

1T-10

アメリカ合衆国政府統計データへのインターネットアクセス  
カリフォルニア大学人口調査データ情報システム片山 博昭  
放射線影響研究所

## 1. はじめに

UCB図書館の政府情報サービスは国内外の統計データや文献、更に人口調査データに対する検索サービスを行ってきた。UCDATAはUCのために社会科学、公衆衛生の統計データベースを提供、援助してきた。最近まで、これらのデータベースの殆んどは大型コンピュータで用いられる磁気テープの形で配布されてきた。過去2年の間、政府情報サービス、及びUCDATAは主として人口調査局から、200枚以上ものCD-ROMを入手している。これら200枚のCD-ROMは800本の12インチの磁気テープ、或は100ギガバイトのデータに匹敵する。この2年間、CD-ROMでのデータの配布はこういった媒体を取り扱う為の適切な人員の増加や手段のないまま25%以上にも増加してきた。この状態は国内での政府公文書図書館、社会科学情報センターでも同様である。CD-ROMによる情報サービスを使うことで社会科学が直面した最初の問題は、データへのアクセスが単一のユーザー、単一のコンピュータでしか許されないということだった。もし数百ものCDが学内の数千ものユーザーから求められるかもしれないと考えると、（数台のPCが限られた時間使えるとして）解決策は見出せないだろう。PCを増やすことは限られた解決策であり、より広範囲な取り組みが要求される。本稿では、我々がどのように複数のユーザーに対し、ネットワークを通じての使用を可能にしているかを報告する。

## 2. CD-ROMへのネットワークアクセス

ほとんどの図書館では、増えつづけるCD-ROMによる情報への対策を模索していたが、ネットワークを利用したCD-ROMへのアクセスや利用はこれまでのところ、ほんのわずかの組織で行われてきただけであった。CD-ROMをネットワーク上で使用する場合に、望まれる機能として次のものが挙げられる。(1) 地域的な制約を受けることなく、キャンパス内、或は組織内のいずれのポイントからでもアクセスできる。(2) CD-ROMデータベースは複数のサーバーから構成する事が出来る。これは、データの中身に精通した管理者をCD-ROM情報データバンクに取り込む事が出来る。(3) アクセスはPC、Macintoshからだけというのではなく、様々なハードウェアやOSから可能である。(4) サーバーは複数のユーザーからの同時アクセスをサポートしなければならない。頻繁に使われるデータベースは速度の遅いCD-ROMから速度の早い磁気ディスクに移動が可能である。(5) データベースはCD-ROMの追加、サーバーの追加があったとしても、費用のあまりかからない方法で拡張が容易である。

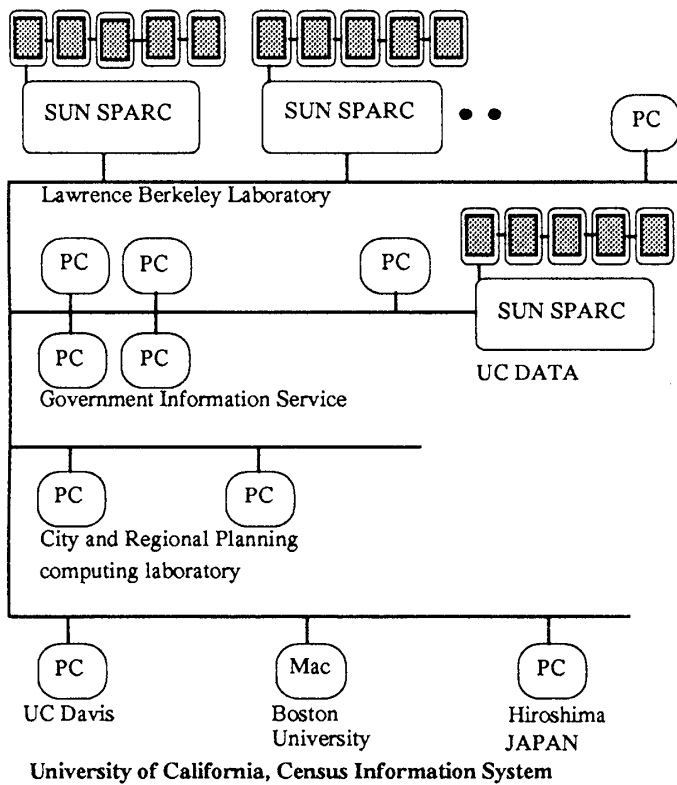
---

Internet Access to Public U.S. Federal Government Statistics

Hiroaki Katayama

Radiation Effects Research Foundation

5-2, Hijiyama-park, Minami-ku, Hiroshima 730, Japan



我々はこれらの機能を満足させる為に、TCP/IPをプロトコルの標準とし、様々なオペレーティングシステムで使用されているネットワークファイルシステム(NFS)を解決に選んだ。NFSは地理的に点在するCD-ROMデータベースを結び付け、CD-ROMが接続されたマシンをサーバーとして働かせることができる。図は我々のシステムの構成を示す。7台のサン製スパークステーションがサーバーとして用いられ、そのうち6台はUCBから半マイル離れた丘の上にあるLBLに、残りの1台はキャンパスから2ブロック離れたUC DATAの建物にある。これらのワークステーションはインターネットに接続され、それぞれのワークステーションには5台のパイオニア製CD-ROMチェンジャー(DRM-600、604X)が接続されている。それぞれのCD-ROM

チェンジャーには6枚のCDが装着出来る。また、キャンパス或はLBLのネットワークに接続され、NFSを使用しているPCはあたかもCD-ROMが直接接続されているかのように、アクセス出来る。これらの大量のCDを管理する為、プロジェクトはPCをベースにしたメニュー志向のフロントエンドシステムを開発した。このシステムは階層的に作られたバッチファイルで構成され、メニューからの項目の選択によりユーザーを目的のCD-ROMディスクに導き、CD-ROM上、或はディスク上のソフトウェアを起動させる。ユーザーにとっての多少高度なデータへのアプローチは人口調査局が提供するEXTRACTプログラムのような汎用目的の為のソフトウェアを使用する事であり、これらのプログラムはレコードグループの選択だけでなく、データベースのフィールドに対してもオンラインでの選択を可能としている。出力様式の変更や他のソフトウェアのフォーマット(例えば、dBASE, SAS等)への出力もサポートしている。ユーザーにとってむずかしいがより効果的な使い方はdBASE, SASのような汎用的なソフトウェアを使用する事である。

### 3. 最後に

この論文ではCD-ROMで提供された政府統計データへのインターネットアクセス及び組織内での処理に必要な機器類、及び機能について述べてきた。システムは約1年稼動してきたが、現在までに100を越すクライアントにより利用され、更に新しいクライアントを加えつつある。我々はいま、サーバー、及びCD-ROMチェンジャーの負荷を検討している。UNIXワークステーションに接続されたCD-ROMジュークボックスを用いる技術はデータ保管及びアクセスの両面に柔軟性を提供し、システムの拡張もそれぞれのサーバーでの最小限の経費のみで可能としている。最後に、LBLのCEDRプロジェクト、UC図書館政府情報サービス及びUC DATAの方々に深く感謝します。