

データ開発と情報システム*

北川 敏 男**

1. まえがきと要約

経営情報システム (MIS), データ・バンク, データ・ベース等が, 時にはジャーナリズムによって社会の注目をうける論題となつてから, すでに数年が経過している。もちろん, 着実なそして実用に役立つ研究開発も進められている。しかし現時点においてこそ, この方面の組織的な研究がスタートしなければならないと思われる。この小論は, このことに関連して,

㊤ データの意味を系統的に再検討し, われわれの必要とするデータとは, 何かを明らかにすること。この観点からいうと, **データの開発**という課題が, われわれの前におかれている。

㊦ 情報システム (information system) の構想を明らかにしたい。とくに**巨大情報システム**に課せられるべき任務を明らかにしたい。この二つの事項を主眼として, 私見を展開したい。組織的に研究の必要を主張する小論の骨子は, 次の五つに要約されよう。

第 1 データ開発と巨大情報システムの設計とは, 社会的な needs を明らかにしてかからなければならない。このためには, **情報化のビジョン**をある程度, 探求し, これに対する**情報化政策**についても, 検討しておかなければならない (第 2 節参照)。

第 2 情報化社会における学問のもつ役割を検討し, 当然, 予想される**知識システムの再編成**についても省察してかかる必要がある。**知識社会の要請**ともいうべきものが, 把握されなければならない (第 3 節参照)。

第 3 データ開発という問題をここに提起しなければならない。データの変換, 伝達, 貯蔵, 検索の諸問題はいうまでもないが, データの生成起源, 利用方式にわたって, 現状および将来像を論じておく必要がある。データを所与とみず, データをつくり出してゆく立場が要請される (第 4 節参照)。

第 4 巨大情報システムのイメージをはっきりさせるため, 一例として**学術情報システム**についてのある構想を述べる (第 5 節参照)。

第 5 データ開発と巨大情報システムは, **情報資源共同利用ネットワーク**の問題に関連し, 情報処理技術に依存するとともに, パターン認識, 自己組織化系, 総じていえば人工知能の理論と技術に対して, 多くの具体的問題を提供し, **情報科学の進展**に寄与することが予想される (第 6 節参照)。

2. 情報化ビジョンとデータ開発に対する社会的ニーズ

「80 年社会の情報化ビジョン」(日本経営情報開発協会 1971年)において, 何故に情報化ビジョンを必要とするかが説明されている。それによると,

(i) 社会経済的諸困難の克服への寄与,
(ii) 工業化社会から情報化社会への移行の効率的達成,

(ii) コンピュータリゼーションの進展に伴い, コンピュータの封鎖の利用から情報ネットワークの形成への進出

があげられている。さらに 80 年社会における中間的な情報化ビジョンとしてあげられて総合的目標とされているのは, 「コンピュータ・マインド (情報化社会に対応してゆくための思考方式) の定着」である。この総合的目標を達成するうえの具体的なターゲットとして設定されたものが, 9 つあり, 全体として**ターゲット系**は, 次のように与えられている。

- (1) ホーム・ターミナルの普及。
- (2) 全国情報ネットワークの形成。
- (3) 行政の合理化。
- (4) コンピュータ志向教育。
- (5) 医療の近代化。
- (6) 公害の防止と制御。
- (7) AVI (自動自動車識別) による交通管制。
- (8) MIS の高度利用。
- (9) 流通機構のシステム化。

* Data Development and Information System, by Toshio Kitagawa (Research Institute of Fundamental Information Science, Kyushu Univ.)

**九州大学理学部基礎情報学研究施設長

(10) コンピュータ・マインドの定着.

ここでは、これらターゲットの意義、相互関連等を論じることは、差控える。ただこれらターゲットの時間的な優先度として、(1)および(2)があげられていることに注意したい。さらに10年間のタイム・テーブルとして、80年までに計画どおり達成されるためには、情報化政策に対する基本大綱の策定にもとづく基本的全国情報ネットワークの整備と政府の行政データ・バンクの形成が75年までにという期限付きで要請されている。これに並行して、社会開発関係のプロジェクトが、少くもモデルケースの形で着手されていなければならないことが指摘されている。その例としてあげられているのは、

- (i) 小中学校のコンピュータ志向教育
- (ii) 救急および遠隔医療システム
- (iii) 公害警告システム
- (iv) 交通管理システム
- (v) スーパーの無人化システム

である。いずれも、computerizationの方向としてはさして異論はあるまい。けれども、しかし、何時から如何にしてというような具体的な点になると、事柄は、そう簡単ではあるまい。いずれにせよ、ここでは、(a)データ・バンクの情報とネットワークの形成が、情報化の前提となるという主張に注意し、かつ、それらは、(b)社会開発プロジェクトの着手と関連していると主張されていることに注目したい。

この小文の趣旨を理解するのには、(a)および(b)という二つの主張のあることを、まず認識しておくことが必要である。しかしそれだけでは充分でない。それゆえにこそ、以下数節にわけて、検討すべき問題を、当てる。大筋の方向は明らかであっても、行く途は、種々ありうる。その選択こそと重大であるといわなければならない。

3. 知識システムの再編成

情報化社会は、学習社会ともいわれ、知識社会ともいわれて、その未来像が論ぜられたのは、つい2、3年まえまでの状況であった。現在、われわれは、経済成長とか技術革新とかに接続した方向線においてだけ、バラ色の未来像を見ようとしても、それは許されまい。1960年代においてある人達は広言した。未来は制御可能であり、設計可能であると。しかしそれは限られた部分的なシステムであっただけではないか。“人類の英知の結集”が要求されること現代ほど切実

なことではない”ということを目指したい。人類の英知の結集には、もろもろの方法があるが、現代の科学、技術が専門分野に細分されて、それぞれが、部分的な知見の集積におわっている状況は何としても打破することが、まず必要であろう。第3節で述べた社会開発プロジェクトというようなものでも、いかにしていわゆる専門分野の知見を結集してゆくかということになると、方法論の開発が要請されているのである。(i)―(iii)の如きにいたっては、教育学、医学、環境科学などが既成のまま、それに利用できるものであろうかどうか、改めて省察する必要があると思われる。このような省察において、提起される問題は、外ならぬデータの開発であり、情報処理システムの設計である。どの学問でもデータを取扱わぬものはない。何等かの情報処理のシステムも存在してきたはずである。しかし、人類社会にとっては新たな生存課題を前にしてこれに対処できる新たな機能と性格をもつデータと情報処理システムこそが要請されるのである。この新たな課題に対しては、基本的には知識システムの再編成が要請されている。この要請にこたえる研究方略がまず設定されなければならないというのが、私の主張してきたことである。一つのサンプルとして、次の研究方略をここに再記しておきたい。(拙論：“環境と営存圏”講座情報社会科学第10巻、学習研究社、1972.)

研究方略 1 自然認識の基柱を物質・エネルギー・情報の3つにとることにより、物理科学中心の在来の自然科学に対し生物科学の独立的な位置づけを確保すること。

研究方略 2 人文科学の構成全体に情報の概念を基軸としてその再編成をはかり、広義のシステム概念を通じてこれと社会学との連関を系統的にすること。

研究方略 3 上記の2つの研究方略を実現させることによって、漸次、自然―生態―社会を貫通する汎科学的な方法論を築きあげる。これによって、自然科学と人文社会科学との間に従来設けられた境界をしないで除去し、必要に応じて結合体系をつくりうるようにすること。

研究方略 4 時間的な軸については、自然史・人類史・社会史などのもろもろの歴史を連結させる歴史科学体系が構想されねばならない。現時点から過去へさかのぼる歴史科学に対し、未来に関与する未来学がその方法論を整理しなければならない。

研究方略 5 科学の従来もってきた基本機能は、認

識・効用・批判の3つにわたってきた。この基本機能のいずれを欠いても、科学は健全とはいえない。この3つの基本機能の結合によって、科学は制御への利用、営存の確保、創造の伸長に寄与すべきである。この目標から、政策科学だけでなく、環境科学、創造科学の開発が必要となる。

この5つの研究方略を前提にしなければ、以下の所論が進められないというわけではない。ただ、情報、システム、結合体系、時間的な連続、基本機能の総合という視座をつくることは、この小文の趣意を活かしてゆくのに必要なように思われる。ここに掲げられた研究方略は、個別科学のモジュール化およびモジュール化された科学のシステム化を意図することによって、実現できるであろう。この二つが、知識システムの新編成の基本となるわけである。これに関連して、“個別的な純粋科学について、その理論のよってたつ前提を明確に抽出し、その理論の演繹の構成を明らかにする。さらに、その実験や観察の対象とする現象が、いかなる事象切断面に関連するものであるかを明らかにする。このためには、その学問の実験観察の機具、装置、組織などを改めて吟味し、相互間の異同、類別などを組織的に整理してみる必要がある。”このことは次節のデータ開発論へつながってゆく。

4. データの開発

コンピュータ・サイエンスの立場では、「データとは、マニュアルまたは機械的に処理・伝送できる形に記写したあらゆる情報である。」さらにデータ・ベースの背景となった基本的な狙いは、プログラムとは切り離して、データの環境や属性を記述し、プログラムを実行する時点においてはじめてプログラムとデータとの結合を工夫するという点にある。データ・バンクの設立に伴いファイルの総合と一元化、オン・ラインシステムの普及と情報資源 (information resources) の共同利用などによって、促進されよう。こうした視点からいえば、利用者とデータ・ベースとを結ぶところのオペレーティング・システムのなかでも、端末管理プログラム、処理管理プログラム、データ・ベース管理プログラムが大切であり、重要な研究課題である。プログラムとデータとを切り離すとき、データ自体に、構造をもたせることも当然おこる要請となる。しかしこの小文で指摘したいことは、計算機科学の当然攻究すべきこのような正統的な課題の重要性だけではない。その基底には、より根本的な問題があるとい

うことである。簡単のため、簡条書にして要点を述べよう。

〔1〕 データを所与として取扱うまえに、データのもとになる情報の発生源を類別しておく必要がある。野外における調査、実験室における実験、図書館、情報センターにおける文献調査、書斎における推理等々の獲得場面を、とりあげてみる必要がある。

〔2〕 情報には、かくして、野外調査による情報、実験による情報、文献調査による情報、推論による情報等が存在するのであるが、これがいわゆるデータ化するまでの作業過程をあたってみる必要がある。

〔3〕 データがいかに利用されるかの利用形態を類別しておく必要がある。データの利用の仕方は、他のデータと組み合わせられ、ある知識システムのなかに投入されて、然るべき役割を果たすというのが、一般の状況であろう。極めて大づかみな見方であるが、データの利用について、次の三つのどれであるか、どこに重点があるかを見ておくことは有意義である。(a) 貯蔵 (storage), (b) オペレーショナルな利用 (operational uses), (c) パターン認識 (pattern recognition)。

〔4〕 現在すでにもろもろの分野において、種々の相異なる視点から展開されているデータ論ともいべきものがある。それらを、情報科学の立場から共通の理論構成のもとに整理しておくことが必要である。例えば、

(i) 統計学では、標本調査、実験計画、品質管理等における経験を通じて、統計資料 (statistical data) の取扱いのために、多くの具体的な方法を築き上げてきた。それらの知見は、むしろ技術者の方面においてより広く普及しているが、基礎科学者の理解を深めなければならない。

(ii) 科学情報のなかでも、物質の物理化学的性質をあらゆる諸定数、天文学や地学において観測され決定された諸量の数値がある。データの信頼性の検定、標準データの確定など、すでに CODATA (Committee on Data for Science and Technology) の活動がある。この方法に広く学ぶべきものがあろう。

(iii) 見落してならないのは、野外科学 (field science) (地理学、文化人類学、民俗学等々) におけるデータの蒐集およびその処理の方法論である。「データをして語らしめる」という態度は、どこまでも失ってはならない基本原則の一つである。

(iv) データを、単に認識と批判という科学の機能だけからみてはならない。効用というもう一つの機能

からもみなければならぬ。例えば **MIS の設計** を想定せよ。そのデータ・バンクを利用する仕方が、第一線のエンジニア、ミドル・マネージメントそしてトップ・マネージメントによって、実際いかに異なるかを観察することも必要である。官庁統計に関して、データ・バンクの設立が課題になっている。そのときにおいても、いかなるデータを結合して利用しようとするかをあためなければならぬ。

(5) 情報処理の機械化は、データ処理の具体的な場面に応じて、それぞれできるだけ導入されるべきであろう。(a)実験、(b)統計、(c)ドキュメンテーション、(d)図書館業務、(e)病院等の事務及び診療の一部のオートメーションが問題になっている。例えば **automated statistician** という新造語さえあらわれている。ここに情報資源の共同利用ネットワークの問題がおこる。

第1節であげたターゲット、(4)、(5)、(6)以下を具体化してゆくとき、必ずや以上あげた諸点からデータ論を展開しなければならなくなるであろう。それは既存のデータ論でない。新しい見地に立つところのデータ開発という問題への組織的な接近が要請されるのである。

5. 科学情報システムへの情報科学的接近と その科学研究に対する意義

標記の論文(英文)を第1回日米協会図書館長会議で発表したのは、1968年10月のことである。紙面の都合でここで立入って紹介することはできないので、以下に目次を掲げるとどめる。

- § 1. まえがき
- § 2. 情報化時代の情報問題
- § 3. 情報問題の三つの視点
 - § 3.1 機能視点 § 3.2 システム視点
 - § 3.3 調整視点
- § 4. 二つの接近法の提唱
 - § 4.1 情報科学的接近 § 4.2 方略的接近
- § 5. 結論(5つの提言)

この文献は、九州大学理学部付属基礎情報学研究施設研究報告、No. 3(英文)で刊行されている。

6. 情報科学的接近

これについても、全般を論じる余地はないので、前掲の論文 § 3.1 に言及するとどめる。それだけからでも、データ開発と巨大情報システムの問題において、

高度の情報科学的接近が要請される理由は、看取できようと思う。

情報問題に対する機能視点として、(i)情報の組織化、(ii)情報の伝達、(iii)情報サービスの三つがあげられよう。ここでは、その第1視点にのみ言及しよう。ここで情報の組織化に関連して、次の(a)、(b)、(c)、(d)4つの問題がある。

(a) 情報の貯蔵を増加させる演繹過程

シミュレーション技法の利用により、既存の情報より演繹できる知見を追加すること。

(b) 情報の簡約過程

このためには、既存の情報貯蔵に対して、真に情報(non-trivial new information)とは何かを問題にしなければならぬ。

(c) 情報の貯蔵に対して適用されるべき自己組織化

non-triviality という概念を導入する。あるデータが non-trivial か否かは固定的ではない。再組織化がつねに貯蔵に対して要請される。

(d) 情報の貯蔵と学習過程

演繹および、帰納については、コンピュータの利用によって、(i)幾何学の定理の証明、(ii)統計資料の解析、(iii)観測データの蒐集および分析等において、自動化がすすめられている。そこでは、学習過程によって、これらの自動化の効率の向上が当然期待される。そうすると、情報の貯蔵の方式においても、構造化が進められることになる。

情報資源の共同利用のために、TSS の利用も当然起る、これからの重大な課題である。ここでは最後に次のことを強調して筆を措きたい。

データ開発と情報システムの設計とは、すでに述べたように、情報化の礎石である。しかしそれは同時に、情報科学に対して、核心的な問題を提起するものである。パターン認識、自己組織化、一般に人工知能論は、ここで当面する課題に対していかなる解答を用意しうるかによって、その実質的な価値を問われることになるであろう。言語論理も言語装置も、同じくこのような試練を通じて、成育するものと期待される。上述の機能をもつ情報システムは、多方面にわたる大量のデータを駆使し、高次の情報処理を行い、広域にわたっての情報流通のネットワークであろう。情報資源の共同利用ネットワークの巨大情報システムのシステム設計がわれわれの目標となるであろう。

(昭和47年8月31日受付)