ミニ・シリーズ／CD-ROMによるデータベース

藤原 譲*

⑪

CD-ROM ミニ・シリーズ総括

Overview of the CD-ROM Series

FUJIWARA Yuzuru

mini-series / CD-ROM-aided Databases

[著者抄録] CD-ROM は、大容量、小型、大量複製、読み取り専用、価格性能比がよいなどの特徴の故に、画期的な情報媒体として注目されている。これまでの大型辞書、オンラインデータベースのように、従来の情報をそのまま CD 化する形から、出版情報、部品情報のように業務に密着したもの、さらには研究開発や、意志決定の支援システムにいたるまで種々の製品や試作品が多様な形で世に出されつつある。技術的に完成度の高い媒体ではあるが利用法との対応ではまだ多くの課題が残っており、さらに情報の流通体制に関連した問題も多い。これらについて、ミニ・シリーズ全体の総括とともに今後の展望をまとめた。

[著者付与キーワード] CD-ROM, ニューメディア, 大容量メモリー, 統合データベースシステム, 研究開発支援システム, 情報流通, コンパクトディスク

[Author Abstract] CD-ROM is remarked as an epoch-making medium because of its advantages such as large capacity, compact size, mass reproducibility, read only memory and cost performance ratio. Some of big dictionaries and online databases have been converted to CD-ROM versions so far, however, information of publication or machine parts are converted recently. Moreover various CD-ROM-aided products such as support system for R&D, decision making and so on are being turned out. Still there remain many problems on sophisticated utilization of CD-ROM and distributive machinery of information. Author reviews this mini-series and describes the prospects of development of CD-ROM.

[Keywords by Author] CD-ROM, new media, large capacity memory, integrated database system, R&D support system, information transfer, compact disk

* 筑波大学 電子情報工学系 (〒305 茨城県つくば市天王台 1-1-1) Tel. 0298(53)5530
Tsukuba University, Institute of Information Science
(1-1, Tennoudai 1-chome, Tsukuba-shi, Ibaraki-ken 305)

1. はじめに

新しい情報媒体 CD-ROM (Compact Disk-Read Only Memory) は日進月歩の情報技術の一つの成果であるが、その性能と技術の展開が非常に急速であったために、技術の可能性と利用の現状には大きな隔りがある。それにもかかわらず、CD-ROM に対する関心の大きさと実用化への試みは予想以上の多くの分野でなされており、その目的も内容も多種多様であることは、これまでのシリーズで紹介されてきた通りである。

CD-ROM の利用面から見た主な特長は、表 1 のようにまとめられる。

第一の特長は大容量で、冊子体や従来のデータベース (以下、DB と略記する) では実現困難であった、専門家が研究・開発または意志決定に必要とする、高度でかつ大量のデータが、容易に利用可能な形で収録できることである。現在普及している大型 DB のうち専門家向きのものは二次情報、しかも書誌情報が中心である。これはデータベース開発の時点では、記憶容量からの制約によって全情報の収録は困難であったためである。もちろん大型計算機の持ち得る補助メモリーは CD-ROM より遙かに大きい、大型 DB のカバーする領域も広いので、利用者あたりの使用できる容量は限定されることになる。

第二の特長は、文字通り直径 5 インチのコンパクトな形状であることで、省スペースの意義も大きい。

第三の特長は、一度マスターディスクを作製すると大量の複製も容易であり、印刷物と同じように多数の利用者が同一のデータを使用できる。すなわち、データが有効に活用されることである。

第四の特長は、ROM すなわち読み取り専用であることで、記憶内容の変化がないことが DB の管理システムを大幅に簡略化することに役立っている。特に、実際に CD-ROM が使われる際は、パーソナルコンピュータに接続されることが多いので、この点は重要である。また、読み取り専用ではデータの更新ができないので事象データ用には

表 1 CD-ROM の特長

1. 大容量 (540 MB)
2. 小型形状 (直径 12 cm)
3. 大量複製 (共用データ)
4. 読み取り専用 (管理容易)
5. 価格性能比がよい (記憶および流通)

不適ではないか、との質問を受けることが多い。CD-ROM 上のデータの更新は、バージョンを変える時でなければできないのは明らかであるが、実際にはそのことによる不都合は起こらない。それは、更新はフロッピーディスクまたは固定ディスク等の書き換え可能な媒体を利用すれば済むことによる。さらに実用的には、大量のデータの更新、例えば 540 MB のデータの僅か 1 % を更新するとしても 5.4 MB に達するので、利用者が個別更新することは得策でないばかりか非常に困難である。したがって、センターで正確な更新を行い、新しいバージョンの CD-ROM または他の補助メモリーの配布を受けるのが現実的である。もちろんオンラインとの併用も、即時性の要求の強いデータに対しては有効である。

第五の特長は、記憶容量についても通信についても価格性能比の良いことで、利用者数が一定数、例えば 20 名を超えればその効果は顕著となる。この条件は、通常の DB であれば十分に満たされている。また、価格性能比のみならず費用の上限が初めから決まっていることも、DB を利用する業務について管理面からの利点となる。出版情報、図書情報など多数の利用者への無料サービスをすることが期待されている図書館などでは、この点も大切な特長とされている。

2. 事例からみた現状

CD-ROM の特長を活かして、種々の実例がこのシリーズでも報告されている。それらを分類すると、辞典型、DB 型、定期型、統合型の 4 種に大

別できる。

辞典型は、グロリアの百科辞典に代表されるように大型の辞書でしかも何冊にも及ぶものは実際に使用する際に不便であり、その小型化と検索の簡便化、迅速化の要求に応える方策としてCD-ROM版は適したものとと言える。同じように、大量データの代名詞とも言える電話帳についても、NTTのタウンページがCD化されている。これらは、当初の目的に沿うことはもちろんであるが、内容へのアクセスの面で冊子体の見出し項目以外の点からの検索も可能である。タウンページの場合では名前は見出し項目となっているが、住所で、しかもそれが不完全であっても検索できることは、内容がすべて計算機可読のコードになっていることから当然とは言え、実用上、冊子体にはない新しい便宜を提供できることになる。

DB型は、DialogのERIC(教育用DB, Educational Resources Information Center), Excerpta MedicaのCD-ROM版であるEMBASE(医学用DB, Silver Platter Information Inc.), 英国ライブラリ協会のLISA(Library and Information Science Abstracts)などに代表されるように、オンラインで提供されているDBを頻繁に使用する利用者や、教育用、図書館用など、使用料金をあらかじめ確定しておきたい場合には、好都合な情報提供の方法として利用されている。もちろん、CD版は更新頻度においてオンラインよりは少なくなるので、最新データについてはオンラインとの併用が必要なので、ソフトウェアは利用者の便を考慮して両者共通化されている。

定期型は、データの量が多く、追加量が月単位でCD枚に相当するような場合で、BowkerのBooks In Print Plus(3か月ごとに更新)、日販のNOCS、特許情報機構の特許情報、などがすでに実用化されている。Books In Print PlusやNOCSは出版情報であり、利用者は主として書店であるから、使用法は検索のみのこともあるが、本の注文と結びつくことが多いので、電子発注のソフトが組み込まれた業務用としての対応ができている。

統合型というのは、研究開発や意志決定のように各種の情報を統合し、かつ高度な処理によらなければ目的に沿わないような場合、必要なデータを各種のソースから収集し収録してCD化するののでこのように名付けた。

有機合成研究開発支援システム(CORES: CD-ROM aided Organic Synthesis Research System)や高分子材料データベース(CAPDAS: CD-ROM Aided Polymer Data System)は、研究用DBの例である。後者の場合、情報源としては学術文献から収録してオンラインの事象DB化されているものをはじめ、カタログ、ハンドブック、辞典、ガイドブックなど、高分子の研究開発に実際使われているデータを収集・整理して、統合して使えるようになっている。このことにより、調査も含め、研究、開発の種々の段階で必要な情報が即座に得られる体制ができることになる。もちろん、現バージョンは量的にも質的にも高度な研究活動をすべて支援するには十分とは言えないが、現バージョンは容量に十分余裕があるので、適切なデータ、情報を補充することによって目標に達することが期待される。

また、これらのシステムの利用は検索中心ではなく、システムに整備してある処理プログラムや体系化された関連知識によって検索結果のデータ解析、新材料の設計、加工法の設計、新用途の開発など、研究開発における中心機能を直接支援できる可能性がある。このことはシステムの機能として、数値演算や文字照合のレベルだけではなく、帰納、学習、類推のように高度なものが必要となる。

特に、二次元、三次元のグラフ、画像、構造情報の表示、処理は、研究・開発に極めて重要であり、しかも黑白では不十分でカラーを要求されることも多いが、CD-ROMでは大型DBに比し、これらの実現も容易である。この点で、CD-ROMがパーソナルコンピュータやワークステーションに接続されて使われていることが、大型データベースシステムに比して有利な点の一つである。つまり、大型システムは多数ユーザの共用であること

と、消去可能型のデータ管理を行うことによる特殊な機能の導入が困難であるのに対し、個人利用が中心となる小型システムに利点のあることを示している。

3. CD-ROM 実用化への課題

以上に述べたように、情報の提供、流通、利用のどの面からみても、CD-ROM は画期的な新しい媒体であると言える。しかしながら、従来の媒体との違いも大きいために素晴らしい可能性を秘めており、技術的に非常に高い水準に達しており、経済的にもオーディオ市場の展開によって極めて有利な状態にあるにもかかわらず、その普及はいま一歩である。

そのためには、装置の普及、豊富な利用可能データ、および各種の用途に応じた高度なソフトウェアといった技術面での課題の解決も必要である。現在の情報媒体の主流である冊子体は、簡便性、速応性など優れた特長を持ち、その利用法も様々な角度から検討され、技術的には完成度が極

めて高く、オンライン DB もその意味で実用的に定着している。それに対し、CD-ROM は外部仕様の技術が優先し、利用法が十分追隨していないきらいがある。

広い分野での応用に対し、必要なデータ・情報とは何であるか、それをどのようなソースからどのように収集し、どのように整理・管理するか、それらを高度に活用するにはどのような方法、ソフトウェア機能が必要であるか、などを十分に検討しなければならない。

また、新しい媒体に適した情報の流通体制、著作権、DB 権などの行使、利用の際の十分な意識、慣習、機構なども確立されなければならない。

4. ま と め

多くの関係者の熱い期待を担って躍進しようとしている CD-ROM はやがて急速に展開され、そのことが研究開発のみならず生活、社会のあらゆる面で有用な、なくてはならない道具となる日も近いであろう。