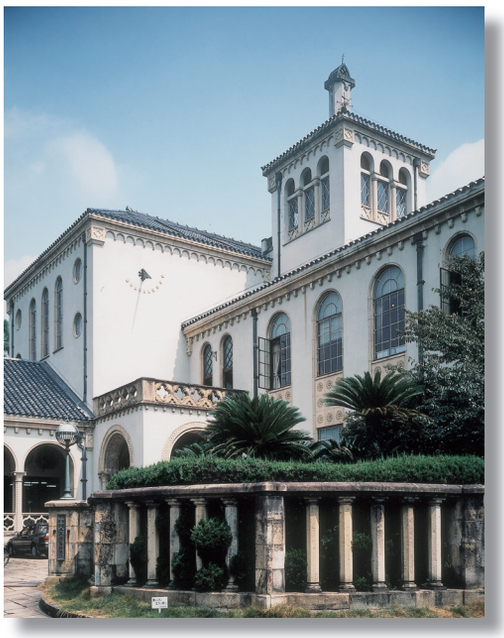
Chinese Characters and Culture
CCC
漢字と文化

漢字と文化

漢字文化の全き繼承と發展のために

京都大學 21 世紀 COE 東アジア世界の人文情報學研究教育據點

創　刊　号



目　次

文部科学省平成15年度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―」に寄せる期待　森 時彦 …… 2
漢字文化の全き繼承と發展のために　高田時雄…… 3
テキストと文字，ストリームとコード　安岡孝一…… 5
唐代研究ナリッジベース　クリスティアン・ウィッテルン／井波陵一…… 6
漢字情報学事始め　武田時昌…… 9

文部科学省平成15年度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—」に寄せる期待

人文科学研究所長　森　時彦

高田時雄教授を拠点リーダーとする「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，文部科学省平成15年度21世紀 COE に学際・複合・新領域分野で採択された。世界最高水準の卓越した研究拠点の確立を使命とする人文科学研究所としては，その責任の一端を果たすべき貴重な機会を与えられたことに心から感謝したい。しかも採択された分野が，人文情報学という学際的な新しい学問領域であることは，とりわけ喜ばしく思われる。

本プログラムの事業推進担当者は，人文科学研究所，人間・環境学研究科，東南アジア研究センターの三部局にまたがるが，主力は人文科学研究所附属漢字情報研究センターのスタッフである。漢字情報研究センターは，平成12年にそれまでの東洋学文献センターを改組，拡充して設置したものである。

漢字文化と情報科学の融合をめざして設立されたこのセンターは，文系の附属施設においてはあまり前例のない試みとして，工学系出身のスタッフを複数招聘するとともに，外国人スタッフも招いて，国際的かつ学際的な研究体制を整えることに意を用いてきた。

周知のとおり国立大学附置研究所，とりわけその附属施設は，時代に先駆けて絶えず新しい研究分野を切り開いていくことを要請されている。設置後３年半を経た漢字情報研究センターは，このような要請に応えるため，スタッフ一同が協力して日本全国漢籍データベースの構築，漢籍の電子テキスト化など，さまざまな新しい事業を立ち上げてきたが，本プログラムの採択はそのチャレンジにさらに弾みをつけることになるものと確信する。

本プログラムでは，（１）東アジアの文字に関する人文情報学的研究，（２）漢字文献ナリッジベースの構築，（３）東アジア人文情報学の人材育成という三大事業を推進することになっている。いずれも，現在の漢字情報研究センターの事業をさらに国際的，学際的な方面に拡大，発展させる可能性をひめた計画である。

本プログラム自体は５年のプロジェクト研究ではあるが，計画している事業が予期どおりの成果をあげるものとの期待のもとに，その成果をふまえた研究分野の拡大，発展にあわせて，現在の漢字情報研究センターを改組，拡充して「東アジア人文情報学センター」（仮称）を設置することもすでに構想されている。拠点リーダーをはじめとして，事業推進担当者各位が，計画にみあう大きな研究成果をあげるべく尽力されるよう祈念してやまない。

漢字文化の全き繼承と發展のために

高田時雄

21世紀 COE プログラム「東アジアにおける人文情報學研究教育據點」は，京都大學人文科學研究所附屬漢字情報研究センターを中核として構想されたものであるが，學内からは人間・環境學研究科および東南アジア研究センターにも參加していただき，多部局にわたる連合體として組織されている。

漢字情報研究センターは，それまでの東洋學文獻センターを改組擴充することで2000年４月に發足した。このセンターは本來，京都における中國學の長い傳統と豊富な中國文獻資料の蒐藏とを基礎に，東洋學文獻の蒐集・整理・公開を主たる任務として1965年に設置されたものである。今回の改組では，情報學を全面的に取り入れた國際的發信力を持つ組織として發展させること，そしてとりわけ漢字文獻に特化した分野での國際的センターとして機能することを目指したものであった。したがって上記の課題により21世紀 COE として認定されたことは，たいへん時宜に叶ったことであるばかりでなく，今後の發展の原動力になることを期待している。

さてプログラム全體の目的とするところは小文の標題に掲げた「漢字文化の全き繼承と發展」である。わが國では早くに中國から漢字を受け入れて，その上に獨自の文化を築いてきた。あらゆる制度や文化は漢字を通じて構築され，あらゆる思想や文學は漢字によって表現されてきた。その遺産は紛れもなく巨大である。漢字文化の存在はもちろん日本だけではない。本家の中國をはじめとして廣く東アジアに全域に長く且つ深い影響を及ぼしている。漢字を用いて書かれた書物や資料の中に體現され，東アジアの人々の血肉となってきた「漢字文化」の全面的繼承を追求することは，我々に與えられた任務でなければならない。

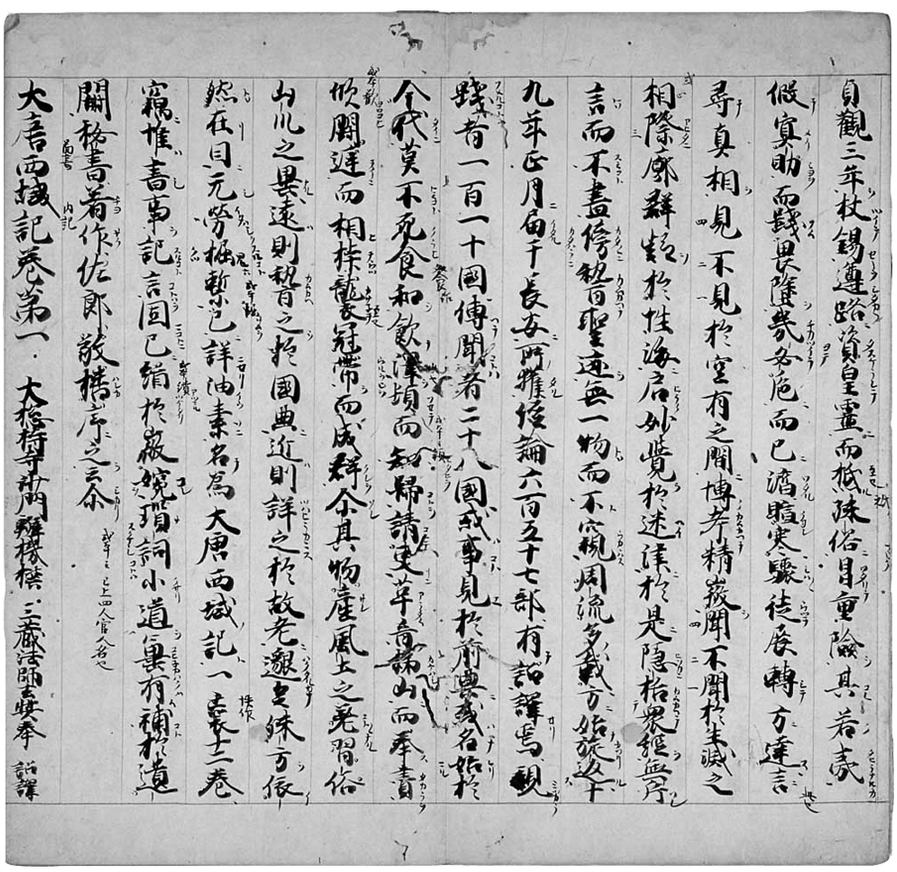
ところで情報學によって漢字文化の發展を目指す上での課題は數多く，個別には様々な問題があり得るが，その中でも當面考慮すべき重要な事柄が二つあると思われる。その一は多くの情報の中から必要とするものを素早く的確に取り出すための仕組みを考えることである。コンピュータが出現してから，やがて漢字を使うことが出來るようになったのが第一期。その頃は文字數の不足が大きな問題であった。日常の現代文を處理する程度のことなら大した不便はなかったが，厖大な文獻を相手にしなければならない研究者には，漢字不足は極めて頭の痛い問題だったのである。しかしそれも近年では大規模な文字コードが實用化され始めたお陰で次第に解決しつつある。ところが一方で巨大な規模のテキスト・データベースが，特に中國を中心として，陸續出現するようになった。研究者はやがて電子テキストのコーパスの巨大な海の中で溺れはじめようとしている。折角の電子テキストの利點を充分に生かしきれない状況が目の前にある。われわれのプログラムでは，多様な檢索を可能とする技術開發を含め，在來のデータベースから更に進化したナリッジベースの構築を進めるべく，すでにその構築に着手した。

第二の問題は我々が傳承してきた漢字文獻を過不足なくデジタル化して後世に傳えることが出來るような方法を考えることである。文獻というものはテキストだけから成立しているのではない。龜甲牛骨から始まり，竹帛，金石など，漢字がその上に書かれた素材は多種多様である。また筆硯紙墨などの用具も視野に入れなければならないだけでなく，書寫の方向や書式など文字外の情報で研究上重要なものはさまざまにある。これがデジ

タル化によって，裸のテキストだけになってしまったのでは，トータルな文化の繼承は難しい。その意味での漢字文獻の全面的デジタル化が可能となる方法を模索しなければならない。

いずれにせよ漢字および漢字文獻の情報化はローマン・アルファベットのそれに比べていちじるしく立ち後れており，漢字と情報に關する諸問題の解決は，すべての漢字を用いる学問分野において重要であり緊急性が高い。しかしその解決は一朝一夕には難しい。そのため，漢字の情報化という課題と主體的に取り組むべき人材を育成することも同時に考えなければならない。また新しい人材は日本という一國においてしか役立たないような存在であってはならない。國境とともに言葉や文化の違いを乗り越えて，東アジア全域で活躍できるような人材を出來る限り數多く世の中に送り出すこと，これは我々にとってまた別個の大きな課題である。そのためには國際的な規模での連携が不可缺であるが，それもすでに動き出している。

思えば20世紀は漢字と漢字文化にとって激動の時代であった。各國の文字改革によって，漢字圏の分裂が起こり，それまでの統一が損なわれた。しかし一方ではまた甲骨文字を初めとして，未知の漢字文化遺産が地下から大量に出現した發見の世紀でもあった。20世紀の100年で漢字の風景は大きく變わったと言えるだろう。われわれは今，新しい世紀の入り口に立って，情報化時代における漢字文化の繼承發展に貢獻したいと思う。



大唐西域記巻一　京都大学人文科学研究所所蔵抄本

テキストと文字，ストリームとコード

安岡孝一

ある文書をデジタル化する，ということを考える際に，避けて通れないのが，テキストをどのように扱うかという問題である。この問題に対するもっとも極端な答の一つは，テキストなど扱わず全てを画像にしてしまう，というやり方である。もう一つの極端な答は，印刷に関する情報を完全に捨て去ってしまって，ピュアなテキストだけを取り出す，というやり方である。前者を「画像データベース」，後者を「全文テキストデータベース」とでも言えば，多少理解しやすいかとも思う。

この二つの極端な答の中間解，それも単なるハイブリッドではない解を，いかに作り上げるか，それが『漢字文化の全き継承と発展』の主眼の一つである。一般ウケする言い方なら「全文検索可能な文献画像データベース」といった趣のものだ。しかし，実際に着手してみると，これはナカナカの難物である。テキストというものをそもそもどうとらえ，それをどうデジタル化すべきか，という根本的な問題に立ち至るのである。

旧来の情報科学においては，テキストは，文字のストリームとして実現されている。文字はそれぞれ，一定長のビットパターンすなわち文字コードで表現されている。つまるところ，鑽孔テープでテキストをやりとりしていた時代——ひとまとまりの鑽孔パターンが文字であり，それを一次元に並べたものがテキストであった時代——を，いまだにひきずっているのである。もちろん現代においては，テキストはXMLによって拡張されているし，文字コードはUnicodeによって拡大の一途にある。しかし，本来の文書の持つ多重構造，あるいは漢字の持つ多層的意味を扱おうとすると，このような拡張では歯が立たない。

文書の持つ多重構造の例として，複数行に渡る長い割注を考えてみよう。論理構造から見た場合，割注は，文の中に入った注である。物理構造から見た場合，割注は，それぞれの行の中に入った複数の行である。「文の中に入った注」という構造，あるいは「行の中に入った複数の行」のいう構造の，いずれか一方ならば，XMLを使ってごく簡単に表現できる。しかし，これらの構造を同時に１つのXMLで表現することは本質的に不可能であり，どちらかの入れ子構造を犠牲にせざるを得ない。つまるところ，ある１つのXML表現は，文書の持つ多重構造のうち，ある１つしか表しえない。それが，ストリームによるテキスト表現の本質的限界でもある。

漢字の持つ多層的意味の例としては，避諱をどう扱うか考えてみるのがいいだろう。漢字は本来，形音義を一体として扱うべきものだが，避諱はむしろ，形を義から遠ざける操作である。その結果，避諱に対する検索は，避諱の形としての検索（そのままの字づらでの検索）と，義としての検索（本来の意味での検索）の２面性を必要とする。しかも避諱は，時代によって，地域によって，あるいは文書によっても異なる。そのようなものを，１字たかだか32ビットの情報量で表しきれるわけがない。つまるところ，漢字の持つ多層的意味を検索に活かしたいとするならば，文字の表現に文字コードを用いるのはナンセンスである。

ではこれらの問題に対する解決法があるのか，と問われれば，それを研究するのがCOEの本旨でしょう，と答えて，本稿を締めくくることにする。まさに「今後の研究に期待されたい」なのである。

唐代研究ナリッジベース

クリスティアン・ウィッテルン／井波陵一

はじめに

大量のデジタル・データの利用が可能になることで，中国研究にまったく新たな方向性が開かれる――そうした期待が高まったのは数年前である。そして現在では，『デジタル四庫全書』をはじめとして，多くのテキストがデータベース化され，その一部についてはインターネット上で無料で利用できるようになった。ところが実際にはあまり役に立っていない。なぜなら，こうしたデータベースは文字列検索こそ可能だが，それ以上の分析ができず，検索結果が大量になると処理が難しいという問題を抱えているからである。

ナリッジベースとは

唐代に関する情報を，包括的な電子アーカイブとして提供することにより，こうした状況を改善することを目指すのが「唐代研究ナリッジベース」である。歴代の王朝の中から唐を選んだのは，それが東アジア全体における制度的・文化的基盤を提供した点で抜きん出た存在であるからにほかならない（たとえば日本の国家形成における役割を考えてみるだけでも，その重要性は十分納得されよう）。研究で主に利用されるのはウェブ・アプリケーションであるが，他のインターフェースについても開発が予定されている。

提供される情報は，初期段階では大半が文字情報となるが，作業が進むにつれて原資料の画像やデジタルマップなどが付加されることになっている。このナリッジベースの最大の特長は，情報を相互に結びつけるのに柔軟で革新的な方法を採用していることだ。

ナリッジベースに収録される情報の概念編成は次のような情報軸に沿って行われる。

- 唐代の人々の姓名，履歴
- 地名および経緯度，地方行政機関，デジタル地図
- 唐代に作られた作品（テキスト，美術工芸品，建築物）
- 暦年および時間
- 国家的重要事件

ナリッジベースの情報は，これらの情報軸の複数にまたがるものが，すべてとは言わないまでもかなりの量に上っており，それらが内部的に相互リンクされてウェブのような構造を成している。さらに，これらの情報は階層的オントロジーにより編成されており，これにより階層内の位置や他の情報との関係をもとにアクセスすることができる。このような階層を構成する項目は，都市など地理的な位置を示す情報の場合はその都市を管轄する上位の行政機関であるし，人名の場合は家系や，出身地，所属する学派，また，僧侶の場合はこれに授戒や嗣法の系統が加わることになる。

こうした階層を縦横に活用することで情報へのアクセスや検索が充実したものとなる。たとえば，則天武后が国号を周と改めて皇帝に即位した時期に権力を握った官吏について概要を得たり，唐朝復興後のその行く末を調べることは，従来の研究ツールではきわめて困難である。そのためには何百人もの伝記を調べ上げ，一人ひとりについてその運命を辿った上で，総合的な概観を得ることが必要だからだ。これに対して，「唐代研究ナリッジベース」を使えば，すべての関連情報をたちどころに融合・咀嚼し，しかも結果を分かりやすく視覚的に表示することもできる。

また，白居易が任官中に出会ったすべての人々

を追跡し，彼らが作った詩をトータルに分析することにより，詩人白居易の同時代的影響について明らかにする，といった利用法も考えられる。その場合，分析対象となる詩を，まず白居易の知己の作とそれ以外の人々の作という２つのグループに分け，そのうえで相互に比較分析することも可能だろう。

もちろん，このナリッジベースを活用できる分野やテーマはそれ以外にも数多くある。ナリッジベースの構成は，データに対して特定の解釈を押しつけることなく，むしろ様々な異説や相矛盾する見解すら成立可能になるように配慮されているからだ。それはあたかも一枚の紙がそこに書きつけられる内容について何ら制限を加えないのと似ている。

もっと技術的な観点から見てみよう。このナリッジベースは次のような３つの主要な柱から構成されている。

・唐代文化に関係する第１次テキストのコレクション―― TEI タグ集を使ってマークアップされ，XML データベースを通してアクセスできる。
・これらのテキストが伝える事物の情報やその他の多くの情報（いわゆるメタ・データ）――これはトピック・マップで管理される。
・ナリッジベースの操作と，テキスト，メタ・データその他すべての項目を結合するためのシステム

したがってナリッジベースの構築に当たっては，その内容のみならず，適切なインターフェースをも構築することが必要であり，最善の結果を得るにはその両方を完全に統合しなければならない。

テキスト・コレクション

テキスト・コレクションの部分の作業は次のように進められる。多くの場合，既存の電子化テキストを利用する。ここでとくに重要なのが，中央研究院歴史語言研究所のテキスト・コレクションだ。ナリッジベースではこれらのテキストに基本的な構造マークアップを施し，データベース内でのアクセスやリンクを実現する。また，システム全体にとっての必要性や協力者の関心に応じて，一部のテキストを選んでその内容情報をさらに処理・マーキングする。

コンテンツ・マークアップは，まず『旧唐書』『新唐書』から開始することになっている。この二書を選んだのは，いずれも唐滅亡後150年以内に編纂され，唐文化の最も広範かつ詳細な記述となっているからである。マークアップの対象としては，テキストの内的構造のほかに，年号，人名，地名，それに作品・著作の標題などを計画している。

他のテキストについては，時間的余裕や必要に応じて対応する。比較的少ない労力で多くの情報を与える（費用対効果が高い）マークアップは，優先順位を付け，それに従って進めることになる。その意味で特に興味深いものとして，当時の知識を効率よく体系的に選録した「類書」と呼ばれるジャンルのテキストがある。

なお，ナリッジベース開発の全体計画であるが，まず最初の２年間であらゆる専門分野で使える幅広い情報を提供できるようにしたうえで，次の３年間で宗教学，制度史，科学史，美術史，文学史の専門的内容を開発することを予定している。

メタ・データ

上述のように，ナリッジベース内のメタ・データはすべてトピック・マップにより管理される。トピック・マップのもととなる情報モデルはきわめて単純かつ抽象的なもので，任意の情報とその関係を記述することができる。柔軟で幅広い応用が可能なため，メタ・データ管理にはきわめて適したものだ。メタ・データの収集と処理は次のように行われる。

・メタ・データは可能なかぎり既存の電子化第１次文献『旧唐書』『新唐書』，文献目録，伝記に現れる固有名詞から収集する。
・こうして得られた情報を処理・高度化する。

２つの史書から採った地名に，史書の地誌の情報を重ね合せることで，位置や上位の行政機関，名称変更，規模変更に関する情報を提供することが可能になる。また人名も同様に，伝記に記載さ

れる重要な事実とリンクされる。

・他の資料（紙，電子データ）から得られた近代以前や現代の情報も，トピック・マップに適宜組み入れる。
・可能なかぎり地理的位置情報を付加する（地図からのアクセスや地名による検索を実現するため）
・年号，日付や関連情報をマーク・正規化する（対照年表や日付による検索を実現するため）
・トピック・マップの情報はいずれの段階でも利用できる。これによりテキストに含まれる他の情報を探したり，その情報をトピック・マップにフィードバックすることができる。

ナリッジベース・管理システム

ナリッジベースのシステムは，既存のコンポーネントやソフトウェア・アプリケーションをもとに開発される。このため，それらをつなぎ合わせて一貫した使いやすいインターフェースを実現することが主な仕事となる。コア部分の Cocoon はアパッチ・ソフトウェア・ファウンデーションが維持管理を行うオープンソース型ウェブ・アプリケーション開発環境である。テキストのデータは eXist などの DBXML 対応データベースに格納されている。トピック・マップ・エンジンおよび編集のフレームワークとしては，今のところオントピア社の Knowledge Suite を予定している。

第 1 段階では，XML テキスト・ベースに対する Cocoon ベースの検索インターフェースを開発するが，これはシステムのその後の開発への足掛かりとなるものである。次の段階では，トピック・マップ・エンジンおよびトピック・マップの情報をテキストと統合する。

最も重要な部分は，利用者がナリッジベースによりデータを検索したり，情報をシステムにフィードバックするためのユーザー・インターフェースである。これによって，ナリッジベース・管理システムは利用者のニーズに対応したものとなり，その内容を進化させることが可能になるからだ。

期待される効果

唐代研究ナリッジベースは，情報アクセスの新しい方法を実現するものだ。単なる文字列検索だけでなく，知識空間の中をなめらかに航行することが可能になる。ちょうど遺跡を発掘する仕事に似て，ナリッジベースは情報の大海の中から新しい発見をもたらしてくれるのである。以下に利用の例をいくつか挙げるが，利用の可能性はそれだけに尽きるものではないことをあらかじめお断りしておきたい。

まず上記の 2 つの史書に現れる人々がどのように分布していたかを中国の地図の上に示すことができる。しかも，一定の地位を得た者や官吏となった者などを異なる色で表示することが可能である。また，この同じ情報を使ってデータベース（あるいはその一部，たとえば『全唐文』）に対する検索範囲を狭めることもできる（たとえば北方出身の人だけを検索し，その結果を南方出身者の検索結果と比較するなど）。比較の目的のために検索範囲を詩の行末の文字だけに絞り込むこともじつに簡単である。メタ・データの重ね合せを行うことで，テキストの中から当面の研究テーマにとって必要な部分だけを的確に抽出することが可能となり，これまでのように何千件もの検索結果の中から苦労して適当なものを探し求める手間が省けることになるわけだ。

また，高級官吏の社会的背景（出身門閥の数やその根拠地など）を比べるのも興味深いことだろう。さらに，ある時点である場所（長安の仏牙巡行など）にいたすべての人を特定し，そのことが各自の著作にどのように反映しているかを明らかにすることもできる。

そのほかにも多くの問いが可能となり，その答えがまたさらに新たな問いを生み出していくだろう。「唐代研究ナリッジベース」のコンセプトは，今後長きにわたって内外の研究者に刺激と指針を与えるものと期待される。中国史の中でもとりわけ興味深い唐代についての理解は，それにより確実に深まっていくはずである。

漢字情報学事始め

武田時昌

2000年４月に東洋学文献センターを改組した漢字情報研究センターが発足し，私は新設の情報部門教授に着任したが，それ以来，センターに関する業務として主に二つの仕事に従事してきた。一つは「漢字情報基礎論の試み」共同研究会の主宰であり，一つはセンター広報誌「漢字と情報」の編集である。いずれもこれまで行ってきた研究教育活動とはまったく異質な体験であり，人文学に押し寄せているIT技術革新の潮流を強く感じさせるものだった。

かつて電子工学の学生だった私がコンピュータ・プログラミングを習ったのは30年も前のことである。長尾眞前総長の授業では，パスカルの三角形をプリントアウトするプログラム作成のレポートがあったし，アセンブラ言語の実習ゼミでは宇治キャンパスまで学内バスに乗って出かけた。卒論では，核融合炉内のプラズマ状態にあるイオンサイクロトロン運動をコンピュータ・シミュレーションし，実験結果の数理解析を行った。その頃は，カードにパンチで穴を開け，札束のように何枚も重ね合わせて読み取り機にかける方式であり，朝早く大型計算機センターに出向き，翌日に出力されてくるのをのんびりと待つという悠長なものであった。

文学部に移籍してからは，中国学特有の文献読解法の特訓によって，大脳の工学的思考はすっかり改造され，ポケコンで処理すれば簡単に算出できる割円術の円周率計算ですら，桁数の多さに四苦八苦しながら筆算でやるほうが楽しいと感じるくらいのレトロな人間に変わってしまった。1985年にPC-9801VM2を購入してからはパソコンユーザーに仲間入りしたわけであるが，用途のほとんどはワープロであり，多種多様な他のソフトは集めるものの，実用に供しない単なるコレクターでしかなかった。

そんな素人同然の私が，漢字情報センターの共同研究の班長役を買って出たのは，新センターに赴任してきた情報学の研究者が互いの研究を知り，中国学研究者とともに有意義に語り合える場を提供しようという情報部門教授としての義務感からにちがいない。しかし，それだけではなく，コンピュータにいったい何ができるのか，あるいは何をやらせようか，という好奇心と期待感があったことも確かである。以前はコンピュータ解析を標榜した文系論文を見かけると，ついつい理系出身者の猜疑心あふれる眼で冷ややかに眺めてしまっていた。数式とアルファベットで記述したプログラムに，漢字の持つ複合的意味や有機的構造が十分に表現できるようには思えなかったのである。しかし，この20数年の間にその難題を見事にクリアーし，漢字が自在に扱えるようになった。さらに様々な情報が潮遊するインターネットの仮想空間が生み出されるに至って，事態は一変したと言っていいだろう。中国学と情報学の融合というのは，研究会を立ち上げた時には夢物語に聞こえたかもしれないが，今や21世紀COEプロジェクトにふさわしい緊急の課題であると認知されるまでになったのである。

共同研究会で行われた論題は，外字・異体字等を含む文字処理法，漢字グリフの自動生成やパブリック・ネットワーク・フォント，マークアップ手法を用いたデータベース化，n-gramモデルと統計解析法による中国古代文献の分析等々，多岐にわたっており，しかも斬新かつ重要な問題ばかりである。その研究発表は，コンピュータの専門知識がないと直ちに理解しがたいこともある。し

かしながら，発表者にはその論点を一般向けに要約して「漢字と情報」に投稿してもらっているので，その編集時に誰よりも早く読んで，なるほどそういうことだったのかと得心することがしばしばである。だから，班長や編集長というと聞こえがいいが，漢字情報学の最前線の講義が学べる最も恵まれた受講生といったほうがふさわしいかもしれない。

ところで，昨年度に，文献類目の遡及入力のアルバイトとして雇った院生を集めて，研究会で知り得た最新情報を紹介しながら，データベース作成に関する簡単な説明会を行ったことがある。すると，知っていれば造作のないことに，とんでもなく遠回りしていたことがわかったらしく，思った以上に好評だった。電子テキストやオンライン検索等のコンテンツは，急激な増加を示しているが，離散的で雑然と浮遊しているだけであるから，有益な情報に容易にたどり着けるわけではない。研究会仲間との間の雑談が，最も有益な情報の入手経路である，というのが本当のところである。

平成 15 年度漢籍担当者講習会

そのことは，漢字情報学の諸技法を身につけなければ，これからの研究者は立ち後れてしまうことを示唆しているにちがいない。しかも，現在の大学でそれらを教わる場所も，機会も，きちんと制度化されて存在するわけではない。センターでは，図書館職員を対象として漢籍目録作成や電算処理のやり方を学ぶ漢籍講習会を長年にわたって開催してきたが，研究者向けの教育を行う必要性が高まってきているのである。

もちろんそれは，中国学に特有の文献考証の手法が時代遅れになったことを意味しない。教育の必要性を言うなら，むしろ逆である。コピー文化の普及によって，本屋に通い詰めたり，書庫をうろうろしたりする学生は，ほとんど見かけなくなった。文献学，書誌学的なモノに対する興味が薄れる一方で，電子化したイメージ資料が増えてくると，これまでのように自ずと習得できるものが，特別の訓練を受けないと身につかなくなる。喪失された知恵は復原するのは実に困難であるから，その方面の教育も忘れるべきではない。

そこで，21世紀 COE プロジェクトの教育部門の試みとして，従来の文献資料学をも包含した広義の「漢字情報学」の講習会を随時開催し，文献情報と漢字情報を横断的に扱えるハイパーな人材の養成を目指したい。手始めとしては，来春から自主ゼミ的なものを立ち上げ，さらに夏休み期間を利用して，若手研究者向けのサマーセミナーを開催する予定である。

現在のところ，具体的な教科内容は思案中であるが，テキストデータベースから階層構造を持つ高次なデータベースを構築するためのマークアップ技法や n-gram によるテキスト解析は，各自の研究テーマを材料として実習してもらうつもりである。もちろん，かくいう私も，若手研究者の受講生に混じって，しっかりと勉強し，研究のグレードアップを図りたいと考えている次第である。

TOPICS

平成15年度開催（予定）の国際シンポジウム・国際ワークショップ等

・国際ワークショップ「書体・組版ワークショップ（Glyph and Typesetting Workshop）」（終了）
平成15年11月28日・29日　於京都国際交流会館
・ 国際会議「ヨーロッパ日本共同考古学会議2003　アジアを掘る（Europe-Japan Colloquium 2003）」（終了）（イタリア国立東方学研究所・フランス極東学院京都支部・ミホミュージアムと共催）
平成15年12月11日・12日　於関西日仏学館およびミホミュージアム
・国際シンポジウム「漢字文化の今」
平成16年2月8日　於京都大学時計台記念館国際交流ホール
・国際ワークショップ「中国語ナリッジベースの可能性（The Possibilities of Chinese Knowledge Base）」
平成16年2月20日・21日　於人文科学研究所大会議室

Chinese Characters
and Culture

漢字と文化

発行日　2003年12月24日
発行者　文部科学省21世紀 COE プログラム
「東アジアにおける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―」
住　所　〒606-8265　京都市左京区北白川東小倉町47　京都大学人文科学研究所
電　話　075-753-6997　　FAX　075-753-6999
Web ページ　http://coe21.zinbun.kyoto-u.ac.jp/