

漢字情報研究センター



北白川の静かな住宅街の中に、「漢字情報研究センター」という研究所があります。共同研究「漢字情報基礎論の試み」の班長をされていた武田時昌教授に、漢字情報研究センターについて、また、ここで行われている研究についてお話を伺いました。



— ここには初めて来たんですけど、すごい建物ですね。

そう、なかなかモダンでいいでしょ。中国学のメッカではあるけども、建物を作るときに「何も中国っぽいものにする必要はない。スペインのロマネスク様式にしようぜ」と、すごい遊び心ですよ。最初は違和感があったかもしれないけど、何も形式にこだわる必要はないんです。

— では、漢字情報研究センターの概要を教えてください。

ここは、もともとは義和団事件の賠償金を使って作られた東方文化研究所というところだったんです。それが戦後京都大学の中に入って、人文科学研究所の一部になった。もとは大学とは別だったんですよ。昔は、有名な先生たちがほんとにすぐそばに住んでいて、昼と無く夜と無く集っては研究をしていた。違う時間が流れていたんです。

人文科学研究所の中では、この建物は2000年まで東洋学文献センターという、漢籍を一般の人々に利用してもらおう場所でした。ところがパソコンが普及してインターネットを活用するようになって、これまでの人文科学の分野でも大きな変化が起こった。そこで、従来の文献資料学を漢字情報学へとグレードアップさせようと漢字情報研究センターに改組され、コンピュータで漢字の処理をする情報学の専門家もスタッフとして新しく加わりました。その人達の研究と、中国学の研究内容を融合させるため、「漢字情報基礎論の試み」という共同研究を始めました。お互い何がやりたいのか、何ができるのか自由に話し合う、そんな場を作ったんですね。といってもこれまでの文献考証学の方法論をすっかり捨ててしまうわけではありません。伝統的な学問スタイルも大事にしつつハイテクも取り入れて、今と昔の融合ができればと思います。

— 実際には、どのような研究が行われているのですか？

コンピュータが出てきて最初に問題になったのは、漢字コードですね。コンピュータで漢字を扱う際は、一字一字に番号(コード)を振って処理するんですけど、外字や異体字なんかの、コードが振られていない文字が存在するという問題。それから、国・地域によってまったく違う漢字を使っていて、それぞれ違うコード体系を持っているという問題。でも、今はあっという間に時代が進んで、その心配はなくなりましたね。7万字以上の漢字を自在に扱えるし、国際統一規格もできて、昔のように外字を個別に作るなどという必要はなくなりました。これは大きな進歩だと思います。

コンピュータってのはやっぱり、アルファベットを扱うようにできてるんですよ。表音文字の世界はすぐにコンピュータになじめるけど、漢字はそれぞれに形や意味も持ってますよね。これまでのコンピュータの技術では、失われるものが多々あった。けれども、工夫すれば漢字の形・音・義をすべて情報として保ったまま扱えるようになってくるでしょう。そこで立ち現れてくる大きな問題があります。近代以降の大きな流れだと思いますけど、コンピュータの世界はやはり、西洋型の思考を中心的に扱うようになってるんです。非西洋型の文化は、一度西洋的な思考様式に置き換えないと理解できなくなっているんじゃないでしょうか。それをいかにして、非西洋型の思考も十分に分析できるようにしていくか。中国にしる日本にしる、漢



▶ 書庫に並ぶ線装本

はみだし
すてーじ

これが載るころ私はいません(そのハズです)。さようなら。
⇒ 旅立つ人がいれば新たにやってくる人も。別れと出会いの季節です。

(法・5 ひっきー)
(新生を見かけるとウキウキする編)

字を用いることによって、思考様式が西洋とは全く異なっています。コンピュータの中に漢字を取り込んで、われわれが頭の中で考えるようにコンピュータにも考えてもらうにはどうしたらいいか、ということをしていろいろ工夫して考えていくと、漢文的な思考というのはどんなものであるのかより深く認識するきっかけになるかなと思います。

あと話し合ったのは、書物をどのような形でデータベース化したらよいかという問題です。テキストデータベースで、コンピュータを使えばすぐに検索ができる。それって感激だったんですよ。特に中国だと、典拠を踏まえた作品を書くんです。読んだ本の主要なものは記憶していて、その言葉を使って、昔の人々のアイデアを内包したような文章を書くわけですね。中国学は、その典拠調べを延々とやってるといってもいいくらいなんです。

それが、今はキーを叩くだけで一瞬で典拠が出てくる。この部屋いっぱい位の書物がこんな小さなディスクに入って、一字一句検索できるわけですよ。するとこれまで分からなかった典拠が分かると同時に、いかに昔の人々が同じ典拠を何度も繰り返し使って文章を書いているかということが分かるわけです。ネタがばれてしまうということですけど、ここが面白いところで、人間の考えることってのは昔も今もおそらくそんなに変わらないんですよ。今は、人と違うこと違うこと言おうとしてそれが個性であり独創だと考えられてますけど、中国の優れた人々の文章を見ると、人が言っていないことを言おうという気は全然ないんですよ。時代に合わせて表現は変わるかもしれないけど、語ろうとすること、伝えようとするのはそんなに変わらない。それはとても平凡で世俗的かもしれないけど、綿々と語り継いでいるものがあるという驚きですね。コンピュータを使って、そういうことが見えてくるのではないのでしょうか。

けれども、いろんなところに現れてくる概念について調べるのは大変な仕事です。よく使われている語句をインターネットで検索すると、沢山出てきすぎて目的の情報に辿り着けないじ

ゃないですか。人々の共通認識を作っているようなものについての情報はなかなか手に入れられないんですよ。だから、テキストデータベースではなく、もう少し有用な情報を含んだデータベースを作らねばならない。

それでmark-upという手法を用いて、文章にコメントをつける。たとえば「月」で検索すると、一月二月の「月」も天体の「月」も出てくるわけです。そこで、それを区別



▲センターの中庭から

はみだし
すてーじ

すりっぱ研究会は別にすりっぱを研究してるわけではないですわ。

(工・1 すりっぱ研究会社会経験部長)

できる情報を付け加えてやります。すると、情報がもう少し絞り込んだ形になる。ただ、それだけでは不十分で、メタ情報というか、目的の情報にたどり着くためのガイドブックみたいなものが必要です。いろんな情報を含んだデータがあって、上からマップ付けしてやれば情報がよりシャープな形で見えてくる。研究所ではその具体的なやり方を話し合っています。

— 書物のデータベース化を通して、実際の書物に対する考え方に変化があったのでは？

ご指摘の通りで、一番議論するところは、データベース化を通じて、失われるものがあるということです。書物を実際に手にとって見るというリアリティーは電子テキストでは味わえない。それに電子テキストだと、目的の情報にすぐ辿り着けるけど、遠回りしないからほかの情報には出会えない。それはやっぱり大きいと思います。我々がIT技術に興味を持っているのはここなんです。新しいコンピュータのスキルを教育したいとは思っていますけど、同時に、これまでであれば書庫に入り、和綴じの本を手当たり次第めくって出典調べするという作業を通じて自然と身に付いた書物に対するリアリティーは、今後教育プログラムとして教えないと身につかないんじゃないでしょうか。文献資料学を漢字情報学にすることによって見えてくる新たな風景があるわけですけど、そのときに失われるものがやはり大きな問題だと。そこを我々はじっくり考えて、連続性を保った形で学問を長く活性化させていかねばならない。

そんな話をしていると面白いのが、電子テキストを画面で見るとのなんか嫌だ、線装本じゃないと、と我々が言っていると、じゃあ本物とそっくり同じものを作ってしまえばいい、と言う人がいるんです。そのうち、仮想空間内で部屋に入っていって手に取れて、においや感触まで再現したようなものもできるかもしれない。そんなのができたら凄いですよね。やっぱりこういう人たちと話したら面白いなあと思うんです。

— 最後にまとめをお願いします。

やっぱりね、まずは知りたいと思わないといけな。今は情報としては何でも流れてるわけだから、何を知りたいか考えなければいけない。絶えず選択を求められることになると思います。そのときは必ずしも今だけ見るのではなくて、古代人の知恵を借りればいいし、そのツールはあります。今後どうなるでしょうね。コンピュータで、これまで分析できなかった人々の思考を分析できるようになると同時に、コンピュータとともに、昔の優れた人々と同じような発想でものを考えたり本を書いたり、そういう時代が来てほしいですね。

いずれにせよ、私は書物を楽しんで読む読書人になりたいし、みんなにも読書を楽しんでほしい。漢字情報学を推進しているのも、漢文を読む楽しさを語り継ぎたいというのが原点なんです。それが最終目的と思っていただけたら。

— ありがとうございます。 (とうめいRunner)