

〈特集：公衆衛生情報ネットワーク・システムの構築をめざして〉

研究を支援する情報・計算基盤環境について

神 沼 二 眞

1. はじめに

公衆衛生分野の研究開発に、コンピュータで代表される情報や計算の技術を如何に役立てたらよいかという問題を考えてみたい。これを敷衍すれば「(科学技術の)研究開発を支援する情報システムやコンピューティング環境を如何に構築し、如何に活用するか」という一般的な問題となる。さらにこの問題は、人間の創造的な知的活動をコンピュータでどう支援するかという、さらに一般的な問題の特殊なケースと見なすことができる。

筆者がこうした問題意識に目覚め、試行錯誤の中でおよそのデザイン原理を確立し、そのプロトタイプと呼び得るシステムを実際に開発したのは、前職の東京都臨床医学総合研究所にあった1980年代のはじめであった¹⁾。このシステムは基礎医学、臨床医学、医学研究さらに疫学の研究支援を前提としていたが、同時にドラッグ・デザイン、毒性研究、素材の研究開発を支援する機能も含まれていた(図1参照)。こうしたシステムの実際とその開発思想は、1980年に設立されたCBI研究会(計算機と化学・生物学の会)の活動を通じて、大学や企業の研究現場への普及が計られた²⁾。さらに、幅広い科学技術分野における研究開発へのコンピュータ利用に関しては、わが国の科学技術の最前線にいる約100名の研究者の協力によって後に編纂された「R&D コンピューティング要覧」によって、集大成を試みた³⁾。この主題のうち、分子モデリングや生物医学に的を絞った解説書もほぼ同じ頃出版した^{4),5)}。1990年までの筆者らのこうした実践活動は「研究開発を支援する情報システムやコンピューティングを如何に構築し、活用するか」という問題に一応の答えを与えたものと考えている。

(国立衛生試験所化学物質情報部)

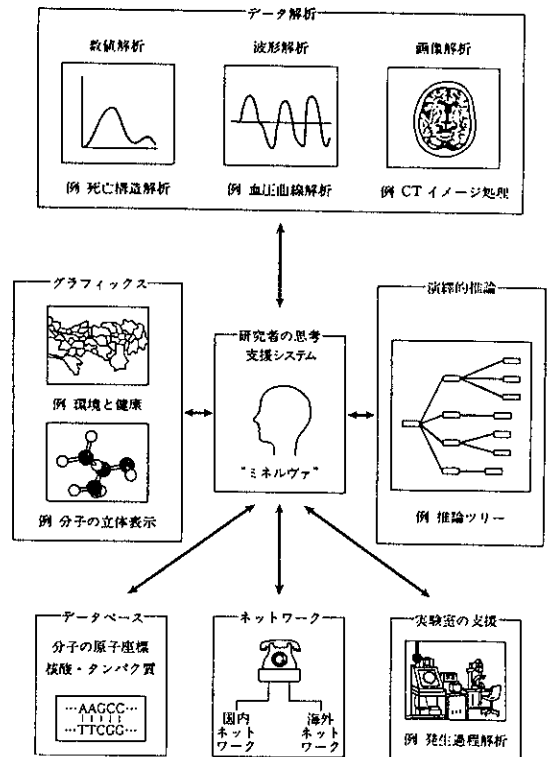


図1 筆者らが東京都臨床医学総合研究所で、1980年代前半に開発整備した、生物医学の研究を支援するシステム。帰納、演繹、発想、と言う普遍的思考支援システムに、分子モデリングなど、分野別支援システムから構成される。国際コンピュータ・ネットワークにもすでに接続可能だった。

もちろん、公衆衛生研究の基盤である分子生物学、基礎医学、臨床医学、疫学などは、現在も文字通り日進月歩の勢いで進歩している。学問の進歩は、常に新しい領域を生み、コンピュータもまたそうした分野で使われるようになる。しかし、それによってただちに新しいコンピュータの利用形態が生まれるわけではない。コンピュータの利用形態の変化は、むしろコンピュータ自身の進歩に依存している部分が大きい。こ

の意味で言えば、現在広域ネットワークであるインターネットの著しい普及がコンピュータの利用「環境」を一新しつつある。以下では、こうした最近のネットワークの進歩を考慮に入れて、研究を支援するシステムをどのようにデザインしたらよいかについて述べてみたい。

2. 研究者の情報環境

公衆衛生には、生命に関する科学的研究と、目標を立ててそれを達成する計画の学問、つまり制御工学的な研究が含まれている。一般に科学研究の目的は自然法則の発見、真理の発見である。これに対して、工学の研究は、新しいシステムや仕組みや道具をつくり、目標を達成することである。いずれも高度に創造的な

人間の活動の一例である。この創造的な活動の最初の索引力は、好奇心やアイデアなどである。

科学研究でも工学研究でも、実験あるいは調査を行う。その目的はデータを得ることである。したがって、一般に科学や工学の研究者は絶えざる情報の流れの中に身を置きながら、情報（データ）をえ、情報を操作して、新しい仮説をたて、実験を繰り返して、価値ある情報（知識）を再び生み出そうとする。工学の研究では、これに加えて、具体的なものをつくり出したり、問題を解決しようとする。

この情報という観点から見た場合、研究者は常に3つの異なる世界とかかわっている。すなわち、外の世界、思考を専らにする研究室(オフィス)の世界、データを生成する実験室あるいは調査のためのフィールド

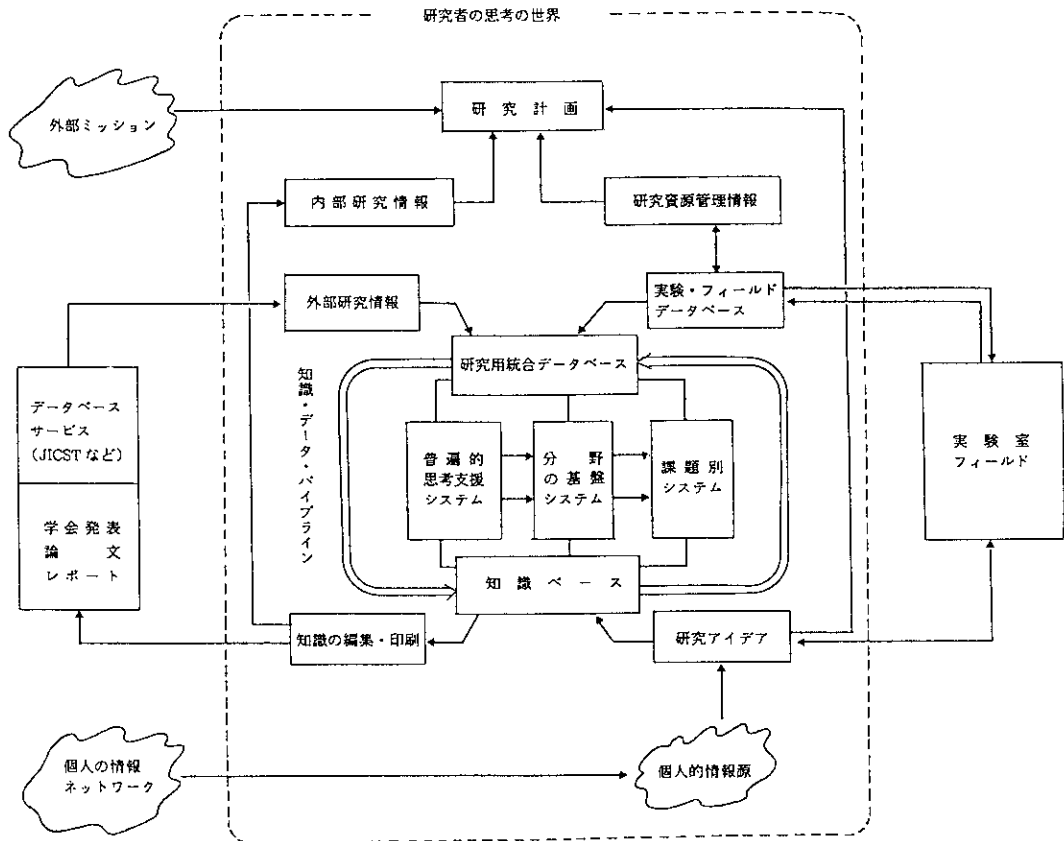


図2 統合的に見た研究者の情報環境

下位の研究者は研究に関する個人的情報源に基づいて研究を計画し、上位の研究者は組織のミッションと、研究資源から計画を考える。中間の研究者は両方の立場から考えなければならず、一番負担が大きく、支援システムのデザインは最も難しい。

の世界である(図2参照)。普通、研究者はまず外の世界から研究情報を仕入れる。一方で、実験室やフィールドでデータを集め、それを自分の経験知識や直感に照らして分析し、そこから新しい知識を生成しようとする。その結果は(組織内も含めて)外の世界への報告書や論文となり、また新たな研究計画や実験へとフィードバックされてゆく。こうした活動の違いを反映して、それぞれの世界におけるコンピュータの利用形態も著しく異なっている。

一般に、研究におけるコンピュータ化は、データ処理を例外として、まず外の世界と実験室から始まった。外の世界におけるコンピュータ化とは、MEDLINEやCASで代表されるような、文献やファクト・データのオンライン・データベース化である。また実験室とフィールドのコンピュータ化は、各種の計測機器のコンピュータ化とそこからのオンラインによるデータの収集である。こうしたコンピュータ化は、現在さらには思考の世界である研究室(オフィス)に及んでいる。

コンピュータの利用形態から見ると、外の世界と実験室あるいはフィールドにおけるコンピュータの利用法は、研究の分野にはあまり依存しない。思考の世界におけるコンピュータ利用は、分野に依存するが、これもよく見ると分野に依存しない部分と分野に依存した部分とから構成されていることがわかる。このように分析してみると、研究を支援するコンピュータ・システムあるいは、コンピュータ環境は、研究分野に依存しない部分が大変に大きいことがわかる。言い換えれば、研究を支援するコンピュータ環境のデザインは、特定の分野に依存せず、共通の思想にもとづいて行えるのである。

いずれにしても、まず考えるべきは、基盤となるコンピュータのハードウェア環境をどう整備するかである。つぎに、これについて考えてみよう。

3. 研究者のコンピューティング環境の整備

3.1 ワークステーションとPCの共存

現在、研究者の標準的な“コンピューティング環境”は、複数のワークステーションと複数のパソコンがローカル・エリア・ネットワーク(LAN)で結合された状況であろう。ワークステーションはしばしば特定の目的の仕事が割り当てられた“サーバー”の役割を

果たす。そうした役割とは、広域ネットワークおよびLANの通信制御の管理、大規模なデータベースの管理、高速の計算能力が必要とされるアプリケーション、例えば分子軌道法や分子動力学のような分子計算、分子グラフィックスなどである。すなわち、ワークステーションは、ひと時代前の汎用機のように研究組織全体に共通する比較的規模の大きな仕事を担当する。

これに対してパソコンは、研究者の個人の事務的な仕事、研究情報の管理、ワークステーションの端末などに使われる。しかし、一部のパソコン自体の計算機能や記憶容量、画面表示機能は、もはやワークステーションと肩を並べるまでになっている。

両者のいちばん大きな違いは、基本ソフト、すなわちオペレーティングシステム(OS)の違いである。すなわちワークステーションはUNIXであり、パソコンはMS-DOS、OS/2、Windows、あるいはマッキントッシュのそれである。後者のOSでは、OS/2を除き、複数の仕事を同時併行的に遂行するマルチタスキングができない。こうした状況は、ネットワークウェア(Netware)やWindows NTなど、パソコン・ネットワーク用のソフトの普及で変化しつつある。ただし、パソコン・ネットワークの中心となるサーバー自体は高性能であってもやはりパソコンであり、これは、普通研究所全体の共同利用マシンである。こうした状況はまだしばらく続くと思われる。

3.2 インターネットへの接続

インターネットのような広域ネットワークは、こうしたクライアント・サーバーの組み合わせからなるLANにルーターを接続した形で構成される。このルーターは、外の専用の通信回線あるいは、公衆回線に接続される。こうした通信回線によって、世界中の研究所のLANが相互に接続されることになる。インターネットは、一種のバケツ・リレーのような方式でデータを転送する。したがって、インターネットに接続されるためには、すでにこのネットに接続されている先達のコンピュータに接続させてもらう必要がある。インターネットの発祥の地は、アメリカである。したがって、日本での場合、アメリカの元祖インターネットか、それに接続されているコンピュータ(あるいはネットワーク)を探して、そこに接続させてもら

う必要がある。

こうした接続先としては、数年前までは、TISN(東大国際理学ネットワーク)、WIDE(慶応大学環境情報学部)しか無かった。現在は、これに、学術情報センターのSINET、JICSTのGDB(ゲノムデータベース)、科学技術庁の省際ネットワークが加わった。その他に、営利を目的とする多数の商業用のインターネット提供者(プロバイダー)が生まれている。しかし、いまのところ国の研究機関や、地方自治体の研究機関が接続しやすいのは、省際ネットワークとGDBである。

3.3 データベース関連技法の進歩

データベースは知識やデータを蓄積しておく、研究支援の基本的な道具である。最近従来のデータベースの概念を全く変えてしまうような、新しい技法がいくつか登場してきた。これらを列挙すれば、

- (1) マッキントッシュのハイパーカード、インターネットのGOPHERやWWWおよびMOSAICのような、ハイパー・テキスト型の情報検索ソフトウェア
- (2) 従来の数値や記述情報だけでなく、動画を含む画像や音声や、音楽を含む音なども扱える、マルチメディア・データベース
- (3) クライアント・サーバー型、あるいは分散型と呼ばれる、ネットワークで結ばれた複数のコンピュータ上にデータベースを構築する手法
- (4) 現在主流を占めるリレーショナル・データベース(RDB)に代わるオブジェクト指向データベース
- (5) SGML(Standard Generalized Markup Language)と呼ばれる、出版やデータベースに適するように、文書(テキスト)に編集や検索用の記号を付する文書処理技法
- (6) 文章だけからなる、いわゆるインフォ・ベースを高速で検索する全文検索技術

などである。

これらの手法の多くは、概念としては昔から提唱されていたが、実用的なシステムが登場してきたのはこの数年の間のことである。この背景には、パソコンやワークステーションの性能の向上、画面を分割表示す

るマルチ・ウィンドウ技術の進歩、ローカル・エリア・ネットワーク(LAN)や、インターネットのような広域ネットワークの普及がある。こうした手法はまだ発展途上にあるが、さらに洗練され、実用性を高めてゆくものと思われる。こうした技術の新しい流れは、従来と大きく異なる、新しいタイプのデータベースを構築したり、新しい研究のスタイルを可能ならしめる可能性を秘めている。

3.4 地球を覆うハイパーメディア情報網

インターネットの普及は著しいが、そのインターネットも1980年代までは、数ある広域ネットワークのひとつに過ぎなかった。この数年インターネットが大変身し、一人勝ちと言えるほど圧倒的な勢いで普及している。その理由としては、米国の副大統領A. ゴアによって、商業利用が解禁されたことと、WWW/MOSAICの組み合わせによるような優れたグラフィカル・ユーザー・インターフェース(GUI)が開発され、無料で手に入るようになったことが大きい。このインターフェースは、画像、動画、音も扱うことができる。この意味で、ハイパーメディアと呼ばれる。

こうしたソフトウェアのもう一つの特徴は、多少の技術を習得すれば、誰もが情報を発信できるようになることである。事実、すでにインターネットに接続している多くの研究機関は、続々と、WWW上にハイパーテキスト型の情報を載せはじめた。この中には、従来データベースやCD-ROMの形で提供されていた情報が多数含まれている。すなわち情報の発信もマルチメディアになってきたのである。

WWWによる情報発信には、これまでになかった利点がある。それは、情報発信しているWWWサーバー同士を容易にリンクできることである。これによって、利用者は、世界中に広がる膨大な情報を、あたかも巨大な単一のライブラリーのごとく、しかも居ながらにして探索することが可能になった。これが、インターネットのナビゲーションあるいは、サイバースペースのサーフィンと呼ばれる利用法である。

この時間問題になるのは、どこにどういった情報があるか、それをどのようにとってくるか、である。こうした問題の解決自身をコンピュータにやらせようとして研究されているのが、エージェントと呼ばれるソフト

である。将来研究者は、自分自身のエージェントによって、望みの情報を地球をめぐり巨大な分散情報群のなかから瞬時に入手できるようになるだろう。現在、これにもっとも近いやり方は、自分自身の WWW のホームページを作成し、そこからよく参照する世界中の

WWW サーバーへナビゲートできるように、リンクを張っておくことである（図3参照）。こうした手法は、異なるネットワーク上のデータベース同士を統合する最も簡単な方法でもある。

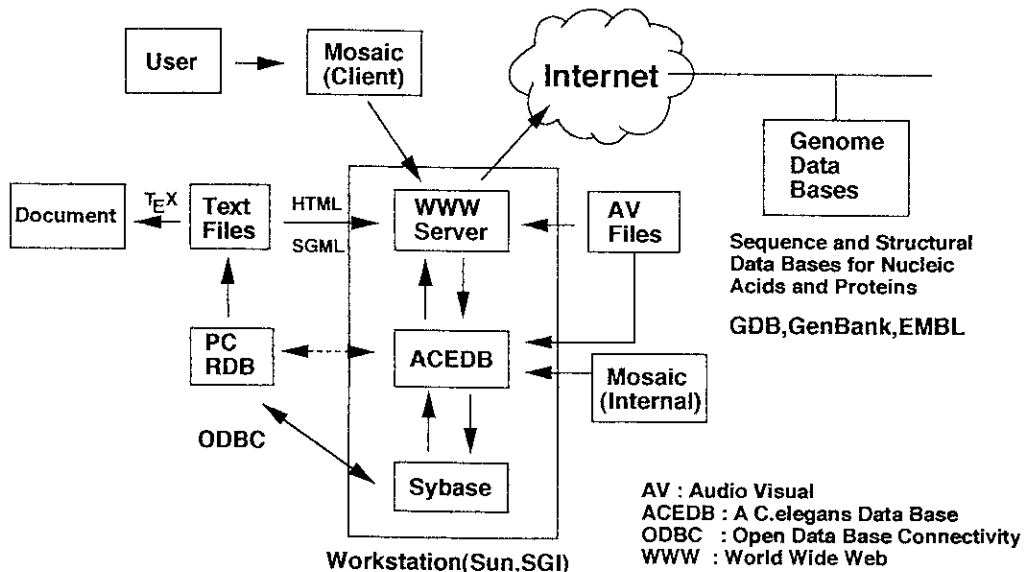


図3 新しいネットワーク環境下でのデータベースの開発技法の概念図。研究者は、手元のコンピュータから、インターネットに接続された世界中の情報資源から必要なデータや情報を瞬時に入手する。

表1 情報・計算のための基盤環境整備に関連した仕事

- (1) 基盤となる通信コンピュータ機器の整備
 - ・広域ネットワークへの接続環境
 - ・所内 LAN, CD-ROM の多重ボックス, グループウェア
 - ・FAX サーバーなどその他の共同通信機器
 - ・無停電電源, その他のコンピュータ関連機器
- (2) 共同利用コンピュータ環境の整備
 - ・共同利用データベースソフト
 - ・スーパー, あるいは並列型の超高速コンピュータ
 - ・高額アプリケーションソフト
 - ・カラープリンターなど高額周辺機器
 - ・画像処理のような専用機
- (3) 研究支援サービスシステムの整備
 - ・図書館システム
 - ・情報検索サービス
 - ・ユーザー教育
 - ・利用マニュアルの作成
 - ・コンサルテーション
- (4) 各部門における研究への応用
 - ・新しいスタイルの研究
 - ・これまでの研究の効率化
 - ・ネットワークによる共同研究
- (5) 研究事務への応用
 - ・全所的なマネジメント
 - ・事務部門の業務の革新
 - ・各研究部門における事務の革新
 - ・事務部門と研究部門とのコミュニケーション
- (6) 広報, 情報発信
 - ・広報誌などの刊行
 - ・WWW などによる情報提供
 - ・問い合わせへの応答サービス
 - ・基盤環境の案内カタログの作成
- (7) 共同研究, 共同事業のための環境整備
 - ・国際機関との接続
 - ・関係する行政部署, 国の研究機関, 大学などとの接続
 - ・地方自治体の関連部署, 機関との接続
 - ・その他の非営利機関, 団体との接続
 - ・細胞バンクのような特別な事業の支援
- (8) 情報部門の拡充
 - ・基盤システムのためのスペースの確保
 - ・コンピュータ関連スタッフの確保
 - ・基盤整備のための予算獲得

3.5 研究機関全体の基盤整備

ひとつの研究機関全体の共同利用を前提にした、情報および計算の基盤環境を整備するには、かなりの仕事がある。これを、筆者自身の経験に基づいて表1に列挙してみよう。

こうしてみれば、必要な仕事が非常に多いことがわかる。情報や計算のための基盤環境整備には、コンピュータ委員会とか、情報委員会などによって企画や運営方式を決定し、あとは外部の企業などに開発を依頼することがよくおこなわれている。もちろん、アウトソーシングは必要だが、効果的な環境づくりには、どうしても専任のチームが必要である。これについては、後でまたふれる。

4. 研究者の思考を支援するシステム

4.1 分野に依存しないシステムと依存するシステム

研究機関全体の話は、ひとまずをき、つぎに研究者の創造的な思考活動を如何にコンピュータで支援したらよいかについて考えてみよう。

創造的な思考過程には、それぞれの専門分野固有の知識と固有の思考パターンが使われる。同時に分野に依存しない普遍的な知識基盤と思考パターンも使われる。そのような普遍的知識基盤とは、数学ないし数式(例えば微分方程式)で表現されたモデルと、帰納、演繹、発想、計画という人間の4つの理性的な思考の基本パターンである。

したがって、研究者の思考を支援するシステムには上記の4つの思考の基本パターンに対応するコンピュータシステムが含まれていなければならない。これらは普遍的思考支援システムと呼べるだろう。なぜなら、これらのシステムは、物理学だとか、化学だとか、薬学だとか、医学だとかいうような専門分野にはまったく依存しないからである。そしてこの普遍的思考支援システムの上位には、それぞれの専門分野の基盤となる知識(モデル)やデータと、それを扱うシステムがくる。思考を支援するシステムは、このような階層性を考慮に入れてデザインしなければならない。

4.2 普遍的な思考支援のコンピュータシステムの構成

つぎに普遍的思考支援システムが、どのようなシステムであるかを、医薬品の開発研究などを例にとりな

がら説明してみよう。

まず数学や数式で表現する手法については、ほとんど自明であろう。すべての現象を、数学モデルによって説明しようというのは、現代科学の時代精神である。この意味で数学的手法は、あらゆる科学的手法の根幹であり、コンピュータの応用としてももっとも歴史が古い。今日では数値計算だけでなく、数式の記号の変換をやってくれる「数式処理」プログラムも広く使われている。

つぎが、帰納であるが、これに関しては、数値、波形(信号)、画像などのデータ解析や、統計学的手法、パターン認識などが対応する。なぜなら、これらはすべて「集まったデータから一般的に通用する規則を導き出す」ための思考の助けをする分析、判断のための道具(ツール)だからである。こうした手法は、あらゆる科学の分野に共通である。

例えば、医薬品や農薬の開発で、リード(化合物の)オプティマイゼーションによく使われるQSAR(quantitative structure-activity relationship)、すなわち化合物の定量的な構造活性相関は、こうした手法に他ならない。この他に、質量、IR、NMRなど各種のスペクトルデータに基づいて、分子式や部分構造を推定する方法がある。研究の歴史としてはこの機器分析データの解析のほうが古い。また、最近では従来のQSARで使われていたものより広範囲な化合物の構造的特徴量を用いて、薬理活性や発癌性、味など生物作用を推定しようという、いわゆる3DQSARという試みが盛んである。こうした手法も帰納支援の例である。

一方、演繹とは、確かとされる前提から出発して、厳密な推論を重ねて新しい知識(結論)を導くことである。コンピュータはもともと(プール)論理演算を行う機械である。これに経験的な知識を収めて、演繹的推論を行わせようというのが、知識(エキスパート)システムであり、しばしば人工知能とも俗称される。この種のシステムは、1980年代のいわゆる第5世代コンピュータ計画で脚光を浴びたが、その原型は、1960年代のJ. Lederbergらの質量スペクトルデータからの構造を予測するDENDRAL⁶⁾や、E.J. Coryらの有機合成探索システムLHASA⁷⁾、そして筆者らのそれも含めて、1970年代初めに開発された医学診断システムである⁸⁾。

このように、帰納や演繹を支援するコンピュータシステムはすでに数多く開発されている。ただし、それらはあくまでも思考の支援ツールであって、コンピュータが人間に代わって帰納や演繹を丸ごと実行するわけではない。それが不可能なのは、帰納にも演繹にも、コンピュータで実行不可能な発想的な要素が含まれているからである。しかし、コンピュータは例えば、有機合成化学者やドラッグ・デザイナーの発想を助けることができる。分子モデリングやグラフィックスはそうした発想支援ツールである。

計画に関する技法もすでによく知られている。それはサイバネックスや制御工学の主題であるが、線形計画法や動的計画法のような具体的な解法がさまざまな分野で応用されている。例えば、ファーマコキネティクスやダイナミクス・モデルを基礎にした薬剤の最適投与計画⁹⁾などは、こうした計画技法の適用例である。

公衆衛生は、「人々の健康や、医療サービスに違いをもたらす」ことを、基本ミッションとする。その意味で、学問として考えるなら、それらはすべて計画技法の応用であり、制御工学に帰せられるはずである。とくに、かぎられた資源を目標達成にどう最適に配分するかという問題が、姿を変えていろいろのところに存在する。こうした問題解決の技法も、すでに数多く開発されている。ただし公衆衛生の専門家は、このことをあまり認識していないようである。

4.3 研究分野に依存した基盤システムの例

4.3.1 分子モデリングとグラフィックス

現代の公衆衛生の基礎には、健康の科学あるいは、医学がある。こうした学問の基礎には分子生物学がある。分子生物学を支援する基礎的なシステムは、生体分子を扱うコンピュータシステムである。その中核となるのは、記号列解析的な遺伝情報解析や、計算化学などとも呼ばれる分子計算や分子のモデリング技法、分子の立体的な構造を表示するグラフィックスなどである。こうしたシステムは、単に分子生物学のみにとどまらず、広く、生物医学に関係する、諸学の基盤にもなっている。

例えば、創薬研究は、薬物と標的となる生体系との分子レベルの相互作用の知識を基盤として展開され

る。当然、薬物分子と生体系分子、およびそれらの相互作用の計算モデルとグラフィックスが重要になる¹⁰⁾。こうした手法は「分子設計」とも呼ばれる。だがこの言葉は、しばしば誤ったイメージと期待を専門外の人々に抱かせる。ドラッグ・デザインは数多くの仮設の生成と実験ないし計算による検証という試行錯誤のプロセスの積み重ねによってなされる。分子モデリングやグラフィックスの技法は、その多くの局面で用いられる基本的な道具である。そこで目的となるのは最終的な薬物分子だけではない。ある場合は、似たような活性を有する多数の化合物の共通部分の抽出であり、ある場合は X 線や NMR のデータに基づいて推定されている生体分子の構造の決定であり、またある場合はリガンドと結合するレセプターの調節部分や、活性部分の解析である。

この意味で分子グラフィックスやモデリング技法は、新しい薬物のイメージを追求してゆくときの基本ツールであるが、研究者の代わりに自動的にピカ新をデザインしてくれる打ち出の小槌ではない。すなわち、ある道筋をたどって、「望みのある分子を自動的に設計する」というような、人間の発想を完全に代替するコンピュータ支援技法は、現実には存在していない。おそらく遠い将来といえども、そのような「自ら創造的に思考するコンピュータ」は実現しえないと思われる。

4.3.2 地理情報システム

公衆衛生の研究を支援する基盤システムとして、もうひとつ重要なのは、地図、地理などを扱うコンピュータシステムであろう。こうしたシステムが使えるためには、まず数値化された地図（境界線図）、等高線図、海底の等高線図などが用意されていなければならない。また、地図を 2 次元の画像、あるいは立体図として扱うソフトと、そのためのハードウェアが必要である。ひと昔前だと、数値地図は、テープで提供されていたし、画像の表示には、専用の高価なマシンが必要であった。しかし、今では、日本の地図ならフロッピー・ディスクで入手でき、2 次元画像ならパソコンで十分あつかえる。立体画像でも、シリコン・グラフィックスの Indy や Indigo のような 3 次元マシンと AVS のようなソフトを使えば、容易に扱うことができる。

こうした、地図情報の上に、気象、環境、住宅、年

年齢人口密度、健康および医療指標など様々な情報や因子を重ねて見れば、それらの変量間の相関にも気づくことがあろう。このような健康地勢学 (Geohealth Tectonics) とでも呼ぶべき解析の支援システムは、公衆衛生研究の基本的な道具であらう。

4.3.3 ネットワーク環境における知識の高次構造化

生物医学分野では、昔からデータベースの開発に多大の努力が払われてきた。とくに生体分子に関連したデータベースは、ゲノム解析計画の進展で、データの蓄積が急ピッチで進んでいる。ネットワークが発達してきた最近では、これまで蓄積されたデータをさらに組み合わせ、高次な構造をもたせるような試みが出始めている。

例えば、われわれは、分子生物学の研究成果をうまく構造化し、コンピュータに蓄え、知識ベース化することを考えている。そして、この知識ベースを薬物作用を実験に先立って予測できるようなシミュレータに進化させることも考えている。これは、従来の計算手法 (分子モデリング) や (分子) グラフィックスに加えて、高次に構造化された知識ベースを構築し、双方をドッキングしようという試みである。

こうした高次に組織化された知識ベースの中核となるのは、疑いもなくタンパク質の構造と機能、とくに受容体のそれであろう。また細胞膜表面から細胞内にいたる信号伝達系、さらに DNA 上の遺伝子の転写や翻訳に至る信号伝達ネットワークであろう¹⁰⁾。こうしたタンパク質の構造と生体内での高次な機能の知識ベースや、それに関連した薬物やインヒビターなどのデータベースは生物医学および創薬研究の基盤となるであろう。

このように、単なるデータベースやその集合体ではなく、ある研究分野の基盤となるような有機的な知識やデータベースの統合化が、他の研究分野でもこれから続々と試みられるようになるであろう。この時、すでに述べたようなインターネットや WWW の技術が重要な役割を果たすと思われる (図 4 参照)。

4.4 課題別の支援システム

以上、分野に依存するシステムのうちでも、基盤的な性格の強いシステムの例を紹介した。このようなシステムのさらに上位にくるのが、ある課題に固有のアプリケーションである。ただし、こうしたアプリケーションは、課題に応じて用意しなければならないので、

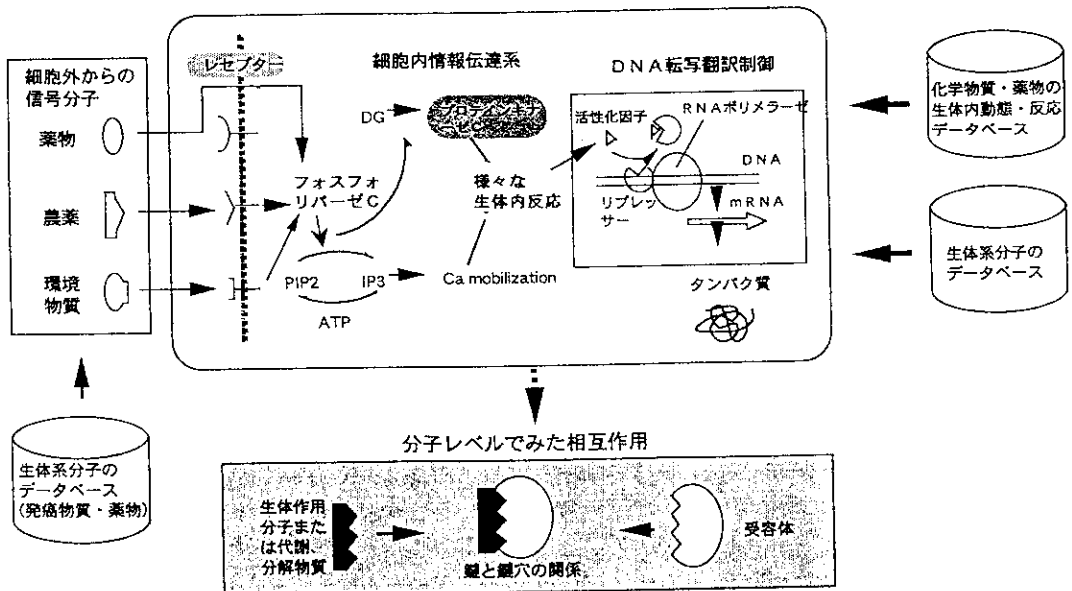


図 4 知識の高次構造化の例、細胞レベルでみた生体外からの化学物質と生体系分子 (受容体) との相互作用と信号伝達ネットワーク。

支援と言う立場からは、共通に取り上げることはできない。したがって、ここでは省略する。

5. これからの課題

5.1 情報学やコンピュータの専門家を擁するチームの必要性

以上、公衆衛生研究を支援するコンピュータシステムについて述べてきた。問題は、誰が、このようなシステムをデザインし、実際に実現するかである。すでに述べたように、わが国の、医学系や衛生学関係の研究機関には、コンピュータや情報学の専門家がいないのが普通である。こういう状況をできるだけ早く改めなければならないだろう。

パソコンの性能が十分でなかった1980年代の中頃までは、研究者がコンピュータを、使いこなすためには、コンピュータの専門家の助けが必要だった。日本では、コンピュータの利用そのものへの認識が遅れていたから、研究機関でこうした専門家のチームを置いたところは、ほとんどなかった。しかし、そのようなチームを置くのは、米国やヨーロッパの研究機関では、当たり前であった。実際に米国の NIH や CDC、ハイデルベルクの EMBL、ケンブリジの MRC 分子生物学研究所、あるいは先進的な医薬品や化学会社などには、こうしたコンピュータの専門家を擁するチームが存在していた³⁾。しかしその後パソコンの性能が向上すると、一般の研究者でも、金さえかければ、高機能で使いやすいアプリケーション・ソフトが揃えられるようになった。リレーショナル型のデータベースや統計解析パッケージ、グラフィックス・ソフト、画像処理ソフト、分子計算ソフト、遺伝子解析ソフトなどは、その典型例である。

しかし最近のように、広域ネットワークが普及してくると、様々なパソコンや複数の UNIX ワークステーションからインハウスの LAN を構成し、これをインターネットに接続し、さらに CD-ROM の多重検索やグループウェアのために Netware のようなパソコンネットワークを共存させ、これに画像処理とかグラフィックスのような特殊なシステムを付けくわえたような環境が必要になった。こうなると、研究者個人個人の努力では如何ともしがたく、組織的な対応が必要になってくる。すなわち、研究機関は全体として情報

と計算のための基盤環境を整備しなければならない状況に再び直面していると言える。

5.2 情報部門の新しい役割

新しい専門家の参加とともに、これまで図書や調査を専らにしてきた情報部門が脱皮しなければならないのも、明らかだろう。これからの情報部門に期待される新しい役割は図書、文献情報サービスだけでなく、そのまま研究のデータとしてつかえるファクト・データを提供することであろう。すでに例にあげたが、例えば医薬品の開発に携わっている研究者なら、従来の書誌、文献、情報に加えて、低分子化合物の X 線結晶解析（構造）データ、核酸やタンパク質の配列データや 3 次元構造データなどのファクト・データや、分子モデリングやシミュレーションのような分子計算環境を必要とする。

こうしたファクトデータの多くが、すでに国際的な協力の下でデータベース化され、インターネットで利用可能になっている。また、良質のソフトウェアも、ネットワークで入手できる。したがって情報基盤をつくるには、これまでのような図書、文献、調査の専門家だけでなく、ネットワークの専門家あるいは、そうした分野の知識をもつ研究者の協力が絶対に必要である。

5.3 国レベルのネットワーク環境の整備

以上、公衆衛生研究へのコンピュータ利用について述べてきたが、新しい環境づくりには、個々の研究機関の努力を越えた、国レベルの努力も必要とされる。そのひとつが、研究者用コンピュータネットワーク（情報ハイウェー）の構築である。現在、国立系の研究機関を結ぶ「省際研究情報ネットワーク構想」が進んでいる。これはまず筑波、東京、大阪を結ぶ通信幹線を敷き、国研あるいはそこ共同研究している機関をインターネットに接続する構想である。また、厚生省も厚生科学課が中心となって、傘下の研究機関をインターネットで結ぶ基幹ネットワーク構想を進めている。内容的にこの計画は、科技庁の省際ネットワークと同じ趣旨であり、両者は相補的な関係にある。しかし、公衆衛生の観点から見て重要な地方の衛生研究所や保健所、一般病院、薬局などとの接続はまだいずれ

の計画にもはっきりとは上っていない。いずれ上記の基幹ネットワークはそうした機関にも解放されると思うが、急ぐのであれば当面、別の可能性を探さねばならない。

もうひとつは、情報リソースやコンピューティングのためのセンターの構築である。それは、研究開発に関する各種の情報案内やデータベースの提供サービス、コンピュータ・パワーに加えて、プログラムやシ

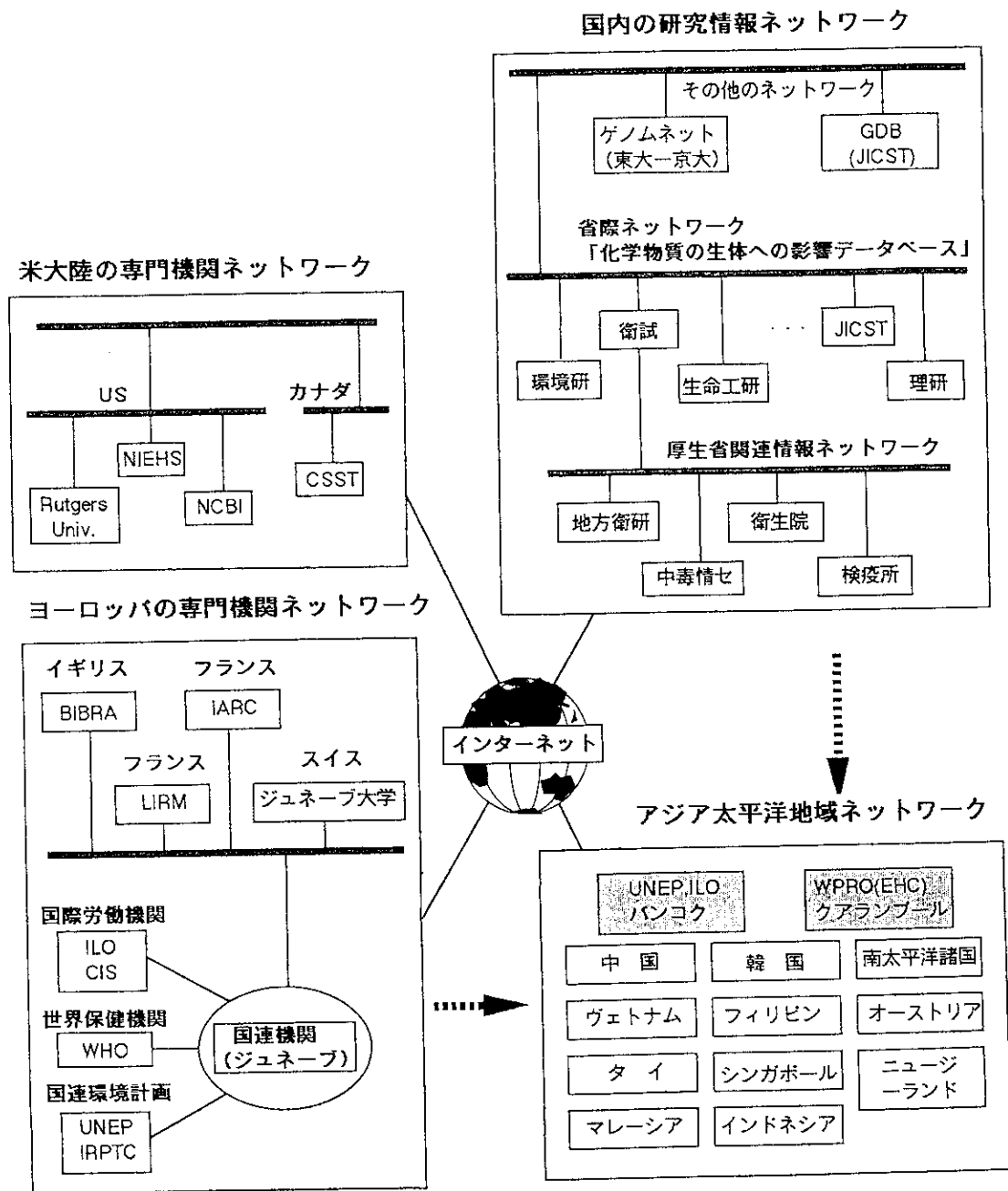


図5 化学物質安全性のための地球情報ネットワークの概念図
 アジア太平洋ネットワーク構築は国連機関の先導の下に日本が協力する

ミュレータなどリソースの活用および支援のための組織を有し、各研究機関の情報・計算基盤を互いに結びつける、積極的な役割を果たすようなセンターを創設することを意味している。

5.4 研究ネットワークの国際的な広がり

研究機関にとって、インターネットはすでに必要不可欠なインフラストラクチャーになっている。インターネットは、もともと国際的なものである。その巨大な情報空間にアクセスできるということは、われわれも、この情報空間にながしかの情報コンテンツをつけ加える義務を負うことになる、考えるべきであろう。

例えば、筆者らは、IPCS と呼ばれる、化学物質の安全性に関する国際的な協力事業に参加しているが、この事業の推進母体である、WHO, ILO/CIS, UNEP/IRPTC などジュネーブにある国際機関のグループと協力して、化学物質の安全性に関する世界中の主要な機関を、インターネットに接続するとともに、情報交換を行う GINC (Global Information Network for Chemicals) プロジェクトを進めている。この計画には、アジア地域のネットワークづくりが含まれており、われわれも協力することになっている。こうしたネットワークは、公衆衛生一般、あるいは他の目的にも利用できる、インフラストラクチャーである (図5参照)。

先に述べた省際ネットワーク計画でも、それをアジア地域に広げることが計画されている。また、この地域におけるネットワーク導入への関心も高い。経済だけでなく、公衆衛生研究においても、アジアとの連携は、より強くなるのは確実であろう。この意味でアジアネットワークも、まもなく有用性が認識されるようになるであろう。

5.5 継続的に環境を見直す必要性

情報技術は、現在爆発的とも言えるほどの発展期に突入している。インターネットやWWWとMOSAIC,あるいはオブジェクト指向データベース、エージェントなどは、これから本格的に発展し、さらに大きなインパクトを与えるであろう¹¹⁾。また、いわゆるグループウェアの実用化が進めば、研究部門と事務部門との文

書交換、事務処理が飛躍的に向上し、現在研究者を悩ませている雑用も大幅に軽減する可能性がある。したがって、研究支援のコンピュータシステムは、定期的に見直してゆかねばならない。

6. おわりに

公衆衛生研究を支援する情報と計算の基盤的な環境をどう整備すべきかについてのべた。一般に研究を支援するコンピュータ環境は、研究領域によらず、共通性が高い。それは、理想的に言えば、情報とデータの収集と蓄積、それらから知識を生成する普遍的思考支援システム、分野別思考支援システム、問題別のモデリングやシミュレーションなどを総合化した情報と計算環境、あるいは思考支援センターだと言えよう。

筆者は、たまたま前職においては、生物医学の研究支援システムを実際に開発した経験があり、また現在は、薬学や安全性の研究支援システムを開発するとともに、研究機関全体の基盤的な環境整備や、研究所間のネットワーク、地球レベルのネットワークの開発も試みている。しかし、この両者を本当に統合した経験はない。ただ、上記のようなデザイン・コンセプトはおそらく正しく、また情報技術が進歩してもあまり変わらないであろう。もちろん、実際のシステムの実現方法はコンピュータとその関連技術の進歩によっても大きく変化してゆくに違いない。

参考文献

- 1) 神沼二真: 医療革新とコンピュータ, 岩波, 1985
- 2) 企業のドラッグ・デザインへのコンピュータ利用に関しては, CBI NEWS, 1993年8月および9月号, 1994年3月号に特集がある (CBI研究会の連絡先は, 〒141 品川区北品川5-3-20 第2エーエスビル3F, TEL.03-5421-3598/FAX.03-5421-7085)
- 3) 神沼二真(編著): R&D コンピューティング要覧, サイエンスフォーラム, 1988
- 4) 神沼二真(編著): 生化学におけるコンピュータ利用, 丸善, 1989
- 5) 神沼二真, 鈴木 勇: 分子を描く, 啓学出版, 1998
- 6) Lindsay, R.K., et al: Applications of Artificial Intelligence for Organic Chemistry: The Dendral Project, McGraw-Hill, 1980
- 7) Cory, E.J., Wipke, W.T.: Science, Vol. 66: p. 178, 1969

- 8) 神沼二真, 倉科周介訳: 診療コンピュータシステム (Shortliffe, E.H.: Computer-Based Medical Consultations: MYCIN, Elsevier), 1976, 文光堂, とくに巻末の解説
- 9) Kaminuma, T.: A Mathematical Model of Gout for Optimal Treatment, Proc. of the Intern. Conf. on Cybernetics and Society, IEEE, Tokyo-Kyoto, 451-455, 1978
- 10) Kaminuma, T. et al.: A Knowledge and Data Base for Cell Signalling Network, IARC Workshop on Receptor Mediated Mechanisms in Carcinogenesis, Lyon, Nov. 14-16, 1994
- 11) 神沼二真: 第三の開国, インターネットの衝撃, 紀伊国屋, 1994