

情報学と文学の融合を考える

人間文化研究機構国文学研究資料館 名誉教授 安永 尚志



私と科研費について、まず思い浮かべるのは、1976年から79年に実施された文部省科研費による特定研究「情報システムの形成過程と学術情報の組織化」である。故猪瀬博先生代表で、全体で10億円を超える経費で、500人を越える研究者が参加した。その総括班に入れていただき、大学間コンピュータネットワークの形成に参画させていただいた。インターネットはまだ無く、Inter-University Computer Networkとしての先駆けの実用化研究である。この科研費は、文部省の重点施策とされ絶大な支援を得た。1973年から75年の特定研究「広域大量情報の高次処理」(故島内武彦先生代表)の研究成果を引き継ぎ、日本の情報学研究を先導し、推進し、多大な貢献を果たした。

研究生活の前半は、大学で主に情報通信工学の研究で過した。後半は大学共同利用機関の国文学研究資料館に異動した。大学での研究は、教育を伴って行われる自由な個人研究が主である。大学共同利用機関でのそれは、教育を持たない目的指向の業務研究である。ただし、最近では大学共同利用機関は総合研究大学院大学に加入し、高等教育の一端を担っている。学術振興会のホームページで、渡辺晃宏先生(奈良文化財研究所)のご指摘に頷いている。とくに、人文科学系の研究機関においては、業務研究しかなく、個人研究はないと言うものである。

国文研では、コンピュータのお守りが仕事であった。これに伴う研究課題は多種多様で、かつ情報学分野での人文学の応用研究は未開拓の領域であった。例えば、古文を含む日本語処理1つとっても、有効な入力手段もなかった(JIS日本語文字規格が制定された頃)。業務研究は業務費という校費であるが、その目的外使用は許されないし、予算も少額である。コンピュータのお守りには外部資金が不可欠だった。科研費は、業務研究の枠をはみ出での申請となった。研究を推進するためには重要な経費であり、獲得に真剣に努力した。お陰様でずいぶんと世話になって、国文学研究におけるコンピュータ活用も端緒が開けた。感謝しても仕切れない。しかし、これらの科研費は機関として突出しており、他の申請に圧迫となった。コンピュータは金食い虫だという批判もさんざん聞かされた。人文学には高額の研究費は要らないとされるが、情報の組織化には膨大な継続的な資金が必要である。

とくに、データベース作りである。初期の学術用データベースは大型計算機センターを中心に、海外からの導入で対処しており、我が国独自の形成は進んでいなかった。1985年に学術調査官に任じられ、学術情報の組織化や学術情報センター(現国立情報学研究所)の創設、研究成果公開促進費(データベース)の充実化の施策に携わった。大学共同利用機関などが有する独自の専門的情報資源のデータベース化を進める必要があり、なかでも日本文化情報資源の情報発信が強求められていた。

(1) 日本文学研究資源の組織化

1983年、国文学研究資料館への転任に伴い、国文学研究に関わる様々な資料、情報の構造分析を行い、情報学からの組織化の基本的理念をまとめた。データベース、一般研究、試験研究、国際学術研究などの科研費を得て、日本語(古文を含む)によるデータベースの構築が可能となった。開発研究した主なデータベースは、古典籍の所蔵目録、総合目録、研究論文目録、全文(日本古典文学大系、断本大系等)、原

本画像など多岐にわたる。とくに、1987年のマイクロ資料目録データベース公開は、日本語による日本文化資源の最初の事例であり、そのCD-ROM化も含め、国内外の利用者から好評を得た。

1988年から5ヶ年で実施したデータベース、試験研究、特定領域研究などの科研費による日本古典文学本文データベース研究は、岩波書店旧版「日本古典文学大系」全100巻約600作品などの全文データベースで、1998年の試験公開以来、海外から高い評価を得ている。そのデータ記述文法(KOKINルール)は最初期の実用化研究成果であり、SGML(Standard Generalized Markup Language)やTEI(Text Encoding Initiative)に先駆け、かつそれら同等の記述能力が実証された。

文科系研究資料情報は一般に大量多種多様で、情報表現も文字、数値、画像、音声など複雑に絡みあった構造を持つ。研究過程で生成される学術資料情報も同様で、これらの総合的な利活用によって教育研究が行われる。基盤研究や重点領域研究による資料情報のあり方と情報構造の把握、それに基づく組織化を研究した。データベースを活用して、国文学研究に新たな知見を得ることができるとして研究を進めた。電子本「漱石と倫敦を作る」は、その一例のシミュレーションモデルである。

(2) 国際コラボレーション

2001年度から5ヶ年に渡って実施した基盤研究(S)「国際コラボレーションによる日本文学研究資料情報の組織化と発信」は、中間評価、最終評価において高い評価を得た(A+)。研究分担者の原正一郎助教授(現京大教授)の貢献が大きい。海外の大学、研究機関、学会などとの国際コラボレーションにより、日本学学術資料情報(研究ディレクトリ、研究論文、全文資料など)を集積し、インターネットにより発信している。とりわけ、イタリア、フランスでの日本学に関わる学術情報をほゞ網羅し組織化し、国際的な活用度も高い。その資源共有化は総研大での教育研究にとっても不可欠な要件となっている。

2004年度より推進した特定研究や共同研究「文化情報資源の共有化システムに関する研究」は、柴山守教授(京大)、原教授など30人を超える各機関の研究者の協力により、人文科学におけるデータベースを統一的、横断的に検索、利用する仕組みを研究し、実用化を目指したものである。この研究では、プロジェクトに参加している8つの人文系研究機関が持つ30個ほどのデータベースを実際に接続し、横断検索するシステムを構築し、実証実験を通じて実用性を確認した。現在では、人間文化研究機構の資源共有化事業として、これらの研究成果を踏まえた総合的な基幹的事業として推進されている。

〈閑話休題：猪瀬先生と俳句〉

国文研に移ってハタと困った。異次元世界に飛び込んでしまったような違和感の連続で、数年戸惑っていた。文学と情報の融合と言うと格好は良いが、そう簡単ではない。よく先生に相談に行った。あるとき、ジョークだと言われたが、「コンピュータで俳句でも捻って来い」というものだ。昨今、将棋やチェスのゲームでは、コンピュータが人間に勝つまでになってきたが、俳句は未だに捻れない。これは実に難題で今では諦めている。