

“データベースと情報知識学”

藤原鎮男

本稿は平成元年9月7日、東京で行なわれた情報知識学会シンポジウムにおける招待講演の内容に加筆したものである。

1. データベースの展開（歴史）

表題中の二つのキーワード「データベース」と「情報知識学」のそれぞれについての考察から議論を始めたい。まず、データベースは今日までに3つの段階をへているという見方から始めよう。

1-1 第I期（1960-1970）

この時期に、世界的規模で情報知識の流通改善の動きが起り、それは先ず図書館の改革から始まり、計算機の導入に進んだ。英国のDainton 卿を委員長とする図書館改革は、知識情報の流通を根本的にかつ広い視野で改革する先達となり、そのわく組みは図書館の改革の指標となるものであった。そこでは明らかに図書と学術雑誌の情報の流通を区別していた。

Boston Spa に Lending Divisin を設け、学術雑誌に重点を置いて中央図書館集書の理念で情報知識の流通をはかろうとするのがその基本的考え方であった。

米国も議会図書館（LC）がこの時期に図書館の改革をスタートさせた。こちらは収蔵図書の整理に改革の主点があり、それは図書館業務の改革であった。しかし、そのとき同時に受け入れ（分類整理）とLC MARCによる内容の要旨の機械処理が含まれ、それは知識の流通に画期的なインパクトを与えるものとなったのである。すなわち、この時期に電子計算機は既に顕著な進展をしており、米国に顕著にその動きが始まった。LCのほか、図書館業務の機械処理を進めるOCLCのサービスや学術論文の要旨を機械処理をするアブストラクデータベースのサービスが始まった。医学情報のMEDLERSや化学情報のCASデータベースの構築はそれである。すなわち情報処理に計算機を導入することが本格的に進んだのである。

これら、上記の動向を反映して、とくに科学技術情報の機械処理による流通をはかるため科学技術情報センターが世界の各地に設立されたのもこの時期であり、これは、既往の図書館の線と別に情報知識の流通の流れを生むものであった。両者の調整や、さらに包括的、統合的なものをつくることは、その後、ユネスコのP G Iのほか、世界の国々の個々がそれ以来とり組まなければならない課題となった。

1-2 第Ⅱ期 (1970-1980)

第Ⅱ期は電子計算機の進歩が著しく、その成果が直ちにMEDLERS, CAS などのデータベースの構築のみならず流通に大きな要素となるに至った時期である。I U P A Cの中に化学情報特別委員会が作られたのは1972年で、C A Sが文献情報から化合物情報を探し出すCompound generation, 化合物名からその構造をディスプレイするStructure generationを行なうシステムをつくったのが1974年である。これによって、化学は電子計算機による化合物設計を可能にすることになったのであるが、それが実際化するのには1970年の後半で74年のC A Sは「C A」の製作を完全に機械処理で実現しえたことを誇示するにとどまっていた。フランスのDuBoisはほとんど時期を同じくして同じ機能をもつD A R Cを完成した。後者はとくに数値情報とスペクトル情報の処理すなわち、オンライン計測の結果をデータベース化し、それをオンラインベースで流通路に載せる所に特色をもち、これは、実験と文献情報知識とを連結するという意味で大きな特色をもつものであった。この点において東大はすでにT O O L

I Rシステムをつくって、それに附随して実験室のNMRの測定結果を、オンラインで解析するシステムを作っており、D u B o i sのD A R Cをつくったフランスは、我々を先導試行として大変高く評価してくれたのであった。

第Ⅱ期のもう一つの特長はC O D A T Aの活動がこの時期に活発になったことである。前記D u B o i sのD A R Cはその一部だが、さらに英国のケンブリッジ大学のX線学者が、構造データの機械処理をすすめる大きな貢献を果し、今日に及んでいる。この時期の特長はX線解析におけるデータ取得のように、高速大容量の計算機能力を必要とするため情報のデータベース構築は、限られた大型計算センターしか寄与出来ない事情があったことである。これは大容量高速の計算機の進歩が、即、情報処理の進歩となる性格を与えることとなった。

もう一つの特長としては、データベースのネットワーク構築の努力がこの頃から始まったことも記されねばならない。そして、このデータベースネットワークが有効に機能する上で化学では化合物のすべてが番号に登録されたことは大きな意味をもつ基盤整備だったのである。これによって、文献情報もファクト情報も、あるいは新論文を載せた学術雑誌と100年200年の記録であるGmelinやBeilsteinの内容が一緒に利用しうることとなりSTNがこれで可能になったのである。

さらに、データベースの利用による新しい学術活動がこの時期に始まったことも注意されるべきであろう。フィラデルフィアでガーフィールドが始めたICIはその例である。

1-3 第Ⅲ期 (1980-)

第Ⅲ期は第Ⅱ期の延長上の時期である。その特長は第Ⅰ期、第Ⅱ期に比較して、その目標が錯然たることである。第Ⅰ期には科学技術情報、とくに論文書誌事項と要旨の機械処理、すなわち電子計算機の活用による流通の展開という明確な目標があった。また第Ⅱ期にはデータベースの連帯、すなわちネットワーク化という、これもまた明確な目標があった。ところが、第Ⅲ期にはこのような明確な目標をあげるのは困難である。勿論、機械処理の導入と展開は依然として進められている。たとえば、学術の諸分野、いろいろなディシプリンの中でそれぞれがもつ情報知識のデータベース化が進んでいる(データベース構築)。また完成された巨大なデータベースCAとBeilsteinなどの地球規模の連繋もある(STN)。あるいは化合物の部分検索の実現、日本語表示まで含む化合物辞書データベースの完成、これを基礎とする化合物設計、反応設計の実施など顕著な業績が上がっているのは事実である。しかし乍らこれらは錯然として展開していて、一括した把握が出来ない。この分散の状況が第Ⅰ期第Ⅱ期の仕事との根本的な違いである。すなわち、第Ⅲ期の仕事は専門に立地しているのである。このことの根本的特長は何かというと、情報学、情報知識学の応用がディシプリンサイドによってなされているということに盡きるであろう。すなわち、第Ⅲ期は第Ⅰ期と第Ⅱ期に展開してきた情報学に学術の伝統ディシプリンとが接触した時期なのである。

さて、ここで演者は考えをたて直して見る必要を感ずる。専門伝統のディシプリンが、整理され、流通路に登載された情報知識の応用をはかることは大変結構なことで大いに推進さ

れねばならない。しかし、それとは別に「情報知識学」自体が一つの専門として立ち、アイデンティティを確立することもまた必要なことである。このシンポジウムがデータベースと情報知識学という表題で設けられ、情報知識学会が誕生し、また日本学術会議に情報学の研究連絡委員会が生まれたということは、一つの専門分野として情報知識学が生れ、生長を始めたといことなのであり、従って今、その関係者は、情報知識学の内容を固め、振興をはからねばならないのである。それについて以下に演者の考えを少し述べさせて頂く。

2. 情報知識学

情報知識学の内容としては、専門ディシプリンの分野に既に蓄積された内容を整理し、処理体系を確立して広く情報知識の応用が出来るようにする仕事が含まれることは先に述べた、しかし、それと別に情報知識の専門家がディシプリンとして学術活動をすることは別途に考えねばならない。その線に沿って、どのような活動があるかという、情報の組織化のための用語の整備、情報表現のための言語のホン訳などがまず、あげられる。そしてこれについてはすでに成果があるといってよいであろう。しかし、学術の一つの専門分野としての成立のためにはさらにもう一步具体的な体系的内容の構築が必要と思われる。

F I Dの理事会は1988年4月に「情報知識」の組織化の上で考えるべき四つの対象として次をあげている。

- 1 resource
- 2 tool
- 3 market
- 4 education and training

高等教育機関におけるDepartmentの構築には、まず4本程度の支柱をたて、その柱に沿って内容をつけるのが普通である。そこで演者はディシプリンとしての情報知識学の柱を何にするかを議することが必要であると提議したい。

今、かりに、F I Dがあげた4つの項目を採り上げてよい。これから一応、情報知識学の構築を果たすことが出来よう。その場合、各項目に応じた内容をつけるところが実際の仕事である。すなわち、この各項目の肉づけを、情報知識学のファミリーの人々が自らの「分野の建設」

の使命として働くことがなされねばならない。これが現在の我々にとって最も必要とされることではないかと思う。

情報知識学が以上で成立したとして、ここで演者は、もう一度、その基本性格について考えてみたいと思う。1-3で述べたように、現在、情報学が立地している場面は、一般化と、個別のディシプリンの接触面である。一般というのは、データベースの形で整理された情報知識が地域や、分野を越えて利用されるに至ったことを指すものとする。その方向に沿う仕事は、集中化であり、大容量、高速の情報処理機能を必須の条件とするものであった。ところが、現時点では、情報知識として求められているものは、ディシプリンの専門に立つ個別的なものであり、それを「一般化」のベクトルに沿って整備されたリソースの中から掘み出すことが仕事なのである。これを実際に有効に行うためには、個別的、専門的な自己の要求を整理し、体系化し、すなわち、自分に即したローカルな“要求”のデータベースをつくり、それを一般化された情報知識のリソースにマッチさせて個別的、専門的な自分だけに有用な情報知識を形づくる作業が必要なのである。くり返していうと、第Ⅰ期、第Ⅱ期の情報知識の整理の流れを一般化の方向とするならば、このベクトルに対し、反対のベクトルがあり、それは専門に立地した情報処理なのである。この二つの相反する流れの接触面が第Ⅲ期なのであると見られる。

従って、専門に立地した情報処理は、分散的であり個別的である。断っておくが、近年盛んにいわれている専門分野毎のデータベース構築は、今迄の議論の上でいうと「一般化」の流れに属する作業で、反対ベクトルの仕事の中のことではない。また、情報検索の業務、サーチャーというのも、「一般化」の路線の先端にいる人であり、これも反対ベクトルの中に入らない。

専門、ディシプリンの側に属して、その分野のニーズを知り、そして、一般化された「情報知識ベース」の活用をなし得る人そういう人が第Ⅲ期の現在はもとより、今後も必要とされる情報知識学の専門家であろう。

情報知識学の柱の中にこれを採り上げるとすればそれはET. すなわち「教育訓練」であろう。具体的に既在の専門分科の分れる前、すなわち一般教養のあとの専門基礎の段階に「一般化」された「情報知識ベースの利用」のカリキュラムを設けること、大学院の修士レベルのカリキュラムに、現在、各種の機器分析のコースが導入されているのと同様に、その専門分科に

即した情報知識データベースの利用法、要すれば自分用のデータベースの構築法を身につけさせる教科を設けること、これが情報知識学という専門分野が構築目標とするカリキュラムとなるのではないかと思う。そして、社会人教育のレベルについても全く同様のことが構想されるのである。演者はこれを、新しい時代のゲートキーパーの登場とよぶことにし、社会がその必要性を認識してくれることを求めているものであり、今年11月にはその国際研修会を行うことを予定している。

3. 情報知識学のテーマ：新安全学

専門のディシプリン、すなわち一つの分科として学術の分野が存在するには、その分科が自らの分野の活動の目標をもたねばならない。研究対象をもつことである。その目標として演者は、「新安全学」をここで提案したい。

近代社会の科学技術の進展に伴って、人間活動の安全は人類の将来まで左右する状況にある。ところで、エネルギー、化学（物質）、構造、輸送のそれぞれは、現在自己の最善をつくして、それぞれの関わる対象の安全を図っている。しかし現実には、その上にもなお安全をはからねばならないのである。それには、現在ある個々の安全の努力を超えて、さらにその上の安全を追求しなければならないのである。ここに情報知識学の出番がある。個々の分野を超えた専門としての情報知識学の活動分野があるのである。そこで、仮称として「新安全学」を演者は情報知識学のテーマとして提唱するのである。