

CD-ROM 長野県植物誌資料集テスト版始末

金井 弘夫 (184-0013 小金井市前原町 5-8-7)

Hiroo Kanai : CD-ROM Materia Florae Naganoensis (Pilot version),
an Introduction

Nearly 430 thousand records of vascular plants indigenous to Nagano Prefecture, central Japan, including of specimens and of field notes, are enrolled. They are retrievable on any one of 23 items such as plant name, place name, collector, date, sheet number, position code etc. Distribution map, horizontal and vertical, of any species is available together with coast lines, prefecture boundaries, contours at choice. Mesh or coordinates can be overlayed. Full spelling of plant name, Japanese and scientific, can be displayed consulting with a built-in dictionary. Distribution map drawing is applicable to any part of Japan when data set is provided in required format.

§ 発端

はじめのうち、資料集のCD-ROMを作ることについて、大したこととは思わなかった。従来知られているフロラのROMは、リストや文章がテキストで入っていて、せいぜい簡易検索ソフトがついていれば上出来という程度だが、我々のものはインデックス検索やら分布図作図やらの機能がついた、ちょっと類を見ないものである。しかしすでに「CD-ROM 日本地名索引」を作った経験があり、同じ会社に発注するのだから、地名と植物という違いがあるにしても、検索の手法も分布図作図の方法も、多少の手直しでよいつもりでいた。43万件というレコード量も、ROMにとってはものの数ではない。受注会社のインテージでも同じ見方をしていた。

ROMを作るとき、印刷の校正刷りと同じで、何度か試作品を作って検討して行く。出来上がりをVersion 1.0と呼び、試作段

階ではVersion 0.1から始めてゆく。Version 0.2まで来たとき、これはどうもおかしいと思いはじめた。私はリスト画面では表操作の感覚を要求していた。つまりカナムグラを検索すれば、その植物名を持つレコードがまとまって見られることは勿論だが、その前後のカナビキソウやカニクサも、頁を前後させれば見られるつもりなのに、カナムグラだけしか見られず、前後には何もないのである。分布図でも同じで、カナムグラの分布図画面からとなりの種類の分布図へ移るには、あらためて検索をやり直さねばならなかった。

文献検索の手法は、目的のデータを持つレコードだけを切り出してまとめ(SELECT)、その中から更に切り出してまとめるという方法で絞り込む。したがってヒットしなかったレコードは切り捨てられてしまっていて、見ることはできない。地名索引CD-ROMはこのやり方だったのだ。つまり私が要求して

いる検索法(FIND)は、地名索引 CD-ROM の設計思想とは根本的に違っていたのである。

§ 出直し

この段階で「客の要求を理解しないで進めるとはなにごとだ!」と強い抗議をした。今から考えると前例があったのが災いして、会社側は「あれでいいんだ」と思い、私の方はどんな検索方法をとっているのかとか、どんな選択肢があるのかも知らずに、自分の思い込みがむこうに通じているつもりになっていたのだ。このスレ違いは、その後いろいろな場面でおこった。要するに同じ単語を使っても、エンジニアと普通人(とも言えない植物人)の常識の違いで、同床異夢に気付かないのである。

とにかくソフトの設計を始めからやり直さねばならなくなった。会社にすれば既製品の小手先の手直しでよかったつもりが、一から出直しである。そうすると、プログラムの各段階で、社内でのチェックを厳重にやらねばならない。ゴールデンウィーク明けには出来ると言っていたのが8月末まで延びたのは、そういう事情である。契約は既に済んでおり、金額も確定していて、いまさら変更できないし、されてはこちらが困る。その上、発注者からは細かな手直しが後から後から入ってくる。こちらが「細かな」と思っている、プログラムを書く側では大ごとな場合もある。担当者は眠れぬ思いだったろう。はじめからそういう設計思想とわかっていたら、見積金額の桁が違っていたかも知れないのである。われわれのような貧乏集団がこれだけの機能をもつ CD-ROM を作ることができたのは、当初の互いの思い違いの産物で、僥幸と言ってよい。それにしてもインテージが、こちらの

出す要求や手直しに一々応えてくれたことには感謝のほかはない。その最たるものが、清水氏がこだわった「年代検索」である。

この機能のおかげで、たとえばヘラオオバコの 1930 年以前の記録と以後の記録を、別個なリストや分布図に作ることができる。この機能を使った分布図は、早くも近田文弘氏によって「小泉標本と温故知新」(国立科学博物館ニュース No. 402, 2002 年 10 月)と題して発表されている。

§ リスト

経費節約のための工夫があちこちに見られる。リスト画面のスクロールが不自由なことは、誰もがすぐ気がつくことだろう。リスト操作の不便さについては、私はあまり気にしていない。なぜならば、レコードの外部へのコピーが可能なので、不満な人は自分の使っている表ソフトへ取り込めばよいからである。このままでも適宜工夫をすれば、見たいレコードやデータを見ることはできるのだから、不便で困るというほどのものではあるまい。

検索は、すべての項目についてインデックス検索ができる。項目を組み合わせたインデックス検索はできない代わりに、同じインデックス項目内なら、逐次検索ではあるが、複数項目にわたる二次的な検索を可能にもらった。レコードはもともと分類順に配列されているので、リストの出力は多くの場合分類順に行える。

§ 分布図

分布図作図機能はこの ROM の最大の特徴である。リストから直ちに、その和名をもつ植物の分布図を作図できる。垂直分布図も同時に作図される。

地図データは海岸線、都道府県界、等高

線(200, 1500, 2500m)を任意選択した上、日本全国どこでも(北方四島を含む)描ける。この他に長野県内の500m等高線と市町村界が描けるようにしてある。これらの線のデータは、20万図から読み取ったものである。現在では国土地理院の数値地図があるので、これを利用するのがあたり前だが、いくらかでも費用を節約するため、むかし私が作ったデータを提供した。日本地名索引CD-ROMの地図は、簡略すぎて私にはなじめなかったのも一つの理由である。つけ加えれば、市町村界や500m線以外は数値地図以前の作品で、おそらく我が国最初の全国的数値地図データなのである。等高線はせっかく作ったものが、機種更新の際メーカー側のデータ変換がうまく行かず失われたので、また作り直したといういきさつがある。

分布図作図関係では、私の作図ソフトKLIPSのノウハウのすべてを投入した。このROMができればKLIPSの出番はなくなるし、普及版となると数値地図を使わざるをえないだろう。だから私はこのテスト版ROMが、1976年生まれのKLIPSの墓標だと思っている。普及版ができた後でも、私自身の仕事にはこのテスト版を使い続けるだろう。

地図の線が実線しか使えないのも、節約の一つである。点線を使うとプログラムが複雑になり、費用がかさむのだそう。その結果、白黒印刷の効果がそがれてしまった。画面はカラー表示だから見ばえはよいが、印刷発表するためには白黒のプリントが必要である。ところがカラー画面を白黒プリントすると、色が薄くて見すばらしい。そこで画面を自動的に白黒に変換する機能

を付けてもらった。ところがその後で点線機能を削除したのだから、点線でメリハリをつけるつもりだった白黒印刷が実線ばかりで濃くなりすぎて、分布図がかえって見にくくなってしまった。色の種類や濃さを工夫してしのぐほかはないだろう。

地図の拡大縮小率の設定画面では、標尺を動かしても拡大率の数字は変化しない。一度確定してからあらためて設定画面を呼び出すと、新しい数値になっている。数字も標尺と共に変化するようにしろと要求したら、この設定画面は無料で利用できるソフトを使っており、数値の変わる画面を使うと経費がふえるのだそう。こんな具合に、われわれの知らないところで経費削減の努力がされているのである。

§ マニュアル

マニュアル制作費を削ってプログラム制作に回したため、マニュアルは私が書く羽目になった。マニュアルというものが、わかりにくい文章の見本であることは周知の事実である。それをエンジニアでもない私が書くのだから、わかりやすくなる筈がない。思い切って、私の前々からの主張である「マニュアルを読む前に使っしまえ」という一章を設けた。初めてのソフトを使うとき、スタートはさせたものの、途中でわからなくて止めたくなっても、止め方がどこに書いてあるのかわからない。面倒になって強制終了させたり、ひどい時には電源スイッチを切ってしまうということをする。だから、「起動」の直ぐ後に「終了」の方法を書いておいた。一度止め方がわかれば、二度目からはあせらずに済むから...。マニュアル本体は、やはり何やらわからないものになってしまった。けれどもひとま

ず使ったあとでマニュアルを読んでもらえば、使い方は一応書いておいたつもりである... と言うよりほかはない。

§ 和名学名対照表

この ROM のもう一つの特徴は、和名学名対照表である。長野県植物誌データベースはわが国最古の起源を持つものなので（もう知る人は少ないだろうが、本当なのです！）、昔々の IBM80 桁カード時代の文字数や文字種の制約が尾を引いている。植物和名は 10 字、属名は 10 字、種小名は 6 字、最小ランク小名は 6 字という貧弱さである。それでも当初の 7, 6, 6, 6 字から拡張したのである。その上和名に用いるカナ綴りは、ソートの便宜上「正しい」ものとはいえない。この「和名」から「正しい」和名や学名のフルスペル（著者名も）を知るために、植物誌リストと対照表は常に連携しており、切り換えればそれらを見ることができばかりでなく、異名もある程度参照できる。

ある地域のリストを印刷出力するとき「フロラリスト印刷」の機能を選べば、植物誌データからではなく、和名学名対照表から植物名の全綴りを出力することができる。こういう機能のついたソフトは、これ迄にないものと思う。

対照表は長野県植物誌の見解を優先しているが、「この見解のままで全国的に通用するのか？」と考えさせられるものもある。無効名と思われるものも含まれている。普及版となったら、執筆者の見解があらためて問われることになるだろう。もはや「地方の一学徒」と、へり下っているわけにはゆかなくなるのである。対照表は先行き日本植物全体に対応し、学名や原典も表示できるような仕様にしているが、まだ中身が整っ

ていない。

§ 外部データ利用

外部データ利用機能も、この ROM の特色の一つである。この ROM は日本のどの地域の地図も出力できるので、県外の人が自分のデータを使って自分の地域の分布図を作る際に利用できる。それが外部データ利用機能である。外からのデータを読むについては、相応の注文がついているが、特にむつかしいものではない。だから植物以外のどんなトピックについても、利用可能なのである。県内のデータでも、たとえばある属の分布図を作りたいならば、関係レコードを一旦外部へ取り出し、整形してから外部データ利用機能を使って分布図を描くことが可能である。

§ デモ

デモについては清水建美さんは「とにかく出来たことを皆さんにお披露目して、ついでにメディアにも来てもらおう」という程度で、アッサリしたものを考えていた。彼のノートパソコンに入れたソフトを見ればよく、それも私に操作してもらうつもりだったようだ。彼はその段階では、この ROM の操作にあまり慣れていなかった。実際まだ制作途中で、予期しない反応が出たり、ファイルがこわれたりというようなことも起こっていて、それがソフトの不備のせいなのか、操作ミスの結果なのかはつきりしなかった。第一、ROM の最終版がいつできるか、まだわからなかった。デモの日時を決めてしまえば、下手をすると前日にならないと ROM が手に入らないかも知れないし、手に入らなければ試作版でやらねばならない。彼はそういう場合には、当日朝会場で、私の持つ最新試作版（Version

0.3.1 まで行っていた)をインストールすればよいと思っていた。

私は「他人のマシンを使ってやる自信はない。清水さんがすべてやってもらいたい。少々失敗しても内部の人相手なら何とかできるので、それで自信をつけてからメディア相手のデモをしたらどうか。メディア相手のデモに、当日朝インストールするなどは考えない方がよい。」とアドバイスした。私の方が彼よりは操作に慣れていることは確かかなのだが、私の使っているマシンは「親指シフト」という富士通特有のキーボードで、日本語入力には大変優れているが、今やほとんど絶滅しており、一般に使われているローマ字変換キーボードを使う自信はなかった。下手をすると余計なコントロールキーを押してしまって、マシンが異常行動でフリーズしかねないのである。当日になってインストールなどということは、恐ろしくてとてもできない。実際ほかの機会に、そういう場面に出くわしたことがある。言うことをきかないマシンを前に「...こうなることになっている筈なのです...」などという説明はやりたくない。

これに加えて、CRT ディスプレイを使ったデモをやることを提案した。CRT ディスプレイは、パソコンの画面をそのまま大型スクリーンに映写する装置である。その理由は、「委員の皆さんは永年にわたってデータの作成や訂正に時間を費やしてきたのに、それがどのような果実に稔るのかは、全くつんば棧敷に置かれたままである。ディスク一枚見せられた上、小さなノートパソコンをのぞき込みながらでは、達成感を共有してもらえないだろう」というのが私の主張だった。彼はそんな装置があることは初

耳だったようだが、とにかくディスプレイを探して借り出し、あらかじめ会場で映写テストをやって、電源やコードやコネクタの接続が支障ないことを確かめ、それからデモのすべてを自分でやれるように練習した。さいわい、ROM の完成版(Version 1.0)は間に合った。練習には一週間近くかけたらしい。万一何かの理由でパソコンがダウンした時に備えて、予備のパソコンも用意した。その自信の上に立って、内部デモとメディア発表を同時に行うことにしたのである。

2002 年 8 月 27 日のデモをご覧になった方は、新しい ROM の機能について、一通りの理解を得られたことと思う。ノートパソコンの小さな画面を覗き込みながらだったら、「見えなくて何だかわからなかった」と言う人の方が多かったに違いない。何も知らないメディアの記者達ならなおさらである。

当日は皇室関係者の出席する「雷鳥と山岳環境に関する国際シンポジウム」があり、「斎藤記念国際音楽フェスティバル」が開催中で、しかも長野県知事選挙の終盤だったのに、信毎では結構大きな記事で扱って貰えたし、NHK 松本は朝昼晩三回も報道してくれたという。清水さんの周到な準備が、効果を上げたのだと言ってよかろう。

§ 普及版へ

この ROM はテスト版で、ごく少数が作られたにすぎず、かつ非売品である。能力としては日本のどの地域にでも適用できるのだから、誰でも利用でき、かつ誰でも入手できるようにするのがよい。そうでなければ宝の持ち腐れである。この意味は、外部データ利用としてではなく、はじめからソ

フトの中に組み込まれたデータベースとして構築できれば、日本全体のフロラ研究に、より有用な武器となるに違いないということである。その様態は少なくとも二通りある。

一つは長野県植物誌と同様な、地域植物誌としてのデータベースである。これはそれぞれの地域研究に直接役に立つものである。もう一つは、たとえば牧野標本館とか薬用植物標本室とか旧制大学の標本室のような、地域に偏らない全国的なコレクションのデータベースである。これは日本全体の植物のあり様を知る上でどうしても必要なものであり、地方植物誌に無関心な人でも利用価値がある。ひとまず長野県版として普及版を作り、その次にはどこかの県のフロラに参加してもらって、その版では長野県データは全国データの方にマージさせて行くという具合にして、次第に日本全体の植物資料集にして行けないものだろうか、私は考えている。だがそのためには、データベースのいろいろな点について、手直しすべきところがたくさんある。

長野県植物誌資料集は、これ迄皆さんが多大な時間を割いてデータの訂正を行ってきた。にもかかわらず、誤りはまだまだたくさんある。ROM ができればいろいろな検索が簡単にでき、かつ各人が自由に操作することになるので、エラーの発見はこれ迄以上に多くなるに違いない。たとえある時期までにエラーを訂正し尽くしたとしても、その後から入ってくるレコードについては、また同じ努力をせねばならないので、終わりというものが見えて来ない。追加レコードはこれからも続々と入って来るのである。

もう一つの問題は、この ROM を県外の

方々に広く使ってもらうためには、データ仕様を見直す必要があるということである。この点では、わが国最古の植物データベースは最も不利である。後発のデータベースほど、仕様の自由度は高いから、そういう人達が共通に参加しようと思うためには、われわれの仕様を変更するほかはない。端的に言えば項目の新設と項目幅の拡張である。たとえば「日」の項目は、増設せざるを得ないだろう。そうすると、他の地域のレコードには日のデータがあるのに、長野県のそれにはないというような「見劣りする」結果になる。今更 40 万件を調べなおして「日」のデータを入れることなど不可能である。産地項目についても、項目幅を拡げれば同じ現象が見られる。

位置座標も、この頃はやりの三次メッシュや経緯度も受入れるようにする。これはテスト版ですでに可能になっている。そうすればわれわれの使って来た Locality Index は、粗雑で見るに耐えないと誤解する人もいるだろう。言っておくけれど、これは「誤解」の最たるものである。でも、こういうことは常に先行者の宿命であると思う方がよくないだろうか？

いずれにせよ、「これが最終で間違いがない」というものは考えにくいし、そうしようとするればノイローゼになるだろう。教科書のように「誤りが無い」「全てが整った」ものにしようと神経を尖らすよりは、「エラーがあれば、次のバージョンで直そう」「新旧仕様が混ざってもかまわない」というつもりの方が、ピリピリしないで済む。そのためにも、よその地域の人達がときどき参入してきて、新バージョンを作るチャンスがある方がよいだろう。

§ 記録用データベースと検索用データベース

もともとの二つは別物と思ってはいたが、ROM を作る過程でやはり別物だとの印象を強くした。植物誌資料集の初期の段階で「産地名を統一して調査票に記すのか、元のラベルのままでよいのか」が問題になったとき、私は「元の記述のままでよい」と言った覚えがある。これは「データベースの記述を統一しなくてよい」「原資料に忠実なのがよい」ということではない。入力するとき統一記述に書き直すとなると、各人の負担も所要時間も莫大になり、しかも人によって解釈の違いがあるので、「統一」するメリットはない。それよりも、頭も時間も使わずに、少しでも数をかせぐ方がよいと考えたからである。そうやって作られた「不統一な」データは、データになってしまえばソーティングを行って比較が容易になるし、必要なら一括して書き直すことは造作ないからである。実際にこうしてたくさんさんのデータが「統一」されている。統一し過ぎて文句の種になったこともある。その理由の多くは、「元の記録に忠実でない」ためだった。

近頃の表ソフトでは、「項目幅」というものを気にしないでデータを入れることができるようだ。「元の記録に忠実に」という立場なら、こうでなくてはならない。しかし今回のように「検索」それもインデクス検索を命とするデータベースでは、検索対象となるデータが一定の位置に納まっていなければ機能しない。項目幅を意識しなければ、検索対象データを選択して入力するという注意は払われまいだろう。どちらの方針を選ぶかは、これ迄議論されたこととはな

い。第一、そういう二つの行き方があることが意識されていなかったのである。

他地域の利用を前提とした仕様改訂に当たっては、私がお預かりしている限りは「検索用データベース」の立場をとるつもりである。他地域から参入する人達は、すでに自分らのデータベースを使っていて、記録用としては自足しているはずである。参入するのは、この ROM の作図機能や対照表機能にメリットを認めるからに違いない。この ROM に参加する場合には、ROM のソフトに都合のよい仕様に変更することを承知してもらわねばならない。

前述した新旧仕様によるデータの不揃いは、検索という立場からは大した問題ではないと考える。例えば採集番号（現在は 4 桁）の幅を拡げれば、新データと旧データで検索結果は異なるが、共通部分を意識して検索を行えば、一回でできなくても少数回のトライで目的を達するだろう。

§ これから？

長野県植物研究会誌 No. 26(1993) と No. 27(1994) に「長野県植物誌データベースことはじめ」と題して書かせてもらった。あの時期に私がなぜああいう文を書いたのかを、察してもらえればありがたい。あの文章は、植物誌の出版が終わったときのために用意してあったのである。20 年を超えるおつきあいの末、ひとまず終着点が見えたのでホッとしている。ここまでよくぞ保ったものだと思う。でもこれで終るわけではない。自然誌の研究は流行のトピックと違って、結果を出せばあとはなにもしなくてよいというものではなく、資料やデータベースの増強はこれからも続けねばならない。標本というものはしまっておけばよい

骨董品ではなく、使うものである。使うためには常に「整理」と、現状に対応するための「補充」をせねばならない。「整理」というと「取り出すとき便利のように」という面しか意識されないが、実は「取り出したあとで元に戻す」エネルギーの方がはるかに大変なのである。カラスライドを仕事に使う人なら、理解してもらえるだろう。放っておいたらエントロピーの増大で、何がどこにあるのやわからなくなってしまう。そういう面倒を、世代を超えて常に見ていなければならない。それには受皿となる施設と組織がしっかり出来ている必要がある。長野県は昔から有名な教育県であるにもかかわらず、残念ながらまだそういうものがない。今回のCD-ROMをきっかけにして、一つはデータベースの全国的展開を図る一方、他方では標本資料のメンテナンス体制を確立する方向への動きを期待したい。

§ 追記

この文章の主体は、ROMの発表デモの直後に書いたものである。続いて普及版にとりかかるべく、2002年9月13日に、援助していただいたABOC社でデモと打合せを行うことになり、12時に大船駅で待ち合わせた。清水氏は時間になっても現れなかった。12時30分、「清水氏に電話するように」との構内放送があった。彼の携帯電話には応答がなく、松本の自宅は通話中だった。大船駅には改札口は一か所しかない。彼は場所を間違えたのだと思った。ABOC社でのデモと普及版に関する提案は、食事をしながら打ち合わせるつもりだったが、一人でやる他はない。急いで腹ごしらえをして、もう一度携帯にかけたが、やはりつながらなかった。再び松本へかけたら奥さんが出

て、清水氏が倒れたことを知らされた。この接触がなかったら、私にとって彼は当分行方不明となったに違いない。奥さんが病院に駆けつけたので、松本の家は無人になってしまったから…。奇妙なことにその時の印象は、「清水さんは来られない」ということだけだった。彼が「倒れる」ことなど、私の辞書にはなかったのだ。ABOC社でのデモや話し合いは淡々と進められた。夜遅くなって日大病院に行ったらご家族が揃っており、詳細を知らされた。彼は地下鉄大手町駅で頭痛に襲われ、駅事務室にたどりついた直後に意識を失い、救急車で病院に運ばれてから自宅へ知らせが入り、奥さんが当日の予定を知っていたため、大船駅へ連絡してくれたのである。脳内出血だった。

これ迄の事業は、清水建美氏の強力かつ周到なリーダーシップによって進められて来た。ここしばらくの間、清水氏は休養せざるを得ないだろう。普及版への動きがそのために停滞することは、彼の本意ではあるまい。同志が協力して、態勢を立て直すことを期待する。§デモの章は§追記と共に書き加えたものである。