

長野県植物誌データベースことはじめ (その2)

金井 弘夫

Kanai, Hiroo: Genesis of the Flora of Nagano Database (2)

植物誌編纂の過程で、研究会には谷間の時期があったと思いますが、私にとっては1986年ころが、データ消滅の危機でした。TOSBACはとうの昔に製造もメンテも止まっており、次の機種への転換も科研費が何年もとれないため果たせませんでした。気密性の悪い機械ですので、ヘッドのごみ詰りやICの劣化による故障が続出しました。東芝のサービスセンターでも、もうこの機種のサービスはしないことにしていたようですが、山下さんという方が、電話をするといつもすぐに直しに来てくれました。部品の交換などはもうストックがなく、あちこちへ電話して探していたようです。あとで気付いたのですが、最後の2年ほどは請求書が来ませんでした。こういう修理は一回来るだけで数万円、部品を交換すれば更に十数万円とられますので、私の研究費では賄えなかったはずですが、山下さんは仕事を離れてサービスしてくれていたのです。近頃になってこういう事情にやっと気付いたのですが、たいへんありがたいことでした。一度、ヘッドの劣化でOSが壊れたときには、故井上浩さんが会計課長と交渉して、ハードの修理費数十万円をもらってくれました。こわれたOSのロードモジュールを、プログラムテープから探し出して読み込む作業は、自分でやらざるを得ませんでした。要するにもう寿命だったのです。分布図作図は当初プロッタを使っていました。当時のプリンタはタイプライタに近いものでした。私の機械はインパクトプリンタといい、幅広のゴムベルトに一文字ずつ文字を鋳込んだ細い板を歯の歯のように植え、これをプラテンの前でグルグル廻し、必要な文字が来たときにハンマーで叩くという方式でした。ディスプレイはありませんでした。こんなものでは今日のようにプリンタできれいな図を出力することなど考えつきません。それにプロッタなら図はいくら

でも自由に作図できます。しかしプロッタ作図は、時間がかかること、多数の図を連続して出力できないこと、プロッタのメンテが面倒なこと、誰でもプロッタを装備できるわけではないので普遍性がないことが欠点でした。そこでプリンタをワイヤドットタイプに更新したのを機会に、TOSBACもプリンタ作図に重点を移しました。

私がワープロを使いはじめたのはかなり早く、キーボード配列のユニークさにひかれて、富士通のOASYSを発売以来使っています。普通植物調査のデータ作りにも、これを利用していました。しかしワープロでいくらかいいにデータリストを作っても、分布図作図のデータにはなりません。TOSBACで分布図を作るには、データカードの形にしなければならないのです。ワープロで作ったデータフロッピーを、外注して大型機でカードにするという、まわりくどいくかつ高価なやり方をとらざるを得ませんでした。まだワープロデータをテキストデータに変換するソフトはなかったのです。この点、長野県植物誌データは、はじめからTOSBACで扱ったので楽でした。FDを使えばことは簡単なのですが、TOSBACはまだFDを想定したシステムではありませんでした。私は今でもデータ作りはワープロの方が楽だと思っていますので、互換性の面からTOSBACの次期機種は富士通製品にするつもりでした。練習のためFACOM9450の下取り品を手に入れ、OASYSとの互換ソフトをつけて、プリンタ分布図を作るようにしました。こちらはBASICを用いました。というのは、パソコンが個人装備となることは明らかでしたし、その主要言語がBASICであることもはや必然だったからです。ですから一ころの私のプリンタ分布図は、TOSBACとFORTRANによるものと、FACOMとBASICによるものとがあります。

もっとも、私が BASIC を使えるようになった頃には、この言語は便利さのあまり方言が分化し過ぎ、共通性がなくなっていました。そのうえデータベースソフトの発達で、プログラムを作る必要がなくなり、近頃では「いま頃 BASIC ですか」と言われることがあります。

こんな下準備のうえ、富士通機を導入するつもりになり、データを移植するために、フロッピー化を東芝に依頼したところ、驚くべき返事が来ました。いわく「1.こんな古い機械を更新しなかったのは、ユーザーの怠慢である。2.当社では TOSBAC の OS を知る者はもはやおらず、したがってその下で働くデータディスクを扱える者はいない。3.ディスクドライブもこの型のものは無くなっているの、当社では読み取ることができない。4.対策としては、全部プリントアウトした上、手で入力しなおすのが最も良い。」

コンピュータのメーカーが、自社の製品に対して言うセリフとは思われません。日本の工業化がはじまった昭和のはじめ頃、「舶来品は何十年前の品物でも部品が揃えてあるのに、日本製品は二年もたつと修理できない」と三流国ぶりを自嘲的に言ったことはありますが、今や世界の最先端に行く会社が、こんなことを言うとは思いませんでした。このままでは永年の蓄積が死んでしまいます。そこで幹部クラスの人に手紙で抗議の直訴をしました。とくにつきあいがあった方ではなく、見学に行つて名刺をもらった程度の人です。東芝の研究所長だった下村さんという方がまじめにとりあげてくれ、結局会社のどこか片隅に生き残っていた機械を使って、すべてのデータをフロッピー化してもらえました。でもこの話しは役員会にまで持ち上げられる騒ぎとなったそうで、解決までに足かけ3年かかり、片づいたのは1989年3月です。さもないければ、私のところの長野県データは生き残らなかったでしょう。

FD化のためには、データのすべてを手持ちのディスクカートリッジに入れて渡さなければなりません。それも一回しかやってくれない約束でした。それまでに作ったデータは、

長野県のものだけではありません。日本地名索引、日本地図（海岸線と県界）、等高線、標本、学名・和名・分類体系シソーラス、文献目録などのデータがありました。

長野県チェックリスト（1985年）を作るとき、植物名や地域コードでソーティングをする必要があり、12万件のソートには、手持ちのディスクカートリッジ16枚全部を動員して数ヶ月かかりました。植物誌以外のデータは当然どこかに退避させねばなりません。ディスクの生産はストップしていましたし、あっても高価ですから、買い足すことはできませんでした。ですから退避先としては紙テープ出力しかなかったのです。何週間もかけてものすごい量の紙テープを作りました。チェックリストの仕事が終わったとき、このテープを作業用に使ったディスクにまた戻したのですが、ディスクの何枚かは傷んでしまって使えなくなっていました。つまりこの段階で既にディスクに戻せなくなったデータがあり、今度さらにFD化に出さないデータを選び出さねばなりません。このことは、選ばれたデータは死ぬということです。その対象になったのは等高線データでした。というのはこのデータは非常に大量でしかも未完成（といっても日本全体の70% くらいはできていた）だったからです。現在の分布図を見て「等高線があるとよいのに」と言われることがありますし、自分でもそう思うのですが、長野県植物誌のデータを優先的に生き残らせるためには、切り捨てる他はなかったのです。もう一つは文献目録のデータです。自費出版していた片仮名の目録を記憶しておられる方がおられるでしょう。ABOC社で植物分類学関係文献目録・索引を出版するようになってからは、このデータは使いにくくなっていました。磁気テープが扱えるシステムならどの機械でも処理ができますので、こういう無用な殺戮をしないで済んだのです。

東芝の約束では、「これらのデータを3.5インチフロッピーにして無料で提供する。ただし変換の際のエラーは補償しない。」というものでした。ところがこちらには3.5インチを使

える機械がありません。「J3100を2週間貸すから、それで確認してもらいたい。」と言われました。J3100は当時最新のラップトップパソコンですが、使い方もわからない機械を2週間でこなすのは無理ですし、エラーを見つけたところで補償はしてくれないので、だまって受け取ることにしました。そんな経過の末、400枚ほどのFDが帰ってきました。長野県のデータは30枚ほどでしたが、ソートの異なるデータセットが2組あって、両方とも保存したので60枚になりました。最近になってその両者のデータの中身が違うことがわかり、両者をつき合わせて（これがまた大変な手間でしたが）、不足のデータを補うことができました。さて400枚のFDがあっても、使えなければどうにもならないので、結局私もJ3100を買うことになりました。その頃には科研費で富士通のFMR-HXというパソコンが入っていて、別に東芝の機械を買うことはなかったのですが、データのエラーの原因が、変換処理のせいなのか、使用機種が違うためなのかかわからないのは面白くなかったからです。20年も昔の、TOSBAC導入のときの磁気テープの騒ぎが、ここ迄尾を引いていました。J3100の値段は給料の n 倍（ $n > 2$ ）でした。「買うけれどコレしか金がない」と $n < 2$ に設定して業者にもちかけたら、ホントにその金額でわけてくれました。儲けどころか手数料にもならないから販売ルートには乗らず、担当者個人の計らいだったそうです。プリンタがないと分布図を保存したり見せたりできないので、これも一番安いものを買わざるを得ませんでした。安いだけにキメが粗く、今となっては不満足です。結局J3100は自宅、FMRは科博で使い分けていますが、両機で使うBASICが異なるため、プログラムは二とおり作らねばなりません。

こういういきさつを経て、新しいパソコンFMR-HXに東芝が変換してくれたフロッピーを、データベースとして移植することができるようになりました。今度の機械は70MBのハードディスクを内蔵しているので、長野県のデータ処理は十分対応できる見込みです。私

は多くの皆さんより早くコンピュータを使い始めたのがたまたま、パソコンの時代に入るのが大変遅れてしまい、MS-DOSというOSはこれが初めてでした。データベースソフトはdBASEIII PLUSで、前にこりていたので、処理レコード数に制約がないことを、購入時にしつこいほど念をおしました。

はじめてのソフトにとまどいながらやつのことで長野県データはdBASEのデータベースに納まりました。リストにしてみたら、大きなエラーはなさそうで一安心しました。ところが最近になって、以前のチェックリストには出ているレコードで、見つからないものがあることがわかりました。そこで東芝が出力してくれた二とおりのFDをしらべたところ、両者は中身がすこし違うことがわかりました。dBASEには量の多い一方しか入れなかったのですが、重複レコードを削除するときに、ソフトに慣れていなかったために誤って消し過ぎたらしいです。そこであらためて二つのレコード群を全部入力し、ついでに既に作ってあったデータベースも混ぜてから（これには漢字データを加えてあったので）、重複レコードの削除をおこないました。このときわかったのは、地名や人名の綴りの僅かな違い、たとえばITO SとITO,Sとか、ハッポウオネとハツポウオネ、が原因で、レコードの他の部分が全部同じでも削除できないことです。結局標本番号とか採集番号とか、レコードを特定する最も頼りになるデータだということあらためて認識しました。

ずっと昔に作ったFORTRANプログラムを思い出しながら、なんとか見られるグラフィック分布図をBASICで作りました。最近の分布図がそれです。

以上書いてきたもろもろの事件は、要するに記憶容量が小さいコンピュータで、身分不相応な大量のデータを処理しようとしたところにあります。今日のように、ラップトップパソコンでも100MBハードディスク内蔵くらい当たり前で、しかも強力なデータベースソフトが使える時代の人からみれば、私の話しはひどくマンガチックなドタバタ劇に見える

ことでしょう。長野県植物誌は20万レコードを目標にしていますが、1レコード100バイトとしても、全部で20MBに過ぎません。しかし、20MBを「に過ぎません」という言い方は、始めた頃には思いもよりませんでした。全部をソートしたとしても、80MBあれば片づきます。インデックスファイルを使えばもっと簡単です。一部屋を占領していたハードウェア群も、今では片手でぶら提げられるラップトップパソコン一台で用が足りるのです。時代の進化の早さを身にしみて体験した15年でした。

しかしいくらパソコンが進歩しても、いや、すればするほど、きわめて深刻な問題があることに気付きました。それはTOSBACからFMRへ乗り換えたときの事件で、たくさんの電算機データを蓄積しても、それは早晚読めなくなってしまうということです。同じ会社でさえそうでした。ましてよその会社の機械で作ったデータならなおさらです。これまでに蓄積したデータがいつでもすべて読めるようにしておくためには、機種やOSが変わるたびにすべてのデータをそれに合わせて作り変えてゆく必要があります。「半年たてば新しい機種が出る」というような時代に、そんなことを一々やってはいられませんし、やればものすごい経費が継続的に必要です。ですからもしやれるとしても「差し当たり必要なデータ」のみを変換し、あとは元のまま保存するということがおこります。保存といってもFDが読めないままに置いておくのですから、後の役には立ちません。これでは文化の蓄積はできません。そればかりか、その時々当事者に都合のよいデータだけが受け継がれ、いわゆる「正史」だけが、誰がいつどのように手を加えたかがわからないまま、残ることになり

ます。

文字や画そのものと、たとえ理解ができなくても、それがあることが認知される形で残ります。だから「四畳半襖の下張り」というような発見もあるし、墓から出た金属板に「太安万侶」と記してあれば、それが意味することはわかります。意味のわからないエジプト古代文字でも、ロゼッタストーンが出てくれば解読できる可能性があります。しかしたかだか百年の未来に、われわれの文明の廃墟でフロッピーディスクが発見されても、それが記録媒体であることはわかっても、そこになにかが書いてあるのかないのかさえわからないでしょう。情報の解読がわれわれの知識ではなく、機械を通じてでなければ不可能になっており、メーカーはロゼッタストーンを残すことなどサラサラ考えていないからです。情報がコンピュータメーカーの独占的（独裁的という方がよいでしょう）支配の下に置かれるペーパーレス社会は、文明を衰退させる敵となるおそれが非常に大きいと思いました。

ですから長野県植物誌は是非とも文字と図の形で残してほしいのです。電算機データがあるからといって、それを頼りにしない方が賢明だと思います。これは手段であり道程にすぎません。できるだけ早く立派なゴールに達するために、これからもお手伝いしようと思います。長野県植物誌データベースの道のりは、私にとっては電算機を独習する勉強部屋でした。その原動力は、あの時の「体が震えるような」嬉しさ、誰も無関心と思っていたのに、私のシステムに利用価値を認めてくれる人達がいたということを、忘れることができないからです。