

特集

失われゆく情報の復元・保存技術

—人文科学における情報処理—

(文献学・データベース共有・史料編纂)

山田 奨治

国際日本文化研究センター

shoji@nichibun.ac.jp

近年、理系と文系の学問の融合が随所で求められるようになってきた。その理由の1つとして、理系の学問には文系的な視点からの問題の発掘が要請されているのだろう。情報処理学会の中では、情報科学と人文科学の境界領域を扱う研究会として「人文科学とコンピュータ研究会」(SIGCH)が1989年から活動を続けている。そのSIGCHでここ数年、ホットなテーマになってきているのが、デジタル・アーカイブをめぐる技術的・社会的な諸問題についての検討である。

デジタル・アーカイブで重要なことは、単に最新の技術を使って文化的な資料をデジタル化し、保存・活用することではない。アーカイブとは、文化的な資料の収集・整理・評価・利用・保存の一連の流れをすべて含む概念で、デジタル・アーカイブはこれらのすべてに関係する。アーカイブに対応する訳語がないことに象徴されるように、日本にはアーカイブという概念もなければ、アーキビストという職種もなかった。日本のデジタル・アーカイブは、土台のないところに楼閣を築くようなことをしているのだともいえる。

しかし、土台がないが故に固定観念に囚われず、文化的な資料のデジタル情報を使ってさまざまな試みができる可能性もある。今月号から2回に渡って掲載する「失われゆく情報の復元・保存技術—人文科学における情報処理—」特集では、何もしないと時間の経過とともに失われてゆく文化的な情報をいかに保存するか、あるいはすでに失われてしまった情報をいかに復元するかについての、いくつかのプロジェクトを取り上げ、それぞれの担当者に執筆いただいた。今月号では「文献学・データベース共有・史料編纂」を、来月号では「博物館・美術館・遺跡・埋蔵品」をテーマにする。

本特集で紹介するプロジェクトは、単なる情報の復元・保存ということにとどまらず、情報科学によって従来の人文科学の枠組みを大きく変える可能性を秘めたものばかりである。そういった意味では、情報科学を媒介にして理系と文系が融合する学際領域の最先端をご紹介するわけである。

はじめに取り上げた、九州大学の竹田正幸氏らによる和歌の研究は、商業ベースで作成された膨大な和歌データベースに対してデータ・マイニングを施すこと

によって、歴史の中で失われてしまった情報を発見しようとするものである。竹田氏らの方法は、和歌研究に革新的な変化をもたらす得るもので、SIGCHでも高い評価を受けている。

従来、全文データベースは、探したい情報のたいへん効率的な検索手段としてしか利用されてこなかったが、デジタル化されたデータを用いれば、あたらしい発見につながる可能性がある。東京工業大学の中川正宣氏による新約聖書の研究もまた、聖書の成立についての失われた情報を復元しようとする、たいへん意欲的な試みである。文化的な資料の全文データが整備されていくに従って、こういった研究は今後ますます増えてくるものと思われる。

さて、古文書を対象にして全文データベースを構築しようとする、難解な古文書の文字を判読する必要がある。日本語手書き文字認識の技術を使って、その作業を支援できないかという研究が進行している。著者はその研究グループの一員であるので、大阪市立大学の柴山守氏とともに、手前味噌ながら研究の現状を紹介した解説を掲載させていただいた。

いったんデジタル化されたデータは、それを作成した機関で保存されるだけでなく、多くの研究者に共有され、活用されるのが理想である。東京大学史料編纂所や国文学研究史料館が進めているデータベース共有の取組みは、先導的な役割を持つものとして大きな期待が寄せられている。これらのプロジェクトの概要について、横山伊徳氏と原正一郎氏にそれぞれ執筆いただいた。

以下は、来月号の予告である。文化財の3次元形状や高精度の画像情報を取得しておく、後世の修復に役立つだけでなく、それをもとに作成された映像を用いると、人文科学の研究成果をたいへん分かりやすく、効果的にプレゼンテーションすることができる。凸版印刷(株)の加茂竜一氏と国立歴史民俗博物館の安達文夫氏らに、いくつかの事例を紹介していただく。

人文諸科学の中でも、特に考古学では情報機器をふんだんに活用した研究手法が芽生えつつあり、失われゆく情報の復元・保存技術という点では、実に興味深い研究が進んでいる。大阪電気通信大学の小澤一雅氏と財団法人元興寺文化財研究所の塚本敏夫氏に、最新の成果について解説していただく予定である。

読者のみなさまには、本特集を通じて、情報科学と人文科学が交わるところに広がっている可能性を感じ取っていただければ幸いである。

(平成14年7月25日)