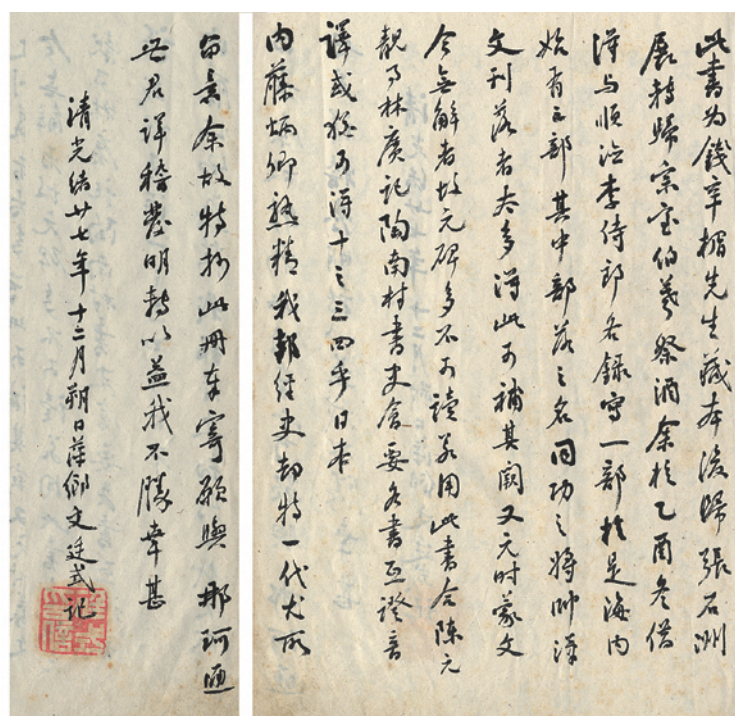
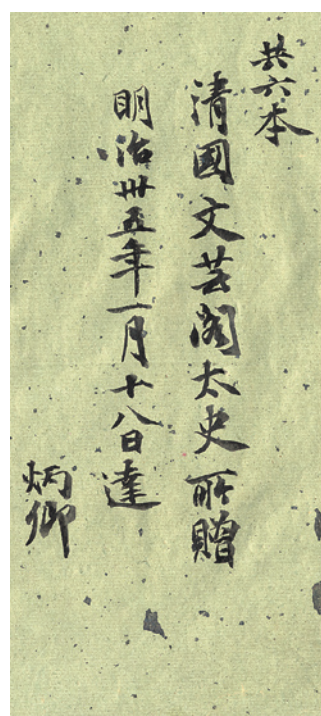


漢字と情報

No. 7
2003・10



京都大学人文科学研究所 Documentation and Information Center for Chinese Studies (DICCS)
附属漢字情報研究センター Institute for Research in Humanities, Kyoto University



- 辞書を「読む」
- 『通鑑紀事本末・卷第八』 宋版本寄託覚書
- 『説文解字繫伝』 データベース構築の試み
- 科学史研究室蔵の能田・藪内関連資料
- 人文研のアーカイブス(7)『元朝秘史』

辞書を「読む」

富谷 至

「漢字と情報」と銘打った本冊子の意図するところ、および現在の私の立場からすれば、以下に述べることは、いささか不似合い、不穏当であるかもしれない。本冊子の意図とは、附属漢字情報研究センターの目指す漢字・漢籍・研究文献などの電算処理に関する成果を伝えることを目的としているからであり、そのセンターの主任という立場に置かれているのが私だからである。

電子辞書というものが急速に普及してきた。高校生をふくめて外国語を学習する者は、もはや旧来の紙本の辞書よりも電子辞書を使う者が大半を占める。大学生の愚息が、新しい外国語を学ぶに当たって、独・仏の電子辞書が英語のそれよりも高いと文句を言っていた。英語以外の外国語のものは3万円から、5万円するのだ。おそらく、彼らの頭の中には、紙本の辞書を買うなどという発想は埒の外、手垢で汚れた、しかし手になじんだ辞書になってこそ外国語が上達するのだ、などとの説教は、携帯のメールを使わず、手紙を書けと言うに等しい^{たわごと}妄言であろう。

かく言う私も遅まきながら電子辞書を使い出した。今夏、国際シンポジウムを開き、そこで発表・討論した時、これは実に役に立った。老化とともに物忘れがひどくなって、日本語が出てこず、ましてや英語の単語はすぐに思い出せない。二種類の外国語を調べたり、言い換えのための類語を見るときにもこの電子辞書は多大の効力を発揮したのだ。また、この前にドイツに出張した折、ドイツ語がわからない私にとって、日常の簡単な単語を即座に調べるのに、これまた有り難い道具であった（以前、スウェーデンのスーパーで歯磨きを買おうとして、チューブに入ったネリモノが靴のクリームか歯を磨く目的の代物か見分けるのに困ったことがあった）。

電子辞書が重宝されるのは何よりもポケットに

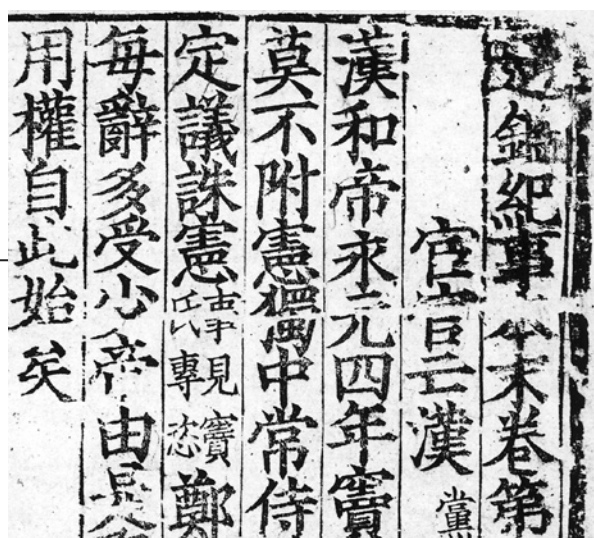
入るほどの利便性と、スペルをすべて正確に引かなくても、初めの2、3文字を入力しただけで目的の単語がでてくるという速効性にあるのだろう。レストランで大きな紙本辞書を取り出して、なかなか出てこない単語を時間をかけて調べるといった不格好を——あまりそのようなシーンを見たことがないが——演じなくともよいのだから。

ところで、これも私事にわたることだが、私は片道2時間の通勤時間をかけて毎日通っている。どうしてそのような小旅行を毎日しているのか、その理由はさておき、誰にも邪魔されない、全く一人の時間が満喫できる4時間の往復時間は実に楽しい（それ故、私は携帯電話など持とうと思わない）。私は、その時間の一部を毎日英語の辞書を読むことに使っているのだが、英語に限らず、辞書をその第1ページから読むことは、存外楽しく、またためになると思う。単語、語法、表現にかんする知識を得るだけでなく、辞書に凝縮している言語の世界、さらには異文化世界を体系的に知る上で、少なくとも私には、辞書はまことに面白い読み物の一つなのだ。

電子辞書を愛用してから、それが紙本の辞書よりはるかに持ち運びに簡便であるが故、電子辞書を読むことに切り替えた。ところが……である。どうも調子がちがう。初めのうちは、慣れていないことからくる違和感だと思っていたのだが、よく考えてみると、どうもこれは本質的なところで、電子辞書と紙本辞書とは相違点があるではないかと思うにいたった。確かに、前者はピンポイントで単語を調べるのには、絶大の効力を発揮する。しかし、それは、紙本の辞書から得る言語世界、異文化世界の体系的把握には、向いていないのではないだろうか。換言すれば、電子辞書は調べる辞書であっても、読む辞書にはならないのである。

ことからは、辞書だけではない。我が研究センターの漢籍目録、文献目録についても同じだろう。目録にはそれが持つ体系、世界があり、一冊の目録にそれは凝縮されている。その世界を理解するためには、冊子本が今なお有効性を有するのだ。

（人文科学研究所教授・センター主任）



『通鑑紀事本末・卷第八』 宋版本寄託覚書

櫻井謙介

私は、かつてシオノギ製菓の研究所で天然物化学研究に従事し、武田健一研究所長の薫陶を受けました。1983年のこと、赤堀昭先生が退社された後を継いで、漢方を学べ、その前に漢文を勉強しろとの社命が下り、研究所勤務のかたわら京大人文研の科学史研究班に参加するようになりました。

山田慶児先生が主催されていた研究会にはいろいろな分野の方が招待・参加されておりました。医学史、道教、建築、植物、数学、天文学、農業、地理等々。それぞれが専門の研究を行うと同時に、一つのテキストをそれぞれの専門の立場から読み、注をつけ、批判を加える、また専門外の発表も聞かせてもらえる。実に楽しく、勉強になりました。共同研究会とは本来、こういうものなのでしょう。皆さんの研究を理解するためには沢山の専門書を読まねばならず、新本では財布が保たず古本屋の利用を覚えました。

さて、6年ほど前になりますか、宋版本を入手したのは、大阪市天王寺区古書店楽人館です。時々韓国・朝鮮系の書籍がありますが、一般書が多く、線装本はほとんどありません。ごく普通の古本屋です。棚に並んだ一冊の大型の線装本に目が行きました。開いてみると巻首に『通鑑紀事本末・卷第八』とあり、品のある文字が並んでいます。文字を追っていくと欠画のある「徴」や「瑗」が見つかりました。これはどうも宋朝の皇帝、仁宗と欽宗の避諱のようです。当時私は龍谷大学大宮図書館の和漢古典籍目録作成のお手伝いをしておりまして、茨城大学の真柳誠先生から古

典籍の手ほどきを受けておりました。それで、事によると宋版ではあるまいかと思った次第です。

しかし考えてみますと、宋版が場末の（失礼）店にあるはずもなく、国内の宋版本はほとんどすべて所在が明らかとなっているはず。ま、字体が気に入ったからいっか、と売値五千円を四千円に値引きさせて購入しました。それから真柳先生に「宋版手に入れたよ」と。「見せろ」とおっしゃるので、レントゲン・フィルムの空箱に入れて送りますと「本物だ。こんな扱いするな。桐箱に入れろ。ナフタリンは使うな。本物の樟脳を使え」。それで桐箱を誂え納めました。五千円也。本より高うつきましたナ。

後に慶応大学斯道文庫の尾崎康先生に見ていただきましたところ、宋版に間違いはない、この版は日本にはないものだ、版木のひび割れの具合から後代の刷りだろうとのことでした。本物の宋版となりますと流石に私には荷が重く、預け先を探しました。たまたま研究会仲間の武田時昌先生のお薦めもありまして、漢字情報研究センターにお預かり願うことにしました。

当時私は父を引き取って（再就職の口をなげうってまで）介護しておりましたので、天がその孝心を嘉して宋版を授け給うたのでしょうか。しかし時に陰で父を罵ることもありましたので、その分差っ引いて端本……でしょうか。

古本屋を丹念に探しておりますと、年に何度か当たりがあります。重要と判断した書籍はしかるべき図書館に寄贈することにしております。今回の宋版は所持していることを自慢したくて、まだ寄託にしております。いずれ死んだ後は寄贈するつもりでおりますので遺族とご相談ください。

なお、会社を定年退職してからは「北奈良式書物貸観処 文庫 半覚齋」主人を名乗り、兼ねて文献調査コンサルタント「鷺屋 芝蘭堂」を経営しておりました。来年には、家人ともども故郷の北海道に帰ることを予定しております。所蔵書は処分せよとの厳命なので、昭和薬科大学に一括して寄贈することになりました。縁あって集まった書物たちと永の別れを惜しんでおります。呵々

『説文解字繫傳』データベース構築の試み

坂内千里

近年、『四庫全書』をはじめとして電子化されたテキストが数多く作成されており、出典調査などは非常に効率良く行えるようになった。例えば、以前「説文」の「徐注」とされる「精者曰綿、繭內衣護蛹者、與其外膜緒雜爲之曰絮」という一条がどうしても見つからず、已むを得ず出典不明としたが、⁽¹⁾『四庫全書』の全文検索で、苦もなく「𦃟」（卷十三 日部）の条の小徐注であることが提示された時には愕然としたものである。しかし、当然のことながら、全文検索は万能ではない。そこで、研究成果を取り入れつつ発展させることのできるデータベースを作成したいと考えはじめ、また、利用者の立場からは、画像による検索ができないものと漠然と考えるようになった。ちょうどその頃、漢字情報研究センターでの内地研修（2002年5月から翌年2月）が決まり、漠然としていた計画を具体化できることになった。以下は、その研修の簡単な報告である。

「マークアップを利用した、画像による検索が可能なデータベースの作成」を基本方針として作業を開始し、先ずその基礎となる画像ファイルを作成した。祁刻本を底本とし、1葉を1ファイルとした。

次に、マークアップの形式と効率的な作業手順を考察するため、取り敢えず巻一を素材として、作業を行うことにした。「画像による検索が可能」という基本方針は、電子テキストと画像それぞれにマークアップを施し、両者を統合することで実現される。そのため、巻一の電子テキストを作成することが必要となる。そこで、OCRソフト「丹青 白金版」（NewSoft）を利用して、先に作成した画像ファイルの電子テキスト化を行った。やはり祁刻本1葉を1文章ファイルとする。

これらのファイルを XEmacs CHISE⁽²⁾ 上で UTF-8 にコード変換した後、校訂作業を施し、更に巻一分をまとめて1つのファイルに編集した。ここで XEmacs CHISE を利用したのは、Unicode 外の文字も画面上表示可能であり『説文解字』のようなテキストの編集には現時点では最適であると考えたこと、画像ファイルのマークアップのために必要なツールは XEmacs CHISE⁽³⁾ 上で動くよう設計されていること等のためである。

完成した巻一の電子テキストに、TEI (Text Encoding Initiative) のタグセットを利用して視覚的構造と論理的構造の両面からマークアップを行った。視覚的構造としては、“巻——ページ——行”，論理的構造としては、階層的に“篇——部——文（その下位構造として、親字・説解・徐注・反切）——重文（下位構造：親字・説解）”にタグ付けした（図1）。ただ、TEI は主として欧米系の文献を対象としており、『説文解字』のような中国の字書の記述に相応しいタグは用意されていない。そのため、論理的構造を適切に記述することが困難で、〈p〉属性を多用し type で区別するしかない；本来〈trailer〉で記述すべきであると思われる巻末の題「説文解字通釋卷第一」もタグの使用制限のため〈P〉属性で記述せざるを得ない等の、限界がある。中国の文献を記述するための新たなタグセットの開発が待たれる。

最後に、画像のマークアップ作業を行った。最初に作成した画像ファイルを OCR ソフト「読取革命 ver.7」（Panasonic）にかけ、透明 PDF ファイルを作成する。やはり祁刻本1葉を1ファイルとする。これを XEmacs CHISE 上で TID (Text Image Description) 形式に変換し、行番

(1)「『古今韻会举要』に引く『説文解字』について」（『漢語史の諸問題』京都大学人文科学研究所 昭和63年3月）361頁

(2)旧 XEmacs UTF-2000。http://kanji.zinbun.kyoto-u.ac.jp/projects/chise/xemacs/。

(3)詳細は守岡知彦「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利用 第14回研究セミナー』京都大学人文科学研究所付属漢字情報研究センター 平成15年3月）参照。

号等の入力及び各文字の座標の修正などの編集を行う。更にそれを SVG エディタ Sodipodi で表示可能な SPS 形式に変換し、Sodipodi 上（図 2）で線のずれなどを確認しながら、TID ファイルの文字座標を修正する。この作業を繰り返して文字の座標を確定する。

以上が、巻一に対して行った作業の概略である。この過程で、1) 視覚的構造のマークアップが非常に煩雑である；2) 同じファイルを二度 OCR にかけるのは、かなりの作業負担となる；という問題点が明らかになった。そこで、巻二以降の作業工程では、OCR を一本化し、可能な範囲でマ

等の論理的構造の一部のマークアップを自動的に行うことにした。漢字認識率が落ち、正字変換を行う必要も生じたため、文字の編集に若干時間がかかるようになったが、全体的にはかなり効率良く作業が行えるようになった。

基礎作業の形式・手順は、ほぼ確立した。今後は、この工程に沿って作業を進めると同時に、明らかになったいくつかの問題点を解決する方法を考え、また、並行して作成している引用書名のデータベース等をどのようにして利用するか、データをどのような形で表示するか等の課題に取り組んでゆく必要がある。

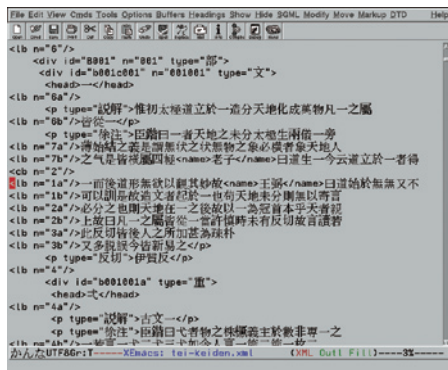


図 1： TEI によるマークアップ例（巻一の一部）

視覚的構造の頁（`<p n="*">`）・頁の表裏（`<cb n="*">`）・行（`<lb n="*">` ab は双行を表す）はインデントを行わず、論理的構造はインデントしてその包含関係を示す。論理的構造としては篇の下に部（`<div type="部">`），その下に文（`<div type="文">`），重文がある場合には更に重文（`<div type="重">`）を設けている。また、文・重の内部は、親字（`<head>`）・説解（`p type="説解"`）・徐注（`p type="徐注"`）・反切（`p type="反切"`）に分けている。巻二以降では、視覚的構造及び親字・説解・徐注・反切のタグは自動的に附加されるようになっている。

ークアップの自動化を図ることとした。OCR ソフトに関しては、「丹青」は漢字認識率はかなり高いが、レイアウト認識に大きな問題がある；「読取革命」は漢字認識率が「丹青」より低く、正字出力が出来ないが、文字の座標情報を記述できる；というように、それぞれ一長一短がある。画像の処理にどうしても必要であるため、「読取革命」に一本化し、電子テキストと画像の処理を同時進行させることにした。つまり、透明 PDF ファイルを TID 形式に変換した後、行番号等の編集をすると同時に、正字変換等の文字の編集を行うことにした。更に、編集後の TID ファイルを TEI の XML 形式に変換し、その際に“ページ・行”等の視覚的構造と“説解・徐注・反切”

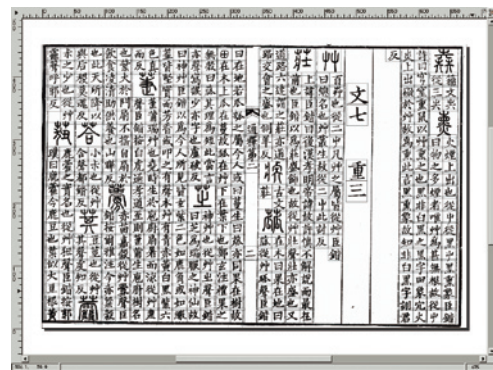


図 2： Sodipodi 画像マークアップの例

自分にとって利用しやすいデータベースを作ろうと始めた計画ではあるが、実際に作成し始めて、想像以上の時間と労力が必要であることに驚いた。しかし、同時に校訂を経た、二次利用可能なデータやデータベースを作成していくことの必要性をも実感した。データベース構築の試みはその緒に就いたばかりではあるが、できるだけ早い時期の完成を目指したい。

最後に、快く内地研修を受け入れ、さまざまな機会を与えて下さった武田時昌教授、TEI に関する指導・協力を戴いた Christian Wittern 助教授、画像マークアップを中心に全般に互に指導・協力を戴いた守岡知彦助手に、ここに改めて謝意を表します。（大阪大学言語文化部助教授）

科学史研究室蔵の能田・ 藪内関連資料

東郷俊宏

中国科学史研究に従事するものにとり、2000年は特別な意味を持つ年といえる。この年の6月、藪内清が94歳で逝去し、10月には『新城文庫目録』が理学部宇宙物理学科の富田良雄助手を中心とする編纂委員会によって頒布されたのである。現在宇宙物理学科が保管する新城文庫は、新城新蔵が生前使用した手帳や多数の書簡を含み、新城の研究経過を詳細に知る上で不可欠の資料である。また人文科学研究所科学史研究室には藪内が研究に用いた写真資料、自筆原稿等が千点以上残されているほか、世代的に上述した二人の間にあって、東方文化研究所における天文暦算史研究を指揮した能田忠亮博士の研究資料が存在する。上の新城文庫とあわせれば東アジア天文学史、科学史研究をリードした三巨人の研究軌跡をたどる上での貴重な資料群となるだろう。研究室では今後これらの資料の整理を進める予定であるが、今回はその中から能田が戦時中関わっていた新東亜暦（国民暦）編纂事業に関係する文書を紹介したい。

1926年に京都帝国大学理学部宇宙物理学科を卒業した能田は、3年後に東方文化学院京都研究所が設立されると、指導員として就任した新城のも

とで旺盛な研究活動を開始する。やがて新城は1935年に上海自然科学研究所第二代所長に就任したため、その後は自分と同じ宇宙物理学科を卒業した藪内清とコンビを組み、1948年に研究所を去るまで暦学研究をともに行った。

近代に入ってから東アジア諸国は、西洋に起源を有するグレゴリオ暦を正式な暦として採用する。日本においては早く明治5（1872）年に採用され、中華民国では1912年に採用された。しかしグレゴリオ暦では正月が意味をなさず、農業立国である中国や日本の暦としては実用性に乏しいとして、日本を主導とする暦編纂の必要性が唱えられる。その中心となったのが能田を筆頭とする宇宙物理学科を卒業したメンバーたちであった。

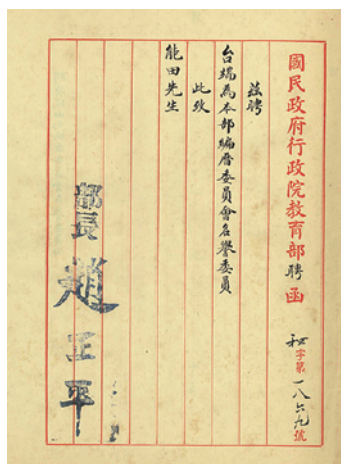
新城が昭和13年に上海で客死した前後より、南京にある紫金山天文台の復興計画が持ち上がるが、これに宇宙物理学科出身の森川光郎、高木公三郎が参加する。森川は藪内と同期、高木はその七年後輩に当たる。そして昭和15年、能田は外務省文化部からの助成金を得て東洋暦術調査を始めるとともに、新東亜暦編纂の主要メンバーとしての役割を演じることになるのである。

写真は能田を国民暦編纂委員会の名誉委員として迎えるために国民政府教育部が発行した任命状である。これに従って能田は昭和15年8月に南京にわたり、9月に同地で開催された国民暦編纂会議に日本側委員として出席した。先述した森川、高木両人も日本側委員として参席している。

9月中に帰国した能田に宛てられた高木、森川の書簡からは、南京における編暦作業の進捗状況と、暦編纂事業の開始に際し、その主導権をめぐる東京側と京都側の専門家との間で微妙な対立状況があったことなどが窺われる。

新東亜暦は結局採用されず、計画は途中で頓挫した。これに対し、能田等がどのような反応を示したのか、現在のところ確かな資料はないようである。能田は昭和23年3月に研究所を去り、代わって藪内が新生人文科学研究所における科学史研究室の屋台骨を背負って立つのである。

（人文科学研究所助手）

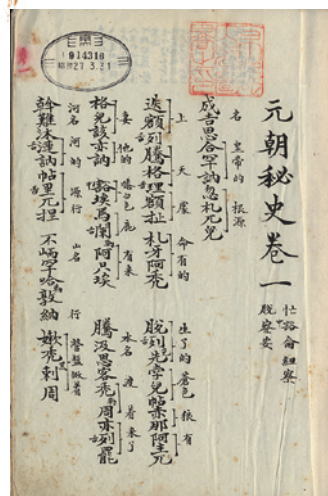


人文研のアーカイブス（7）
元朝秘史十卷 續集二卷

元 闕名 撰

鈔本

梶浦 晉



『元朝秘史』は、モンゴル族の遠祖の伝承からチンギス・ハーンの時代にいたるまでの事績（後にオゴタイ・ハーンの事績が付される）を記した史書で、歴史のみならず、文学・語学などモンゴル研究全般にかかわる重要文献として知られている。元来はモンゴル文字で書かれ、13世紀前期に成立したと推定されている。しかし現存の本は、明代初期に編纂されたもので、モンゴル語を漢字音であらわした文と、漢語による逐語訳および意訳の部分から構成されている。正集10巻続集2巻の12巻本と『永楽大典』所収の15巻本がある。

ここにあげた『秘史』は、内藤湖南（虎次郎）旧蔵書である。本所には、数量はさほど多くはないが、湖南旧蔵書のうち、満蒙関係書がまとまって収蔵されている。

掲出の『秘史』は、光緒中の鈔本でさほど古いものではないが、清の文廷式が湖南の請をうけ、所蔵の本を転写し送ったもので、第一冊第一葉に文の識語がある。文廷式、字は芸閣。咸豊6年～光緒30年(1856～1904)。江西萍郷の人。光緒16年(1890)の進士。一時、日本に亡命する。日本滞在の記録に『東遊日記』がある。

日本における『秘史』の最初の訳注は、那珂通世『成吉思汗実録』（明治40年）であるが、那珂は湖南からおくられた転写本を用いて研究をしており、湖南旧蔵本は『秘史』研究史上、意義のあるものといえる。なお、那珂が用いた本は筑波大学に収蔵されている。

この『秘史』は、明治41年6月7日開催の史学研究会《那珂博士追悼会》の展覧に出陳されている。現在、本所にはこの本と併せて、以下の資料が付されているが、同時に展陳された〈文廷式手簡〉はない。（『目録書譚』所掲の書簡か）

那珂通世書簡四通（内藤虎次郎宛）

那珂通世手校成吉思汗実録校正刷

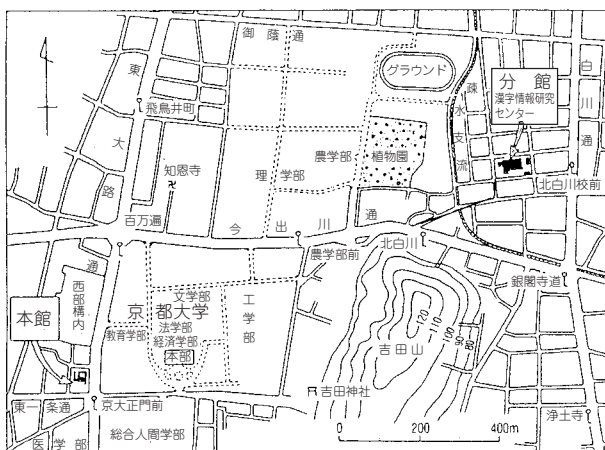
大阪朝日新聞切抜（明治35年2月3日）

大阪朝日新聞切抜（明治40年4月28日）

那珂博士追悼会列品目録

肅親王書翰（那珂博士宛：石印）

（センター助手）



- ・2003年度の21世紀 COE プロジェクト複合・新領域に高田時雄人文科学研究所教授を拠点リーダーとするプロジェクト「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」が採択された。本センターのスタッフも中心的な事業推進担当者に加わっている。詳細はホームページ(<http://coe21.zinbun.kyoto-u.ac.jp/>) 参照のこと。
- ・本年度は本センターが五センター会議の当番校であるが、6月13日(金)午後2時より全国文献・情報センターワーキンググループの会合を行った。また、9月19日(金)午後2時より第39回全国文献・情報センター長会議を京大館101号室にて開催した。
- ・本年度の漢籍担当職員講習会は、10月6日(月)～10月10日(金)に初級を実施し、19名の修了者があった。所外からの講師には、宮沢彰国立情報学研究所教授、井上進名古屋大教授、森賀一恵富山大助教授を迎えた。中級は11月10日(月)～11月14日(金)に行う予定である。文部科学省から研究振興局情報課学術基盤整備室長の當麻維也氏が来られるほか、所外の講師には、高津孝鹿児島大教授、平田昌司文学部教授、高島航文学部助教授をお願いすることになっている。なお参加予定者は20名である。
- ・11月26日(水)～27日(木)に2003年度全国文献・情報センター人文社会科学学術情報セミナーが一橋大学国際共同センター1階多目的ホールにて開催される予定である。全体のテーマは「センター所蔵資料の活用と人文社会情報」であり、本センターからはセッションⅡ『『東洋学文献類目』の編纂と歴史』において、井波陵一教授、安岡孝一助教授の二人が研究発表を行うことになっている。
- ・最新のセンター刊行物
『東洋学文献類目』2000年度版(2003年3月)
『東洋学文献類目』2000年度補遺版(2003年3月)

<http://www.kanji.zinbun.kyoto-u.ac.jp/>