

連載報告

情報学とその課題



3. わが国における情報学の課題†

尾 関 雅 則¹⁾ 田 畑 孝 一²⁾ 根 岸 正 光³⁾
 藤 原 讓⁴⁾ 神 田 利 彦⁵⁾ 横 井 俊 夫⁶⁾

1. はじめに

米国の大学図書館へ行くと、ならんでいるコンピュータ端末がクレジットカードを切って情報検索をしている学生で賑わっている。料金を払ってでも情報を得ようとする気概がひしひし伝わってくる。そしていくつもの情報提供者が利用者を獲得すべく情報の質と量で互いに競いあっている。

ひるがえって、わが国の場合、そのような学生の光景は想い浮かばない。無料だとしても賑わいをみせるであろうか。検索に足るものがない、有料に値するものがないなど、情報提供者側のせいにするものもいよう。情報を大事にする慣習が根付いていないわが国の利用者側の状況がそのまま提供者側に反映してか、わが国の情報資源、特に日本語の情報資源の整備、流通はまだまだこれからという状況にある。有用なものが整備されれば、おのずから利用が喚起されるとの期待も含めてここでの議論を進める。

2. 情報資源の整備

(1) 情報資源の基本的認識

情報は、社会経済活動を支える基本的要素の1つである。近年、コンピュータと情報通信技術の飛躍的な発展は、従来の数値や文字情報のみならず、音声や画像情報を含む大量かつ高度な情報処理や情報流通ネットワークを実現した。

このような情報技術の高度化は、「情報」を体系的に収集し、加工し、処理しやすいように整備・蓄積することを容易とし、利用価値の高い「資源」と呼ぶにふさわしい状態にすることができるようになった。

「情報」資源は、「物質」、「エネルギー」などの天然資源と異なり、研究開発の進展や技術進歩、社会システムの高度化・多様化などにより、その拡大再生産と新たな蓄積を図ることができることに大きな特徴がある。

(2) わが国における情報資源整備の現状
 情報技術の進歩により、現在様々な情報が様々なメディアにより容易に資源化できる環境にある。このような情報資源は、科学技術分野はもちろんのこと、産業・経済、文化・芸術、教育、観光・レジャー、生活関連分野など幅広い分野で形成されつつある。

これら情報資源の充実・整備が人間や社会に及ぼすインパクトは、(a)人間の知識や研究開発・技術開発に及ぼす影響、(b)社会システムの情報資源への依存性の増加、などに代表される。

研究・技術開発の成果である科学技術情報資源は、新たな研究開発の糧であると同時に人類共通の知的財産でもある。一方、医療・福祉、防災、環境、行政、教育などの各種社会システムの有効性は、その投入する予算規模や人材と同様、システムの保有する情報資源の質と量にも依存するといわれ、今後この比重はますます大きくなると予想される。

わが国の科学技術分野のデータベース化は、米国より約10年遅れ1970年代後半より本格化し、現在、科学技術以外の分野のデータベースの整備も急速に進んでいる。わが国で利用可能な国産デ

† Information Science: What has to be done? (3)—In Case of Japan—by Masanori OZEKI (a member of 15th term, Science Council of Japan and also Railway Technical Research Institute), Koichi TABATA (University of Library and Information Science), Masamitsu NEGISHI (National Center for Science Information Systems), Yuzuru FUJIWARA (Institute of Electronics and Information Science, University of Tsukuba), Toshihiko KANDA (The Japan Information Center for Science and Technology) and Toshio YOKOI (Japan Electronic Dictionary Research Institute Ltd.).

1) 第15期日本学術会議会員(第4部)、(財)鉄道総合技術研究所、2) 図書館情報大学、3) 学術情報センター、4) 筑波大学、5) 日本科学技術情報センター、6) 日本電子化辞書研究所

データベースは、平成3年度のデータベース台帳総覧によると、文献データベース152(科学技術分野56, その他の分野96), 全文データベース342(科学技術分野33, その他の分野309), ファクトデータベース345(科学技術分野27, その他の分野318)となっている。しかし、国内で利用できる海外製を含めた全利用可能データベースのうち、国産製は、文献で21%(科学技術分野15%), 全文で25%(科学技術分野12%)にすぎない。ファクトデータベースについては、総数では国産製が多いが、科学技術分野では24%となっている。

全文データベースでは、新聞編集工程での副産物としての各種新聞情報が、ファクトデータベースでは、市況、投資、財務などビジネス系を中心に民間企業でコストリカバリーベースで整備されつつあるが、採算性の悪い基盤的情報資源としての科学技術情報の整備はまだ不十分といわざるをえない。

(3) わが国における情報資源整備の課題

個々に存在するだけでは利用価値の低い情報を、利用価値を高め、利用に便ならしめるように体系的に収集、加工、整理・蓄積するためには、長年にわたる多くの知的労働と資金が必要となる。

オリジナルな情報が、ほぼそのままの形で資源化される数値情報や事実情報、一部の全文データベースはあるものの、情報そのものを圧縮化したり、利用しやすい形態に加工する機会が多く、オリジナルな情報を資源化するための新たな、ある意味では非生産的な、作業が必要となる。こうした作業は、人的・知的作業が多く、合理化・機械化が困難あるいは部分的にしか適用できない。

データベース構築上の問題点として、データベース白書'93版のアンケート結果によると、(a)データの収集、入力などの構築作業にコストがかかる(87.5%)、(b)構築後のメンテナンスコストが負担(61.3%)、(c)初期投資が大きく回収困難(45.0%)、(d)構築に関しての国の助成がない(22.5%)、(e)DBMSなど効率的ソフトの不足(22.5%)、(f)標準化の検討が不足(17.5%)、(g)インデксаなどのデータ作成者の不足(13.8%)などとなっている。

オリジナルな情報を資源化する作業は人的・知

的作業が多く合理化・機械化が困難で、事実、アンケート結果もそれを裏付けている。このような状況を克服していくためには、(a)研究者や研究機関など情報発信者への、あるいは防災、環境、医療機関など各種情報収集機関への情報資源化のための財政的、人的、技術的な支援の強化、(b)大学や国立試験研究機関などが積極的に研究情報を公開できる環境の整備、(c)情報資源化のための情報表現形式の標準化の推進、(d)わが国の主体性の確保と国際的役割を果たすために基盤となる情報を整備する公共的な情報資源化専門機関に対する助成の充実、などが考えられる。また、技術的な課題として、情報そのものの自己組織化によるさらに高度な情報の創出メカニズムの研究なども検討に値しよう。

3. 日本語情報の収集と提供

日本はその情報資源のほとんどを欧米のそれに頼っているのが現状である。情報資源の整備は今後、特に日本語のそれに努める必要がある。そのため公的機関(あるいはそれに準ずる機関)における日本語情報の収集と提供活動の現状を概観し、今後どうあるべきかについて触れることにする。現状は、満足すべき状態とはいいがたく、今後の努力に期待することが大きい。ここでいう日本語情報とは日本語の言語情報のことである。

(1) 日本語情報への需要

日本語情報への需要の源は、大きくは日本語教育、言語学教育、コンピュータによる日本語処理の3つである。従来は、日本語教育も国内の国語教育、言語学研究も国語学研究という観点からの需要であり、コンピュータに関してもごく限られたものであった。しかし、ここ十年程で大きく状況が変わってきている。国際化と情報化がその要因である。日本経済の国際化の中で外国人に対する日本語教育の大きな需要が生じた。しかも、ビジネスの現場や日常生活に役立つ実用日本語に対する需要である。コンピュータによる日本語処理は日本語ワープロの普及によって大きく加速された。機械翻訳をはじめとする日本語処理ソフトウェアの本格的な取組みが始まり大規模で高品質の日本語データが求められるようになった。日本語教育と日本語処理という現場のニーズに刺激され、また、学問自体も大きく進展する中で言語学

研究の現場も大きく様変わりをしつつある。理論研究がある種の成熟に達しつつある中で、日本語学という観点を踏まえた実データに基づく研究が注目を集めるようになってきている。

(2) 日本語情報の供給

需要サイドの変化に十分に対応しきれているとはいえないが、供給サイド、すなわち、日本語情報の収集と提供の活動も新たな様相を見せつつある。ここでは、最近の動向の特徴的なところにしてその活動を概観する。

外国人の日本語教育に関しては、日本語情報は教材、教授法などやそれらの裏付けデータの形をとる。教育課程の研究を含め体系的な整備の役割を担うところとして国立国語研究所がある。たとえば、外国人の日本語語彙教育のための基本語用例データベースの作成などが行われている。一般的な教材に関しては外国人に対する日本語教育の中心機関である国際交流基金が供給源である。また、教育現場の具体的なデータは外国人留学生の日本語教育のセンター機能を持つ各地の大学からも提供されている。

コンピュータの言語処理に関しては、日本語情報はコンピュータ処理可能になった辞書、コーパス (corpus)、テキストデータ、そして処理ソフトウェアなどの形をとる。日本電子化辞書研究所では日本語と英語を対象に次世代自然言語処理技術のための大規模な電子化辞書の開発が行われている。現在、50カ所ほどの国内外の大学、研究機関などに評価研究用として供給されている。また、小規模ながら特色のある辞書作りが各所で行われている。情報処理振興事業協会技術センターが代表例である。また、京都大学をはじめとする大学の研究グループがデータやソフトウェアを公開し、共同利用の環境を整えつつある。テキストデータに関しては各種の機関がデータの蓄積をし、一般利用に向け供給を始めている。さらに、音声言語に関しては、国際電気通信基礎技術研究所、電子技術総合研究所、日本音響学会、東北大学などで音声データの収集と提供がなされている。

言語学研究に関する日本語情報の供給は国立国語研究所や大学などにおける研究成果の従来からの蓄積活動の延長線にある。むしろ、これに関しては出版社における大型辞書などの刊行事業に負

うところが大きい。

(3) これからの活動

日本経済の不調から外国人の日本語教育の過熱状態は納まったが、その重要性はむしろこれからであろう。分野や外国語ごとの対応での辞書や教材、教授法などの開発、整備を着実に進めねばならない。情報化もこれから本格化する。日本の国家情報基盤のその基盤となるのが日本語情報である。日本語処理用言語データの開発、整備にもいっそうの努力が求められる。これには言語学研究との密な連帯が重要になる。言語学にとっても言語データの分析に基礎を置く実験科学への展開が今後の課題である。しかも、どの方面からも海外からの日本語情報への需要が生まれてきている。これには、日本として無条件に応えるべきであろう。

4. 情報流通の整備

流通すべき情報として、灰色文献 (Grey Literature) と呼ばれるカテゴリの文献がクローズアップされつつある。灰色文献とは通常の出版・流通の経路で扱われていなく、またそれについての検索手段が整備されていないので、入手が困難な文献を指す。これにはテクニカルレポート、会議資料、学位論文、大学出版物、官公庁資料などが含まれる。正規の流通経路にある学術雑誌は研究成果が得られてから出版するまでに時間がかかり、現場の研究者にとってはテクニカルレポート、会議資料などがより重要となる。また官公庁資料は民間の企業活動にとって貴重であるがなかなか入手できない。灰色文献の収集と蓄積、検索、配布については、従来、英国の BL (British Library)、欧州の EAGLE (European Association for Grey Literature Exploitation)、米国の NTIS (National Technical Information Service)、わが国の JICST (日本科学技術情報センター) などがその事業の一貫としてそれに努めている。灰色文献の二次情報をオンライン情報検索システムや CD-ROM で提供し、ものによってはその一次資料のコピーを郵送あるいは FAX で配布している。しかしながら、対象とする文献は限られたもので、またその事業の採算性に難点があり、このままの体制では問題の解決にほど遠い。最近になって、問題の解決を求めた動きが活発

になってきた。1993年9月には米国NASAの主催によるワークショップ「外国灰色文献へのアクセスの改善」が開催され、また同年12月にはEAGLE主催(JICST, NTIS共催)による「第1回灰色文献国際会議」が開催されている。いずれの会議においても、国際学術情報ネットワークを介した全文データベース(一次資料)の相互共有、所在情報サービス、一般利用者への文献本文(一次資料)の提供の必要性が強調されている。つまり、最近注目されているインターネットにおけるWAISやWWWなどの形態が問題の直接的な解決の糸口となるというのである。「第1回灰色文献国際会議」では、INGLA(International Network for Grey Literature Association)の設立を目指して、会議参加者にアンケート調査を行っている。

あるべき1つの姿として、国際学術情報ネットワークを設置し、灰色文献の生産者となる組織が加入する。各組織は自身が生産する文献を格納する全文データベースを維持し、ネットワークからのアクセスに応じる。望むらくは、各国の産官学のすべてがそのような組織として加入する。このネットワークにはもちろん一般の利用者がアクセスでき、文献本文が電子的に入手できる。

この方式を進めるにあたって解決すべき点がいくつかある。

- (1) 文献データの国際標準化
- (2) マルチメディア情報
- (3) 所在情報、検索手法の開発と整備
- (4) 著作権使用料
- (5) 一般利用者からの料金徴収方法
- (6) データベース提供者への利益還元方法
- (7) ネットワーク運営・財政の管理

特に、わが国としては日本語の問題をかかえているので、欧米に対し積極的に働きかける必要がある。

5. 著作権の権利処理と情報の標準化

情報化とは、情報を技術的・経済的に処理可能な客体に化することであり、これは情報の電子化によって具現された事態であると解される。情報学は、このような現代的状況の要請に応えるべく構成され、またその振興が図られるべきものである。ところで、情報の発展のためには、情報学研

究の素材あるいは対象となるべき電子化情報が、情報学研究者に豊富に提供される必要がある。一方、情報学研究には、基礎科学的研究だけでなく、社会の具体的な要請に直接に応える応用科学的領域も大きいと目される。このような応用科学的研究の成果が社会的にひろく実地に応用されることにより、わが国における情報化が適切に助長されるものと期待される。したがって、情報学関連の制度、機構の検討にあたっては、情報学の研究促進を目的とするものと、情報学の成果を社会に還元、定着させることをめざすものという両側面からの検討が必要である。

(1) 情報学研究と著作権処理機構

情報の電子化にともなって、情報の処理・提供方法は飛躍的に多様化し、この傾向は、昨今多用されるマルチメディアという用語に象徴されるように、現在ますます強まっている。そして、この種の応用を研究的にまた事業的に推進しようとする立場から、著作権の処理が重大な制約要因と見なされるようになってきている。こうした事態は、これまで著作権の問題に直接関係しなかったような企業、研究者の参入、すなわち、彼らにおける著作権処理に関する無知、誤解に由来する部分が多いようである。一方、著作権者の側においても、マルチメディアなどの新しい公刊形態に対する誤解・不安という要因がある。これらは当局の啓発活動によって逐次解消されるべきであり、各般の措置がすでに講じられているところであろう。しかし、啓発のみによっては解決し難い部分もあり、法的あるいは制度的な方策も検討されつつある。

情報学の研究においては、研究の素材あるいは実験試料として、各種の著作物を多種・大量に利用することが不可欠である。現状では、何分新しい事態であるため、研究者側と著作権利権者側の双方に上記のような無知や誤解があり、個別的な利用交渉は一般にはかばかしくないようである。情報を保護する制度によって、情報学の進展が阻害されるとすれば、これは重大な矛盾であろう。

元来、著作権法には、たとえば図書の売買契約に際して、これに並行して、その内容の利用方法を個別に著作権者と契約交渉するという煩瑣な過程を省略するための、著作権利用約款の役割があると考えられる。このような著作権処理の定型

化、簡素化によって、従来型の著作物の流通は良好に維持されている。これに対して、いわゆるマルチメディア・ソフトをはじめ、電子化された著作物では、利用者側において、これらの各種を素材として組み合わせ、様々に複製・加工・編集して利用することが可能であるし、また電子化の利点はむしろこのことにあるともいえる。現行の法制はこうした利用態様に対応していないので、利用契約は個別的交渉にならざるを得ず、研究的利用に関しても、これが不調となることが多いのは上述のとおりである。この際、法律、制度の整備により、この種の利用契約に関する手順の明確化、標準化、簡素化を図ることが望まれる。

著作権審議会マルチメディア小委員会の第一次報告（平成5年11月）にもこうした指摘があり、その初段階として、著作権の所在をデータベース化した「著作権権利情報集中機構」の試案が提示されている。マルチメディア・ソフトの制作では、多数の著作物を素材として利用するために、それぞれの著作権者と交渉する必要があるが、現状ではそれらの著作権者を同定するだけでも、実務的に多大の労力を要する。著作権権利情報集中機構は、こうした権利所在情報を集中して、データベース化し、一括して提供できるようにしようとするもので、著作権処理の面からのマルチメディア振興策の手始めといえる。

著作物は著作権者の私的財産であると同時に、社会的な知的資産でもある。したがって、これが新しい情報技術によりいっそう有効に活用されることは、著作権者にとっても社会全体にとっても望ましいところであり、著作権保護の意義もここにあると考えられる。この際、情報学研究を助長するような著作権処理制度、体制の整備が望まれる。

(2) 情報試料の蓄積交換機構

前項にのべた著作権処理機構を、情報学研究振興の立場からさらに発展させると、これは、情報研究者間で共通に利用されるような情報試料を集中したセンターの設置という構想に至るであろう。全文データに関して、オックスフォード大学のテキスト・アーカイブズのような事例がすでに外国にあることから見ても、このような機構は情報学研究の振興に有効であると考えられる。

この種の機構では、研究者相互間での情報試料

の交換を簡便に行いような機能がまず想定される。しかし、より重要であるのは、研究者社会の外部、具体的には企業などから多様な情報試料の寄託を受け、これを研究者間の共用に付すという機能であろう。前記のように、研究者と企業など著作権者間の研究的データ利用契約交渉は難航することが多い。大学研究者は、もとより営利的利用を目的とするわけではないが、権利者側では前例のないこととして理解に至らず、商業利用のために設定されるような高額の使用料を提示することもある。このような価格は研究者からすればほとんど禁止的なものであり、結局利用を断念するというような例が聞かれる。情報試料の公的な蓄積交換機構の整備は、このような事態を打開するために有効に機能すると期待される。

(3) 標準化の研究と標準化機構

情報の良好な流通を確保するには、適切な標準の制定と普及が不可欠である。したがって情報に関連する標準化は、情報学における重要な研究テーマとなる。これまで、標準化は当該技術の実用化、普及段階での事項とされ、標準化それ自体は研究の対象にはなっていない。産業論、技術論の一端として事後的、歴史的な議論がなされることはあるが、現在および将来の標準化は扱われない。情報化の進行にともなって、情報関連の標準化は社会的に重大な意味を持つようになっていくから、これを情報学の研究領域に正当に位置づけて、科学的基礎に立脚した検討を行うことは大いに有意義であろう。

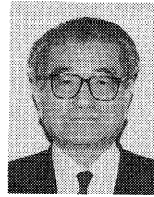
ところで、近年情報関係では、公的な立案・審議過程を介さない、デファクト・スタンダードが重要な地位を占めるようになってきている。これには、技術進歩の高速化に対して、審議に時間のかかる公的な制定機構が対応できないということが要因であると考えられる。また、ファイル形式やプログラム・インタフェースのようなソフトウェア的に簡単に実現される規約の類については、研究者の任意団体による規約が次第に普及し、企業の製品化がこれに追従するというような、新たな現象も現れている。このような情報分野における「標準のソフト化」ともいべき状況に対して、標準化機構はどのようにあるべきか、さらに検討する必要がある。

6. おわりに

インターネットの普及にともない、世界のいたるところで情報の提供、流通に劇的な変化が起こりつつある。遅れをとっているわが国にとって遅れを取り戻すまたとない機会が到来している。この機を逃すことなく新たな情報資源、流通のありかたを含め課題解決をはかって行かねばならない。

以上、3回にわたって情報学とその課題について述べてきた。情報学の間口の広さと奥行きへの深さはだれにも分かっていないというのが本当のところであろう。ここで論じたことはその一部に過ぎず、なおまた問題提起のみに終わり、具体的な行動計画、体制作りに至っていない。第15期日本学術会議情報学研究連絡委員会のわれわれとしても心残りである。

(平成7年4月10日受付)



尾関 雅則 (名誉会員)

1946年東京帝国大学第1工学部電気工学科卒業。同年運輸技官。1972年国鉄電気局長。1975年常務理事。1979年日立製作所入社。1982年OA事業部長。1983年常務取締役。1987年より(財)鉄道総合技術研究所理事長、現在に至る。国鉄時代、マイクロ波通信の実用化、電話の自動交換網の推進・普及、貨物操作場の自動化システムの指導などに従事。また、みどりの窓口などの旅客総合予約システムを開発。この功績により紫綬褒章と電子通信学会業績賞を受賞。現在は主に超電導磁気浮上式鉄道の開発プロジェクトの指導・推進に従事。運輸技術審議会委員、情報処理国際連合(IFIP)理事など。

