研究所の必要から中国古典をも検索する必要のあるユーザーにとっては福音となろう。

日本語版は西岡漢字情報工学研究所から発売される。価格は中国語版より高くなる見込みで、パソコンにプレインストールしての販売も検討されているとのことである。ただ、日本の学术界に果たして日本語版制作のコストに見合うだけの需要が存在するのか、少々疑問である。

■ 書同文 Viewer、Writer

書同文が開発中のユーティリティソフト。Viewer は IE のアドオンソフトで、Web 上の書同文独自形式のファイルを縦書き版本のように表示して、閲覧・検索できる。Writer は、一種の XML エディタで、書同文数碼太師で作成した XML ファイルを Viewer 用の独自形式ファイルに加工することができる。

書同文にデータ入力を委託した場合、CJK+ に基づく異体字処理が可能であるが、しかし、そうしてできたファイルは漢字テーブルが無いと検索・加工が困難である。その欠点を補うツールとして期待される。

■ 進行中のプロジェクト

書同文は中国国内の研究機関・図書館と共同して、さまざまなデータベースを構築している。現在、構築が進んでいるものに、以下の二種がある。

● 北京地理歴史信息系统

北京大学と共同で、本年より構築が始まっている。北京歴代の地方志・地域史文献を網羅的に収録する計画であるという。

● 地方志データベース

中国国家図書館との共同で、同図書館が所蔵する地方志の全文データベース化を進めることが決定したという。プロジェクトには 5 年以上の期間が見込まれ

ており、完成は 2010 年頃になるとのこと。

奇しくも両者ともに地方志デジタル化プロジェクトであるが、これは、基本的な古典籍のデジタル化が急速に進展した結果、次なる対象として地方志が浮上してきた、ということなのであろう。

このほかにもいくつか、档案などを対象とした大規模古典データベース構築の計画が持ち上がっているとのことである。

□ 電子書籍

■ デフレの定着と入力の手遅れ

前号でも紹介した十元正規版シリーズ「芝麻開門」は、その後も着実に数を増やしている。オンライン目録 (<http://www.2688.com/>) によると、その総数は 1,500 点にも達している。

「芝麻開門」には文献データ CD-ROM も次々と収録されている。この一年間では、古典文献では新華世紀の各種 CD-ROM^[2] が収録された。現代文学では、青蘋果社の『家庭蔵書集錦』から抜粋した魯迅・老舍らの著作集などが目立つ。

しかし、「芝麻開門」の価格破壊ぶりが目立つばかりで、新たな文献データ CD-ROM はほとんど目にしない。古典の定番書籍が、様々なソフトハウスから組み合わせを変えて出ているに過ぎない。おそらく、中国における教養としての古典への要求は、既に発行されたものでほぼ満たされているのであろう。

また、いずれのソフトも価格競争には熱心だが、品質の向上にはさほど熱心でないように感じられる。量が質に転化してくればよいのだが、その日が訪れるのはまだまだ先のようだ。

注

[1] 同事件を裁いた裁判長自身による解説が、<http://www.ccw.com.cn/news2/zl/dongtao.asp> に公開されている。また、本号掲載の著作権講座報告もあわせて参照されたい。

[2] 本誌第 2 号 p.185 参照。

電子辞書 2 機種

田邊 鉄

□ 学生が使う電子辞書

3 月に CASIO から XD-R7300^[1]、5 月にセイコーインスツルメンツ (SII) から SR-T5030^[2]、と、相次いで市販中国語辞書 (テキスト部分のみ) を完全収録した電子辞書が発売された。

最近、高校生や大学生が電子辞書を購入するケースが目立つという。ある家電店によると、これまでビジネスマン中心に売っていたのが、昨年末ごろから突然学生の購入が増えたという。古語辞典や単語練習など、受験生を意識した機能のついた機種が売れ筋とか。ここに来て発売された、英語以外の外国語辞書を完全収録した機種も、大学生を主要なターゲットに定めているのかもしれない (その証拠に、CASIO 製品のショッピング・サイト e-casio では、中国語電子辞書は「第二外国語」というコーナーにある)。

ここでは中国語授業を履修する大学生向けのツールとして、この 2 機種を比較してみた。

□ 収録辞書

中日／日中辞典はどちらも小学館の第 1 版である。かつて唯一の中国語電子辞書だった SONY の電子ブックと変わらない。どうせなら 2 版を収録してほしかった。

両機種とも内容に違いはないが、CASIO は、親字

表 1 中国語以外の収録辞書

	CASIO XD-R7300	SII SR-T5030
国語	広辞苑・逆引広辞苑 分野別小辞典	広辞苑・逆引広辞苑
漢和	漢字源	漢字源
英和	ジーニアス 3 版	ジーニアス大辞典
和英	ジーニアス	ジーニアス
英英	×	OXFORD 現代英英
英類語	大修館英語類語	OXFORD Concise
その他	JTB ひとり歩きの中 国語自由自在	学研カタカナ語 DHC 英語・中国語 とっさのひとつこと

に付いているはずの GB コードと“電碼”が付いていない。大学 1・2 年生の中国語授業には、あまり関係ないだろうが、なぜ外したのか気になる。

中国語以外の収録辞書を表 1 に挙げる。特に英語辞書の違いが目を引く。同僚の英語教員に聞いてみたが、大学 1 年の後期レベルでは、ジーニアス英和辞典では語彙数が足りないという。特に、リーディングやビデオ教材のリスニングでは、「10 万語 (研究社) レベル」が必要で、その他「しっかりした英英辞典の併用」も必須だという。また、CALL 教室での作文練習では、類語辞典をよく使うとも聞く。大学生が「これ一冊」と持ち歩くには、英英と大英和を収録し、英語関係が充実している SII が相応しいと言えるかもしれない。

□ 中国語辞書の使い勝手

中国語辞書としての検索や閲覧の使い勝手は、大きく異なっている。

■ 中日辞典

両機種とも「ピンイン」「漢字の日本語読み」「部首画数」「総画数」による検索ができる。SII は部首の日本語読みによる検索にも対応している。

ピンイン、漢字日本語読みでは、逆引きもできる。CASIO は専用の入力欄を用意しているのに対し、SII は通常の入力欄に、? や * のワイルドカード^[3]を加えて指定する。“w?men”のような中抜き指定もできるので、より柔軟性が高い。

検索結果のリスト画面で SII はカーソル位置にある単語の本文を先頭 5 行だけ小さな窓に表示してくれる。基本の意味だけ知りたい、という時に、いちいち本文を開く必要がなく便利だ。細かいことだが、CASIO はリスト画面で文字が拡大できない。似たような漢字が並ぶリスト画面でこそ、拡大文字が活きると思うのだが。

本文の構成では、「熟語」の扱いが異なっている。CASIO は親字を引いた時、「その漢字で始まる熟語」をリストアップする機能はない。熟語キーを押すと「そ

の漢字で終わる熟語」がリストアップされる。SII は紙版と同じように、「その漢字で終わる熟語」は本文の後にまとめて表示される。熟語キーを押すと「その漢字で始まる熟語」がリストアップされる。小学館の中日辞典は親字の下位に熟語がリストされる構成なので、SIIの方がより紙版に忠実であると言える。逆にCASIOは辞典の重要な独創である語の配列について紙版を無視している。電子化するのだから、印刷物の構成にこだわる必要はないが、完全収録を謳っているのだから、インデックスも含めて紙版との同一性が保持されているべきだろう。

■ 日中辞典

「日中辞典は使わせない」という教員は多い。入門・初級段階の学習者が「今言いたい」「書きたい」表現に、日中辞典の訳語を大した吟味もなく適用しても、ほとんど通用しないからだ。日中辞典は上級者が使うもの——と私も考えていたのだが、日中辞典の訳語から、直接中日辞典をひいてみることができる電子辞書に出会って、これなら初級者にも使えるかもしれない、と思った。

日中辞典の検索では、SIIのみ中日辞典と同じワイルドカードが使える。ただし「??? にん」のように“?”を複数使うと、極端に検索スピードが落ちる（この例の場合で実測18秒）ので、実用的とは言えない。使えるのは、せいぜい“*”を使った逆引きまでだろう。

□ 全体的な操作性、外観など

両機種の外観は、幅、重さはほぼ同じで、筐体の違いはCASIOの奥行きが若干長いこと、SIIがわずかに厚いくらいである。キー数はSIIが52、CASIO

は50なのだが、CASIOは大きさや色がばらばらのボタンがごちゃごちゃした印象。SIIのキーボードには、シフトキーがない。限られたキーに多くの機能を割り当てるため、CASIOはシフトキーを多用する。声調を入力する時は、例えば1声ならシフト→Qのように入力する。SIIは独立した声調キーを設けているが、1回押すと1声、2回押すと2声が入力される。辞書の切り替えも、CASIOは主要な辞書（広辞苑、英和、和英、中日、日中）に専用キーを割り振り、他の辞書はシフト操作で呼び出すのに対し、SIIは中日／日中、英和／和英の切り替えは、それぞれ1つのキーのトリグル動作で行う。好みの問題ではあるが、メインの日中／中日の起動だけはCASIOのように専用キーを当ててほしいと思う。

□ 大学生にはSIIか

スイッチオンで、すぐにピンインを入力できる手軽さも含め、全体的に後発のSIIの方が利用者への配慮が行き届いている印象を受けた。英語辞書の充実もあり、大学生向けツールとしてはSIIに軍配が上がる。ただ、中国語辞書の内容は決して十分ではなく、今後小学館2版や他社の中国語辞典、さらには中国語ソーラスなどが電子辞書として登場することを期待している。

注

- [1] CASIO Exword <http://www.casio.co.jp/exword/>
- [2] SII 電子辞書 <http://www.sii.co.jp/cp/>
- [3] “?”は任意の1文字、“*”は任意の1文字以上を表す。たとえば、「??? にん」なら「苦勞人」「遊び人」などがヒットする。なお、“*”は0文字にはヒットしない。

字通

山田 崇仁

□ はじめに

CD-ROMあるいはWeb上で入手可能な漢字文献の電子テキスト（テキストコーパス）は、ここ数年でかなり充実してきた。それに対し漢字文献を読みこなす

ための辞典（字典）の電子化は遅れている。

特に日本人にとって「古典漢文」を読みこなすための必須の工具である「漢和辞典」にいたっては、学習研究社の発行する『漢字源^[1]』（JIS X 0208の範囲内のみ対応）『学研漢和大辞典^[2]』（JIS第一～第四水準まで対応）くらいしか見あたらず、しかもWeb上

では infoseek の漢字字典^[3]があるが、「対応範囲が JIS 第一・第二水準」「提示される情報が単漢字の極簡単な情報のみ」なので、「古典漢文」を読むための工具としては全く役に立たない。この状況は、国語辞典や英和辞典のソフトウェアが各社から何種類も発売され、かつ Web 上でも気軽に検索できるのに比べてお寒い限りである。その中で 2003 年 7 月 30 日に平凡社から白川静著『字通』が CD-ROM 化された事は歓迎すべきだろう。

『字通』は、漢字や詩経に関する研究で世に知られる文化功労者白川静氏の著作のうち、いわゆる「辞書三部作」の掉尾を飾る著作である。『字通』は「漢和辞典」だが、一般的な漢和辞典とは異なり「癖」がある辞典である。そのため他の漢和辞典と同じように利用すると違和感を抱くかもしれない。

しかしこの「癖」こそが、著者白川氏の学問観と深く関わる部分であり、それを理解して初めて『字通』のみならず白川学の世界に歩みを進ませる事が可能なのである。『字通』は漢和辞典である以前に、辞典の形をとった白川学そのものだ。CD-ROM 版のレビューをする前に、『字通』がいかなる漢和辞典なのかを見ることがしよう。

□ 『字通』の編纂方針

『字通』の編纂方針は、凡例の「編集の方針」に以下のように書かれている。

漢字の体系的な理解と、字義の展開のあとを考え、またその用語例を通じて、漢字文化の特質にふれることを目的として編纂した。漢字による用語の理解を深めることは、国語の表現をゆたかにする基礎になるものと考えられるからである。

これを読むと、他の漢和辞書がよく使われる意味においての字義あるいは熟語の解釈に中心を置くのに対し、『字通』が文字の原初的な意味を提示しそこから様々に派生する字義や熟語の展開を示す事に重点を置くことがわかる。この「根本を押さえる」という方針こそが白川学の基本であり、『字通』の編纂方針も実はその例に漏れないのである。

また「編集の方針」に見るとおり、『字通』は国語

という側面にも重点を置くことが理解されよう。

すなわち『字通』という書物は、漢字文献を読むための専門辞書であるというより、漢字をという文字を媒介にした漢和辞典と国語辞典という二つの性格を持ち、それを合わせる事によって日本人が伝統的に使用してきた文章語についてもう一度見つめ直すための「読む」辞書なのだとはいえる。そもそも『大漢和辞典』のように伝統的解釈を基準に編纂され、その解釈や用例を「調べる」ための辞書と、編集の根幹自体が異なるのだ。

従って極論してしまえば、経学的な解釈に従って古典を読みたい場合には余り役立たないかもしれない。そのよしあしはともかく、『字通』はそのような「癖」を持った辞書だという事を念頭に置いて利用する必要があるといえる。

□ CD-ROM 版『字通』

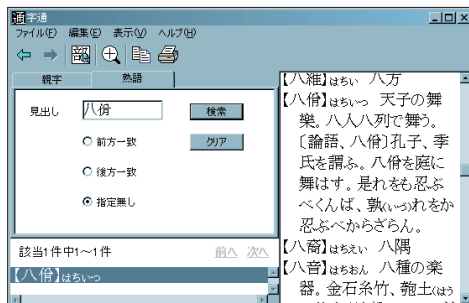
ここからは CD-ROM 版のレビューである。

『字通』は Windows 専用のソフトウェアだが、動作に必要とされるスペックはここ二三年以内に購入したパソコンであれば問題ないレベルである^[4]。ただし Windows Me に関しては、長時間使用すると動作が不安定になる可能性があるので、推奨はされていない。できれば Windows 2000/XP 環境で利用した方がよいだろう^[5]。

また、別途 Internet Explorer Ver.5.01 以上 + Microsoft .NET Framework 1.1 をインストールする必要がある。Microsoft .NET Framework 1.1 は『字通』本体と同時にインストールすることも可能だが、Internet Explorer は利用者が個別にインストールする必要がある。

ここでは、自作マシン上で動作する Windows XP

図 1:「熟語」で「八佾」を検索



にデータを全てハードディスク上にインストールした状態で動かしてみた。

●メイン検索画面

『字通』CD-ROM版を起動すると「メニューおよび検索ウィンドウ」画面が表示される。

画面は「入力」「検索結果一覧」「内容表示」に三分割される。以下、個別に説明していこう。

●入力

ここで各種キーワードを入力する。キーワード画面は「親字」「熟語」の二つあるので、それぞれをタブで切り替えて検索する必要がある。キーワード自体は「親字」で「漢字」「読み」「部首」「総画数」「四角號碼」^[6]が、熟語で「見出し(読み・漢字いずれかで)」がそれぞれ指定できる。漢字の部分にいくつか入力してみたところ、補助漢字程度ならUnicodeでそのまま入力できるようだ。ただし、以前に使用した検索語を再利用できないのは、検索インターフェースとしてやや物足りない。

文字の読みや部首・画数が不明な場合、文字の各パーツを組み合わせて検索するための「部品検索」機能がある。これはメニューの「表示」→「部品検索」から呼び出される。実用上の問題はないものの、先行する今昔文字鏡やATOK文字パレットとは異なり、ややぎこちない使い勝手だと感じた。

●検索結果一覧

キーワードを指定して検索を実行すると、この部分に検索結果の一覧が表示される。目的の親字や熟語が見つかったら、該当箇所を左クリックすると「内容表示」部分に実際の内容が表示される。HTMLの左右に区切られたフレームページのようなイメージと思っていたらよい。

●内容表示

内容表示画面に表示される内容は、書籍版『字通』

をそのままテキスト化したものである。

テキストは、JIS X 0208 + 独自フォント形式^[7]である。デジタル版『平凡社世界大百科事典』と同じ形式だが、七つもの独自フォントを用意してあるのは、甲骨文・金文・篆書に加え、より活字体に近い明朝体まで用意し、書籍版の見た目をなるべく忠実に再現しようとしためだろう。今昔文字鏡の甲骨文フォントが甲骨文の作成に最低限必要な十干や数字の部分が欠けていたのに比べ、『字通』CD-ROM版は代表字ではあるが収録されているので、これらの文字を個人的な目的で利用したい場合には重宝するだろう。

テキストデータのコピー＆ペーストも可能だ。但し、テキスト形式でエディタなどに貼り付けると、独自フォントの部分が■に置換されてしまう。リッチテキスト形式やHTML形式での貼り付けに対応したアプリケーション上へ貼り付けると文字化けしない。Wordや一太郎はもちろん貼り付けが可能。

●付録

書籍版『字通』には、同訓異字、書籍・人名の索引、平仄一覧などが付録として巻末に収録される。これらはCD-ROM版にも収録されているが、本体とは別途HTMLヘルプ形式のデータとして提供されており、併せて書籍版巻頭の白川静氏の「序」も収録されている。中身はメイン検索画面と同じJIS + 独自フォント形式で電子化された本文に必要なリンクをつけてHTMLヘルプ化したものである。また、検索機能はHTMLヘルプのそれを利用しているためそれほど使い勝手はよくない^[8]。

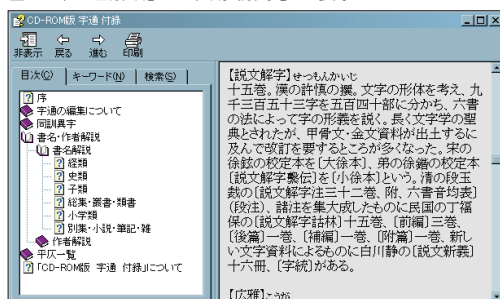
□ おわりに

以上、『字通』CD-ROM版のレビューを簡単に書いた。『字通』の辞書としての特徴は既に述べたが、CD-

図2:「王」の一部をコピーしてWordに貼り付け



図3:「書名解説」の「説文解字」を表示



ROM 版はよく言えば「できるだけ書籍版に収録される文字情報をデジタル化して収録した」といえるだろう。ただし検索のインターフェース自体は、HTML のフレームページに近いシンプルな物であり、デジタル版として書籍版の特徴をどう表現するかという視点では作られてはいない。

そのため CD-ROM 版『字通』は、一般的なデジタル版辞書と同じく「引く辞書」としての側面を重視したものとなってしまった。

このソフトウェアを実際に購入して利用するだろうと開発元が想定した層にとっては、シンプルな方が飽きが来ずかえって使いやすいのかもしれない。しかし、「読む辞書」の側面も持っている『字通』ならではデータの表現方法、特に「引く」よりも「読む」部分を重視して書かれている「付録」の扱いをもっと工夫すべきだったのではないか。

辞書とは「検索するためのインターフェース」と「辞書の中身」の二つがあたかも車の両輪のような関係であってこそ初めて機能する代物である^[9]。インターフェースが洗練されてこそ、辞書は真価を発揮すると評者は考えるのである。

またデータ自体で不満なのは、せっかく JIS X 0212（いわゆる「補助漢字」）程度まで検索可能なにもかかわらず、データ自体が相変わらずの JIS + 独自フォント形式な点である。

Windows 98SE 以上が動作環境として指定されており、甲骨金文など出土資料の文字が含まれている以上、Unicode ですら全ての文字を収録することが不可能なのは明らかだが、それならせめて補助漢字ベースにまではテキスト化を行い、独自フォント形式も異体字ならば元の文字のコードポイントをできるだけ利用する方法を採用して欲しかった^[10]。

しかし、漢和辞典のデジタル版自体が非常に少ない今日、著名な漢和辞典が CD-ROM 化されたことだけでも喜ばしい事ではないだろうか。

日本の誇るべき知的財産である『大漢和辞典』のデジタル化は資金難で難航しているという噂を聞く。日本漢学の伝統と神髄が蓄積されている漢和辞典のデジタル化の遅れは、日本における中国古典学の衰退とも無関係ではないだろう。これは、本場中国の漢字文献のデジタル化の急速な充実ぶりと好対照である。日本中国学の側から声を積極的に上げないと、この状況を

打破するのは難しい。

せめてハンディながら評価が高い『新字源』や『全訳漢辞海』といった漢和辞典のデジタル化を希望する次第である。

注

- [1] システムソフト電子辞典シリーズ。現在ではロゴヴィスタ株式会社取り扱い。
http://www.logovista.co.jp/jiten/products/Jiten_Kanjigen_W.html
- [2] スーパー日本語大辞典全 JIS 版所収
<http://www.gakken.co.jp/soft/jiten/>
- [3] http://jiten.www.infoseek.co.jp/Kanji?pg=jiten_kjtop.html&col=KJ&svx=100302&&svp=SEEK
- [4] 詳細なスペックは以下の Web サイトを参照。
<http://www.heibonsha.co.jp/jitsuu/guide.html>
- [5] Windows XP の場合、メモリーは 256MB 以上搭載することを推奨している。
- [6] 四角号碼の検索方法については、メニューの「付録」を選択して HTML ヘルプ形式の付録を開いて「四角号碼の検字法について」の項目を選択してその解説を参照すればよい。
- [7] ここでの「独自フォント」とは、JIS 未収文字などを JIS のコードポイント上に独自に文字を配列して、見た目上文字が表示されるようにしたものを目指す。例えば『字通』上の「八角」の実際の文字コードは「哭咽」であり、それを「JITSUU11」というフォントに変換する事で「見た目上」「八角」として表示しているのである。
- [8] 付録がメイン検索とは相容れない以上、ヘルプ形式になるのも仕方がなかったのだろう。どのような形式であれ書籍版全体をデジタル化しようとする行為には敬意を表することとしたい。
- [9] インターフェースとして書籍版は「五十音引き」という漢和辞典としては捉破りとも言える手法を採用しており、これが存外評者にとって使いやすかったりする。
- [10] この辺りは技術力やコストの問題が大きいものかもしれないが、日本のソフトウェア界の Unicode に対する一種のアレルギーみたいなものが原因にあると思うのは考えすぎなのだろうか？

青森県史資料編古代1補遺（青森県史叢書）

青森県発行 2003年2月

二階堂 善弘

紙媒体と電子媒体の組み合わせによる出版の試みについては、昨今ではそれほど珍しいものではなくなっている。しかし「県史」という、やや保守的とも思われる分野で、最先端の試みがなされていることには、さすがに驚きを禁じ得ない。

この『青森県史資料編古代1補遺』は、「青森県史叢書」の中の一冊であり、『青森県史資料編古代1文献史料』に収録されなかった文作作品その他の資料で構成されている。

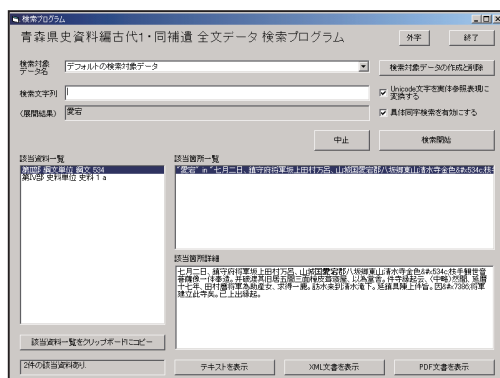
しかし、より重要なのは付録とされたCD-ROMに収録されるデータベースである。このデータベースには、『青森県史資料編古代1文献史料』及び『古代1補遺』の電子データをすべて収録し、表示や検索が可能となっている。しかも、単なる電子テキストではなく、XMLやPDFなど多彩なフォーマットにより、より高度なデータ利用ができるようになっている。そのため全文をXML、或いはPDF、或いはHTML、といった任意の形式により閲覧することが可能となっている。

更にそのための専用の検索ソフトやビューアまで備えられている。しかもUnicodeを使用しての多漢字

文書の対応もなされている。特筆すべきは、独特の異体字検索機能を有していることであろう。

ある意味で、このデータベースは現時点での古典データベースの「手本」となりうるものであるかもしれない。ただ、いまひとつ入手しづらいデータベースであるのが残念である。それから、マシンスペックによっては若干インストールなどに問題が出る場合もあるので注意が必要である。

専用の検索プログラムで語彙を検索



文字コード「超」研究 深沢千尋 著

ラトルズ 2003年8月 ISBN4-89977-051-0 2980円+税

二階堂 善弘

これまで文字コードに関する本には、どうしても理論面に偏りがちな傾向が強かったようだ。むしろ、文字コードを処理するためにはある程度の理論について理解していなければならない。しかし、実際にテキストを処理するための応用面に関しては、これまでの書籍では不十分であったように思える。

その点では、この『文字コード「超」研究』は、書名こそ「研究」と銘打っているが、徹底した実践面からの視点で書かれたもので、実際に文字データを処理

する現場で重宝されそうな書である。かといって決して理論面がおろそかになっているわけではない。そもそも「文字コードとは何か」、「コンピュータで文字を扱うとはどういうことか」など、非常に詳しく説かれている。ただ、それはあくまで「実践から見た理論」が重視されている、という感じであるのが特色である。扱っている文字コードも、ASCIIやJISコードから、Unicodeの拡張面まで、非常に幅広いものがある。

ただ、実に丁寧に説くためか、時に冗長に感ずる部

分もあるが、これは作者自身が、読み飛ばしを奨励しているように、読者のスキルに合わせてどんどん読み進めてよいと思う。

この書には、また「実際に Perl を駆使してデータを処理してみる」という特色がある。参考として扱っている Perl のスクリプトもかなりの数にのぼる。中にはわざわざ Perl を使わなくてもよさそうなものもあるが、逆にそこに作者のむしろ徹底した姿勢が感じ

られるかもしれない。

もっとも、ちょっと難を言えば、せっかく Unicode の拡張面について言及しているのに、多漢字処理については GT 明朝などを使っているのが気になる。実際にはすでに CJK の拡張漢字が使用できるようになっているのであるから、もっとこれらを用いた処理の実例を出してみてもよかったのではないだろうか。

電腦日本語論

篠原一 著

作品社 2003 年 3 月 ISBN4-87893-511-1 1800 円＋税

師 茂樹

本書は、ジャストシステムの有名な日本語入力システム ATOK の開発史・秘話を、当事者たちのインタビューを中心に振り返ったものである。著者が小説家であるためか、この手のインタビューものとしてはだらだらとしており読みづらい印象がある。しかし、開発現場の素朴な作業がやがて、規範性を求められ、日本文化を担うという意識を戸惑いながらも持とうとするにいたる過程がなかなか面白い。コンピュータが社会に浸透することで、日本のみならず各国の言語政策などに大きな影響を与え、また教育現場における役割も大きくなっている今日、日本一の看板を背負った ATOK がたどってきた道を振り返るのは、示唆に富む点が少なくない。

ただ本書では、ATOK の、特に辞書開発の面に力点が置かれているために、日本語入力システムを使った言語活動の多様性についての考察が少ないのが残念だ。最近 ATOK は方言や古語への対応を進めているが、その伏線は本書の前半に見える、国語学者たちの言語に対する洞察があったはずだ。国語学者たちは日本語のコンテクスト依存性を説き、また ATOK 開発を一貫して率いてきた小林龍生氏は ATOK を「楽器につかって良い」と言う。「氏ね」などの新たな表記は、日本語入力システムがあったからこそ生れたものだろう。

先日、前期末の課題として出したレポートの採点を

した。情報関連の授業であり、また手書きはかさばってかつ読みにくいので、ワープロ書きを推奨したところ、集まったレポートの 7 割方がそうであった（書いている途中でパソコンが使えなくなり、ワープロと手書きが半々という学生もいたが）。レポート用紙をめくっていくうちに、ひとつの“誤変換”に目を奪われた。「埒問題」。「対象」と「対照」、「意外」と「以外」などのよくある誤変換は、視覚的に似ているために起こるものだと思っていた。ワープロを使うと日本語能力が落ちるとよく言われるが、それは日本語入力の際、ずらっと並ぶ変換候補から「どこかで見たことがある」ものを動物的に選んでしまい、しかもそれがそこそこ当たってしまうからではないかと思っていた。しかし、当のレポートで執拗に繰り返される「埒問題」は、私の思い込みを粉碎するのに充分だった。「拉致」と「埒」は視覚的に似ても似つかない。しかも前者は、最近のテレビなどで見ない日はないのに、だ。

思うに我々の想像以上に、コンピュータは我々の言語の質を変えている。オングは現代の我々の言語活動を、視覚に依存した「文字の文化」であると喝破したが、いまやそれも徐々にずれつつあるようだ。ATOK はどこまでこの点に深入りしていくつもりなのか、もうひとつの電腦日本語論を読みたくなった。そしてそちらの方が、著者の筆も冴えるような気がしてならない。

漢字文献情報処理研究会彙報

2002.10～2003.9

2002 年 10 月 1 日

機関誌『漢字文献情報処理研究』第 3 号出版。

2002 年 12 月 7 日

第 5 回大会、2002 年度総会開催。

2003 年 5 月 20 日

会員名簿発行。

2003 年 8 月 2 日

夏期公開講座「東洋学情報化と著作権問題」、
臨時総会開催。

第 5 回大会・2002 年度総会

日時：2002 年 12 月 7 日（土）

会場：花園大学 無聖館 403 教室

■ 第 5 回大会

1. 研究発表（13:00～14:00）

- ・花園大学禅デジタルアーカイブ・プロジェクト
の構想 師 茂樹（花園大学専任講師）
- ・統計学的手法を利用した先秦文献の分類に関する
研究——『論語』を例にして—— 山田 崇
仁（立命館大学講師）

2. 特別セッション「CHISE が開く新しい漢字の世界」 （14:30～16:30）

- ・CHISE とは何ではないか 師 茂樹（花園大
学専任講師）
- ・XEmacs UTF-2000 の紹介 守岡 知彦（京都
大学人文科学研究所附属漢字情報研究センター
助手）
- ・漢字字形の自動生成とネットワークフォント
上地 宏一（慶應義塾大学大学院政策・メディ
ア研究科博士課程）
- ・漢字のネットワーク 藤原 義久（ATR 人間
情報科学研究所）
- ・漢字構造の可視化 (work in progress) 江渡
浩一郎（独立行政法人産業技術総合研究所研究
員／メディアアーティスト）

■ 総会（16:30～17:00）

- ・2002 年度事業報告、会計報告（会計監査：小
口雅史、山崎直樹）
- ・執行部改選
代表：二階堂善弘 副代表：千田大介
副代表兼会計：師茂樹
幹事（名簿担当）：山田崇仁
幹事（サーバ管理担当）：上地宏一
幹事：小島浩之、佐藤仁史、田邊鉄、平林宣
和
- ・新監事承認
小川利康、金子真也
- ・2003 年度事業計画、予算案承認
- ・『漢字文献情報処理研究』の発行形態について

2003 年夏期公開講座・臨時総会

■ 夏期公開講座

題目：東洋学情報化と著作権問題

日時：2003 年 8 月 2 日（土）

会場：慶應義塾大学 日吉キャンパス 来往舎中
会議室

講師：石岡克俊（慶應義塾大学産業研究所助教授）

1. 講演「著作権法のパースペクティブ」（13:30
～15:00）
2. 質疑応答（15:30～17:30）

■ 臨時総会（17:30～18:00）

- ・会則改定
- ・BBS 運用ルールについて
- ・JAET サーバ設置について
- ・名簿 PDF の会員への公開について

著者紹介

秋山 陽一郎（あきやま よういちろう）

1975年生まれ。立命館大学大学院文学研究科 博士後期課程に在籍中。1998年に立命館大学文学部中国文学専攻を卒業ののち、大学院は東洋史学専修に転向。専門は中国古代の政治思想史と文献学。現在は、一時一手に成らずとされる先秦古文獻の山から「人」を発掘する手段として、劉向校書事業を研究中。「孫徳謙 劉向校讐学纂微訳注〔一〕」（『立命館東洋史学』26、2003）

池田 証寿（いけだ しょうじゅ）

栃木県生まれ。北海道大学大学院文学研究科言語文学専攻助教授。国語学専攻。『一切経音義索引』（共著、汲古書院）。1991年より高山寺典籍文書総合調査団団員として高山寺経蔵調査にかかわる。また、1994年より2003年3月まで、日本規格協会情報技術標準化センター JCS 調査研究委員会委員をつとめ、JIS X 0208:1997の改正、JIS X 0213:2000の開発に関与。

石岡 克俊（いしおか かつとし）

1970年生。慶應義塾大学産業研究所助教授。専攻は、経済法、知的財産法、消費者法。著書に『市民カレッジ・知っておきたい市民社会の法』（共著・不磨書房）、『著作物流通と独占禁止法』（慶應義塾大学出版会）がある。著作物に関する再販売価格維持制度の研究を皮切りに、人間の知的・精神的活動の成果（著作物）と競争秩序との関係や、広く文化政策に関する市場のあり方、また政府の関わり方などについて関心を持っている。

石川 正敏（いしかわ まさとし）

1972年、兵庫県生まれ。2000年、奈良先端科学技術大学院大学情報科学研究科博士後期課程単位取得退学。同年、島根県立大学総合政策学部助手を経て2003年から同大学総合政策学部非常勤講師、北東アジア地域研究センター客員研究員。データベースシステムの研究に従事。

Wittern, Christian（ウィッテルン、クリスティアン）

1962年生まれ。中国の哲学と宗教、ドイツ文学、西洋哲学をチュービンゲン、ハンブルグ、杭州、京都、ゲッティンゲンで学んだ後、『景德伝灯録』巻28に関する研究でゲッティンゲン大学より中国学のPh.D.を取得。現在、京都大学人文科学研究所附属漢字情報研究センター助教授。ZenBase CD1の開発に携わり、CBETAの開発協力、Text Encoding Initiative (TEI) のtechnical councilのメンバーを務める。最近はテキスト・マークアップ理論及び知識表現に関心がある。

江渡 浩一郎（えと こういちろう）

1971年東京生まれ。慶應義塾大学大学院政策・メディア研究科修了。メディア・アーティスト。「WebHopper」(1996),「RemotePiano」(1996),「SoundCreatures」(1998),日本科学未来館「インターネット物理モデル」(2001)。2002年よりCHISE Project 可視化 subprojectに参加。現在、独立行政法人産業技術総合研究所特別研究員。情報流デザイングループにてネットワーク上のコミュニケーション・システムの研究を行う。

小川 利康（おがわ としやす）

1963年生。早大大学院博士課程中退、大東文化大学を経て、現在、早稲田大学商学部助教授。専攻は現代中国文学。現在、周作人・松枝茂夫往復書簡のデジタル化を進めている。来年には冊子＋CD-ROMで刊行予定。<http://www.f.waseda.jp/ogawat/>

上地 宏一（かみち こういち）

1976年大阪府生まれ、シンガポール育ち。慶大大学院政策・メディア研究科博士課程在学中。漢字グリフの自動生成に興味を持つ。

川幡 太一（かわばた たいち）

現在、NTTサイバーソリューション研究所に勤務。マークアップなどのテキスト情報処理システム、漢字・多言語の表示・編集処理や、ネットワーク分散オブジェクトシステムなどに関心を持つ。

小島 浩之（こじま ひろゆき）

1971年岐阜県高山市生まれ。東京大学経済学部資料室助手。東洋史学と図書館情報学の狭間で書誌学、目録学、資料保存などの分野を中心に彷徨している。主要論文は「唐の玄宗 - その歴史像の形成 -」（『古代文化』52-8, 2000）、「大学図書館利用者のためのオンライン目録学」（『漢字文献情報処理研究』2, 2001）、「現代中国書の書誌的特徴」（『大学図書館研究』64, 2002）など。

佐藤 誠（さとう まこと）

1973年愛知県生。花園大学歴史博物館研究員。花園大学非常勤講師。花園大学文学部史学科、花園大学大学院文学研究科日本史学専攻を経て、八幡市立ふるさと学習館勤務の後現職。日本民俗学専攻。論文「動物霊の伝承と構造～葬送儀礼をめぐる猫の伝承を中心として～」(『花園史学』第23号、H14)

佐藤 仁史（さとう よしふみ）

1971年生まれ。滋賀大学教育学部講師。専門は中国近代社会史。現在は郷土意識の形成という視点から清末民初の地方エリート層の地方志編纂や郷土教育に関する活動や、彼らの地域社会観や国家観などを調べている。近著に「清末民初江南の地方エリートの民俗観——「歌謡」をてがかりに」（『史学』第72巻2号、2003年）がある。

高橋 由利子（たかはし ゆりこ）

1948年愛媛県生まれ。慶応大学文学部中国文学専攻、お茶ノ水女子大学人文科学研究科修士課程修了。現在上智大学比較文化学部教授。大学では中国語と漢文を担当。研究領域は説文解字の文献学的研究。老舎とキリスト教の関係にも興味がある。数年前から大学内でのコンピュータ関係の委員会の担当となり、授業や研究にどのように生かしていくか模索中。

田中 正流（たなか まさる）

1974年大阪府生まれ。宝鏡寺門跡百々御所文庫学芸員、花園大学非常勤講師、中世日本研究所特別研究員。花園大学大学院文学研究科修士課程日本史学専攻修了後、各地の自治体史編纂、花園大学歴史博物館研究員を経て、現職に至る。専門は日本民俗学、最近は特に宝鏡寺門跡の年中行事や生活について研究している。中世日本研究所では、情報収集・公開全般を担当し、「女性と仏教」に関する文献目録データベースの構築をしている。

田邊 鉄（たなべ てつ）

1963年京都府生まれ。北海道大学情報基盤センター助教授。CALL、情報メディア論、ネットワーク・コミュニケーション論の「三結合」で法人化後のサバイバルを目指す。最近の著作は「中国語教育と情報処理」（『漢字文献情報処理研究』3, 2002）、「カジュアル・コンピューティング時代のインターネット・ガバナンス」（『北海道大学国際広報メディア研究科ジャーナル』1, 2003）など。

千田 大介（ちだ だいすけ）

1968年生。慶應義塾大学経済学部助教授。専門は中国文学、研究テーマは中国通俗歴史物語の変遷と受容だが、近頃は影絵人形劇などの芸能も扱っている。電腦そのものよりも、電腦を巡る社会と人々の動きの方により興味を感じている。著作など詳細は、Webサイト「電腦瓦崗寨」<http://wagang.econ.hc.keio.ac.jp/>を参照されたい。

苦米地 等流（とまべち とうる）

1965年青森県生まれ。京都大学大学院文学研究科博士後期課程（仏教学）単位取得退学。ローザンヌ大学文学部助手を経て現在、同共同研究員。専攻はインド・チベット仏教学。インド後期密教の文学的研究を主要テーマとする。著書に『大乘仏典 中国・日本篇 15 ツォンカパ』（共著、中央公論社 1996）など。GNU Emacs のチベット文字サポート実装を担当。

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単位取得退学。但し、その間は大学院と会社を行き来し、貿易事務やシステムエンジニアなども経験。東北大学大学院国際文化研究科助手を経て、現在、茨城大学人文学部助教授。専門は中国の民間信仰研究で、著書に『封神演義の世界』（大修館書店）や『中国の神さま』（平凡社新書）などがある。コンピュータ関連でちょっと仕事をしているせいか、外ではそちらが専門のようになっている。しかし実は自分の大学では電腦はほとんど教えていない。

松川 節（まつかわ たかし）

1960 年東京生まれ。東京外国語大学、大阪大学大学院（文学博士）、国立民族学博物館講師を経て現在、大谷大学専任講師。専門は東洋史（モンゴル時代史）と人文情報学。著書に『図説・モンゴル歴史紀行』（河出書房新社）など。

宮崎 泉（みやざき いずみ）

1968 年和歌山県生まれ。京都大学大学院文学研究科 COE 研究員。現在の研究テーマは後期インド仏教とそのチベットへの伝播。著作に「『禪定灯明論』漸門派章について」（『西藏学会々報』第 48 号）など。コンピュータ上の文字・言語処理にも興味があり、平成 14 年より CHISE プロジェクトに参加。

守岡 知彦（もりおか ともひこ）

1969 年生まれ。北陸先端科学技術大学院大学情報科学研究科修了。博士（情報科学）。京都大学人文科学研究所附属漢字情報研究センター助手。現実逃避のままに流れ流れて今に至っており専門は謎だが、現在は一応『多言語情報処理』ということにしている。この数年は特に文字符号に依存しない文字処理技術に関心を持って取り組んでいる。

師 茂樹（もろ しげき）

1972 年生まれ、大阪府生まれの福島県育ち。花園大学専任講師。日本を基点とした東アジアの仏教思想史、および計算機上の文字処理やテキスト処理・分析、そしてこれらすべてに関する方法論について思考をめぐらすことが目下の課題。近著に、方法論懇話会編『日本史の脱領域 多様性へのアプローチ』（共著、森話社、2003）

山崎 直樹（やまざき なおき）

1962 年生。大阪外国語大学外国語学部所属。専攻は、中国語学および中国語教育。現在の興味の対象は、電子媒体上の構造化テキスト (HTML/XML) の言語教育への応用。

山田 崇仁（やまだ たかひと）

1970 年、愛知県生。立命館大学文学研究科修了。博士（文学）。現在は日本学術振興会特別研究員（PD）・立命館大学非常勤講師。専門は中国先秦史のはず。現在は「統計学的手法を利用した先秦文献の分類に関する研究——儒家文献を中心として——」が正式な研究テーマ。N-gram や統計手法などを駆使して先秦文献の分類や成立について探るのが目的。

編集後記

『漢字文献情報処理研究』第4号をお届けする。
本号の刊行に先立ち、大修館書店月刊『しにか』7月号に「中国系お役立ちサイト70」という特集が掲載された。本会会員も記事を執筆しているが、国内サイトを中心とする紹介であったため、題材集めに苦労したと側聞している。インターネットが騒がれはじめてはや十年、未だに有効な教育・研究コンテンツを発信しえない我が国東洋学界の意識の低さが、図らずも露呈されてしまった。

近年、多くの大学ではデータベースの公開やe-learningの導入などを進めているが、一部の電腦盲信派の暴走の結果、あるいは研究・教育の先進性を宣伝するためのカンバンではないかと疑われるような事例も多い。このような見せかけの学術・教育情報化を排して、いかに新たな学術成果を生み出しうる本質的なコンテンツを構築できるか、個々の研究者・教育者の意識が問われようとしている。

本号の各特集は、東洋学研究・教育の情報化を進めるにあたって参考となる事例や最新の技術情報、さらにはクリアしなくてはならない法律問題などを取り上げている。これらが、本質的な人文学研究・教育情報化を目指す方々の一助となれば幸甚である。

本号の編集が終わろうというところに、勝村哲也先生の訃報が飛び込んできた。勝村先生は、e漢字やCHINA3の開発を通じて、漢字文献情報処理のみならず中国学全般の発展に大きく貢献してこられた、まさにこの分野のパイオニアであった。今後とも、その国際的な視野と行動力とでリードし続けてくださるものと期待されていただけに、急逝が非常に惜しまれる。謹んでご冥福をお祈りする。

さて、編集子は今号をもって本誌編集長を退く。振り返れば編集上の不手際も多々あったが、三号雑誌の壁を突破したことで、ひとまず初代編集長の最低限の職責は果たせたのではないかと思います。次号は新編集体制のもと、更にパワフルかつ「とんがった」内容に仕上がることだろう。読者諸賢にはご期待いただきたい。

本誌の刊行は、好文出版の尾方社長、日本中国語CAI研究会、さらに漢情研会員各位から有形無形の援助のたまものである。末筆ではあるが、御礼申し上げます。(●)

漢字文献情報処理研究 第4号

発行日	オリジナル版：2003年10月1日 PDF版：2005年10月10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 上地 宏一 佐藤 仁史 二階堂善弘 師 茂樹 山田 崇仁	小川 利康(CAI) 小島 浩之 田邊 鉄(CAI) 平林 宣和 山崎 直樹(CAI)	
デザイン DTP	電腦瓦崗寒 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540-106 TEL：03-5273-2739 FAX：03-5273-2740		

●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、

- ・送付先住所
- ・氏名
- ・年齢
- ・職業
- ・勤務先
- ・必要部数

を明記の上、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)。

●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。

ISBN4-87220-072-1 C3004 ¥2000E