

# あたらしい標本管理・検索システム～GEMS～の誕生

坂 卷 幸 雄 (鉱床部) ・小 野 晃 司 (地質部)

地質調査所は いま約3万3千点の標本を持っている。この標本を有効に利用するために 昨昭和50年8月からコンピュータを使った新しい管理・検索システム GEMS—(EDP System for GEological Museum Samples) がスタートした。それ以来 新規に受け入れる標本はすべて新システムで処理されている。また 以前に受け入れられているものも これからだんだん新システムに移しかえられてゆくことになっている。

昭和47年度にはじまる GEMSの開発には もちろん所外のシステムエンジニアやプログラマの力を借りたが基本的な構想を支えたのは 実際に標本室の運営にあたっている2人の専任担当者を含む標本室運営委員会 および そのなかの作業部会である通称「コード化委員会」であった。その意味で GEMSは これらの人々全部のチームワークの産物である。

私たちは 標本の提供者・利用者であって 残念ながらコンピュータや情報処理の専門家ではない。そのこともあって GEMSの開発にあたっては システムを使う側としての発想が常にあった。1年間動かしてみたが 利用者の側から問題を考えていたことに それほど大きな間違いはなかったように思う。では 「考えていた問題」とは何か それらは 実際にどのようになっていたのかを 今の時点でふりかえてみたい。

## これまでの標本管理

博物館や大学をはじめ 地質標本を管理している研究機関や社会教育機関は多い。わが国の場合 そのほとんどは 標本の登録や検索に カードか台帳を使っている。進んだところで ホールソートカードかマークカードを採用している例も わずかだが ある。反対に現物とラベルがあるだけという場合も めずらしくはない。

地質調査所も これまでは 台帳とカード(標本名産地 鉱種等 の3枚が1個の標本について作られる)の併用方式で 同時に 国内産の標本で産地が明らかなものについては 5万分の1地形図上にその位置がプロットされていた。この方式は きわめて手がたいやり方で 当面特に大きい支障が生じていたわけではなかったが 将来の利用度の向上という点からは いくつかの不安材料があった。

その1つは 研究者が標本室にいて 台帳やカードや地形図のファイルを繰らなければ 目的とする標本の有無がわからない—という点である。仕事の密度が年とともに濃くなっているいま これはなかなか気の重い作業である。収蔵されている標本の数がふえればふえるほど(このことは本来は喜ぶべきことなのだが) 逆に 検索の能率はおち 見落とし 誤認の確率もふえる。

もう1つは 記録されている内容が基本的な項目に限られている—という点である。その標本 その産地などについて 分析値などの実験結果や 記載された文献がほしくなるという例はしょっちゅうある。標本自体についても 大きさ 産状 共生関係などは カードや台帳からは読めない。どこにしまっているのか。自分の研究に 少しカケラがほしいのだが もらえるか— このような情報はすべて専任担当者の献身的な努力によって記憶され提供されてきていた。

しかし 筑波研究学園都市へ移らねばならない日は 昭和54年度と せまっている。

移転にあたっては ここにのべたような点が 差し当たっての障害になるだろうと 当然予想された。コンピュータによる標本管理は この対策をたてる

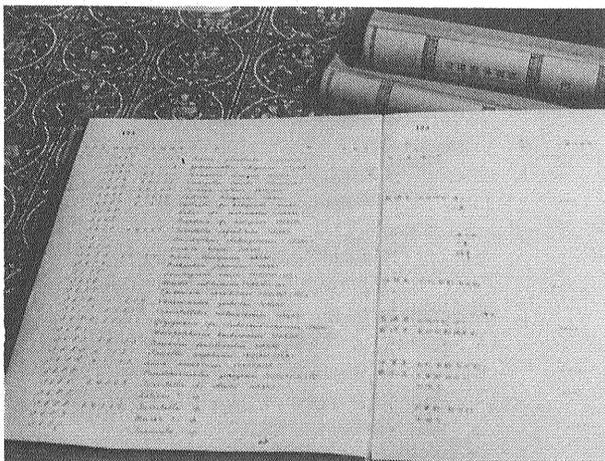


写真1: これまで地質調査所で使われてきた標本台帳。岩石・鉱物・化石の3分冊になっていて ずっしりとお厚い。震災 戦災の2度の災厄をのりこえて 今日まで蓄積されてきた標本の 全記録である。

にあたってでてきた筈のうちの一つである。

これまで どこでもこのような試みがされなかったことの理由の一つは 文献関係の情報検索とちがって 標本などの場合は 費用をつぎこみ 手間ひまをかけても恐らく引き合うまい一との しごく常識的な読みがあったからだと思う。 さいわいにして GEMSの開発では 目先の経済効果をみてプラスかどうかという議論はほとんどしないですんだ。 しかし プラスは 他の面で着実にあったとみてよい。 それは 標本の果たす役割を動的なものとしてとらえる視点が生まれてきたこと そして 地質学で扱う情報のデータ構造について かなりはつきりしたとらえ方ができるようになったこと一の 2点に象徴されるものであった。

### 「静」の標本から「動」の標本へ

地質の標本は 人間の手の加わった 工学分野の標本とも また おなじ自然科学のなかでも 動植物の標本などともことなった いくつかの特徴をもっている。

それをまとめて一口で言えば ひとつひとつの標本の独自性が きわめて強い一とでも表現されよう。 そのため いったん一つの標本をなくしたり 使い切ったりしてしまえば それと同質の標本をふたたび入手できる機会是非常に少なくなる。 極端な場合には 石の切れ目目が研究の切れ目にもなりかねない。

このような危険は 最近 特に切実なものになった。

たとえば こ二十年ほどの間に 大きな鉱山が次々と閉山していった。 鉱物産地として名の通ったところばかりである。 このようなところでは 今後 研究上の必要がおこっても 再び坑内にはいって必要な標本を採集してくることはできない。 一般の想像とは正反対に いわゆる国土の「開発」が進めば進むほど フィールド

の条件は悪くなる傾向がある。

一方 標本を使って研究を進める側でも 近年 テーマの分化と フィールドと実験室の分業化が進むようになってきた。 この傾向が強まってゆくと 一人の研究者にとっては 標本のもつ多様な情報の一部を活用しただけで もうあとはご用済みとなる場合が多くなるだろう。 そのうちの一部は これまでも標本室に提出される習慣があった。 きちんと記載がされた標本 いろいろな意味で典型的なもの あるいは特異なもの 外見が特に美しかったり 大きかったりするもの などである。 そのほかの大部分に相当する標本は 死蔵されたり はなはだしい時には惜しげもなく捨てられたりしてきた。 一つには 標本の収納スペースが不足していることもあるが 基本的には 研究者のイメージの中に 標本室とは標本をしまう(収蔵・保管)ところか みせる(展示・陳列)ところだという認識が根づよく残っていることが このような結果につながっていたものと思われる。

一方では 標本を入手する途がますますせまくなる。 しかし 標本を入手したいという要望は少しも下火にはならない。 そうなると やはり しまいこまれてホコリをかぶっている標本を 研究者の手もとから広い場へ移す必要がある。 理想をいえば いま研究中の標本でも 状況に応じて標本室に納め またスムーズに引き出すことができなければならないのである。

これまでは 標本の扱いはいわば「静」の立場で行なわれていた。 今後は これを「動」の立場でみてゆこうというのである。 最初の予想をはるかに上まわって この「動」の立場での標本の取扱いにとって GEMSが有力な武器であることがだんだん認識されるようになってきた。

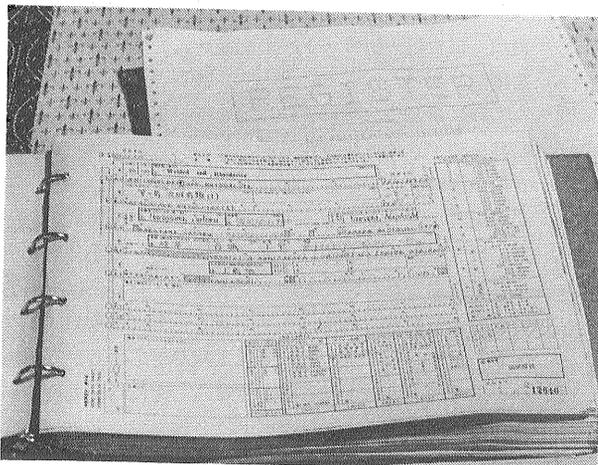


写真2 写真2・3 新しいデータ・シートのファイルと磁気テープ。

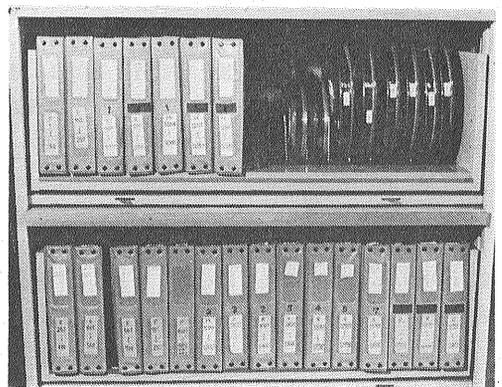


写真3

これまでの台帳一カード方式にかわって 昨50年8月から切りかえられた。





は地学の教育を受けた人という考え方が前提にあるので通常の情報検索のように たとえば同義語の提示のようなキメこまかさは見せていない。

産地リストでは JIS コード 海外コードにそって出力される。同一コード番号にいくつもの標本がある場合には、分類コード順にソートされる。

地質単位リストは 入力するときBカードに記入した地質単位に対応して それに属する標本のデータを打ち出す。採集者リストも 地質単位リストと同じように今回あたらしく GEMS によって追加された部分である。

現在までのところ このように出力は5種類のリストの形で与えられ それ以上の高度の検索機能はもたせていない。したがって 利用者はリストを参照しながら自分の必要とする標本をえらぶことになる。一見 これはいままでの台帳—カードシステムと大差ないようにみえる、しかし もっとも決定的なちがいがあ

それは 均質で 加工可能なかたちで情報が入力・収納されているので 今後は「おしきせ」の出力形式だけではなく 処理手法に応じた利用法が期待できることである。いわば 良い材料を豊富に仕入れておけば あとは料理人の腕次第で 良い料理がどんどんできるようなものである。

これらを通してみると 第1段階として考えていた目標は ひとまず達成されたといえるであろう。

### G E M S の 問 題 点 と 将 来

実際に提出されてきたデータをみて まず最初に気がつき そして 対策を迫られたのは データの「ゆれ」であった。

「ゆれ」がおこる原因はいろいろある。一番根本的なところでは 基本となる分類体系自体が確定してなくて 細部で「ゆれ」がおこることがある。分類体系自体が きちんとしたトリー (樹木) 状の階層構造をもっていれば それでは「ゆれ」がおこらないかといえ

決してそうではない。上位の名称を使うか 下位の名称を使うか 異名の取扱いをどうするか—などという点では 研究者それぞれの間に微妙なちがいがあ

これらをムリに統一することはとてもできないので 明白なミスや 新しくデータが付加された場合を除き 標本名については提出者の表現をそのまま採用している。適当な字引をコンピュータに記憶させておけば これらの「ゆれ」の一部は解消できよう。だが 一定程度は利用者側の常識でカバーしなければならない部分が残るであろう。

受入側で一番「ゆれ」がおきやすいのは 分類コードをあてはめる時と考えられる。一方では ローマ字化に際して 地名 人名等の不統一がおこることもありうる。たとえば 普通のラインプリンタでは長音符号が使えないが その際の書きかえの基準となるような有力な規範はまだない。これらを含めて さまざまな点についての標準化を いま急ぐ必要がある。

次の問題は 入出力の機構の部分にある。

現在ある約3万3千点の標本についてのデータをすべて入力するためには 相当の費用と手間がかかる。とりあえず これまでの台帳の内容の切換えは 標本整理のすんだ部分を重点に 年代の新しいものから古いものへと順にさかのぼることを基本としたが ネットはカー

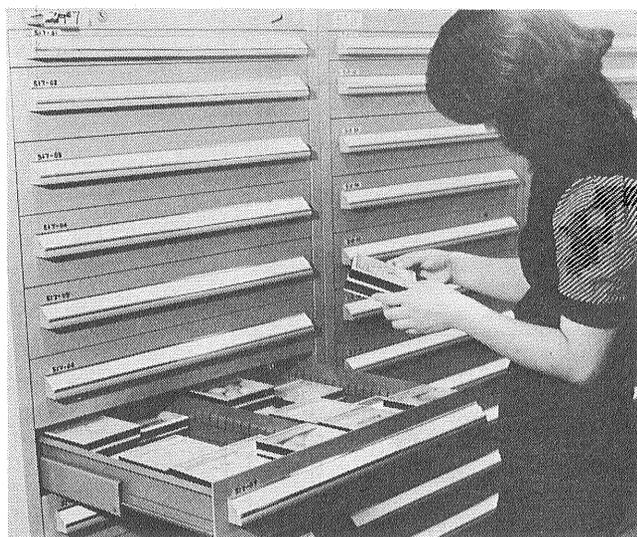


写真4

写真4・5 新・旧の収蔵設備。GEMSでは 収蔵場所も大事な入力データである。そのためモロブタと棚の一つ一つに全部番号が打たれた。

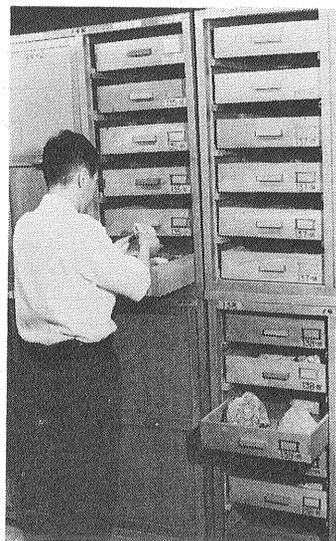


写真5

写真4・5 新・旧の収蔵設備。GEMSでは 収蔵場所も大事な入力データである。そのためモロブタと棚の一つ一つに全部番号が打たれた。

25/06/1975

\*\*\* REGISTER NUMBER LIST \*\*\*

PAGE RGNO R0001

R09301	310, ---	SANDSTONE (B) GIONYA 4A F (C) NOZAWA=TAMOTSU TN58022105 (D) 45443 MIYAZAKI-K NISHIUSUKI-G GOKASE-T KURAOKA GIONYAMA (E) N152-06-13- KURAOKA PLOT -	STLUR	B18-0108-01	1944
R09331	420, ---	SCHIST, PIEMONTE CHLORITE PLAGIOCLASE (B) RYUHOZAN METAP. RKS (C) NOZAWA=TAMOTSU TN65031805B (D) 43466 KUMAMOTO-K YATSUSHIRO-G TOYO-V ABURADANI (E) N152-12-05- YATSUSHIRO PLOT -		B18-0159-09	1944
R10021	340, ---	LINESTONE (B) ISOMI=HIPOSHI 40-42(38) (C) 14441 FUKUI-K MIKATA-G MIKATA-T KAWACHIMINAMI (D) N153-13-03- NISHIZU PLOT -		B18-0144-04	1944
R10111	010, ---	ADAMELITE, BIOTITE (B) YAKUSHIMA GRANITE (C) NOZAWA=TAMOTSU TN66031003 (D) 46504 KAGOSHIMA-KKUMAGE-G YAKU-T KOSUGIDANI		B18-0136-06	1966
R10141	420, ---	SCHIST, PIEMONTE MICA QUARTZ (B) IIMORI MINE (C) NOZAWA=TAMOTSU TN66051301 (D) 30323 WAKAYAMA-K NAGA-G NAGA-T IIMORI-KOZAN (E) N153-15-11- KOKAWA PLOT -		B18-0106-10	1966
R10171	480, ---	HORNIFELS, BIOTITE (B) NOZAWA=TAMOTSU TN67022106 (C) 46341 KAGOSHIMA-KKAWANABE-G KASA-T KOSAKI (E) N152-07-16- NOHADAKE PLOT -		B18-0106-09	1966
R10231	120, ---	ANDSITE, AUGITE (B) ASO VOLCANO (C) YATSUBARA=HIDEKI 6902 HAD (D) 43422 KUMAMOTO-K ASO-G ASO-T SAKANOSHITA (E) N152-11-05- KIKUCHI PLOT - (F) LOWER HORIZON OF SOIMA DEPOSITS, STORE/6 SPEC. IN 0104-04, 3 IN -05.		B18-0104-04	1966
R10321	010, ---	GRANITE, BALL (C) OYAMA=GOKICHI 432(12055)H (D) 08303 INARAKI-K MAKABE-G MAKABE-T (E) HJ54-24-15- MAKABE PLOT -		B18-0136-05	1966
R10471	460, ---	SCHIST, ACTINOLITE PIEMONTE QUARTZ (C) NOZAWA=TAMOTSU A64080606-4 (D) 21626 GIFU-K YOSHIKI-G, KAMITAKRA-V, NAKAO (E) HJ53-06-07- KAYIKOCHI PLOT -		B18-0104-11	1972

第3図  
出力された登録番号リスト

ドのパンチにある。近い将来の入力方式はマークカード方式から光学的文字読取装置(OCR)でデータ・シートの内容を直接読みこませる方向へ進むであろう。

出力側でも近い将来もっと完全な形で多重検索機能が要求されるであろうしその段階では当然のことながらラインプリンタによる表作だけでは満足できなくなると思われる。この目的に不可欠なのはGEMSにふさわしい会話型検索プログラムの開発とそれに見合った入出力機器の整備であろう。CRTディスプレイやプロッタ、ディジタルなどが自由に使えるようになればこの面での急速な前進は十分期待できる。

さらに最近文部省特定研究「広域大量情報の高次処理」の進展によって姫路工大の弘原海(わだつみ)清氏らを中心として地学分野でのデータ・ベース・システム—GEODAS—が開発された。先のべた地学情報のデータ構造の特徴もこの研究班での討議のなかで明らかにされてきたものである。

GEODASはそれぞれ独自性をそなえた研究者ファイルを入力データとしてそれらを総合的に管理し要求に沿ったデータを選別して提供する機能を持つ。今回くわしくふれる余裕はないがGEMSのデータ・フ

ァイルを他の二、三の研究者ファイルと共にGEODASのテスト・データとして用いた結果は予想以上に満足すべきものであった。このことは将来各研究機関・各研究者の間でデータを集積し視点をさまざまに変えた多様な検索をこころみることによって新しい課題やテーマを発見できる可能性を明確に示している。

一定の方向づけを持った量の集積は必ず質的な飛躍をもたらすという。その意味でGEMSでいま試みられている実験はただ標本の取扱いの利便ということだけにはとどまらない。この試みを通じて地学データの基本的な性質とそれらを取扱う際の具体的な手法が今後一層明らかになるものと私たちは信じている。

・GEMSのデータ・シート記入のためには次の所内資料が刊行されている。  
地質調査所標本室運営委員会(1975):「地質標本用データシート」記入の手引

・GEODASについては下記等を参照  
弘原海清(1975):測定尺度のラティスを持つ地学用データ・ベース・システムの基本設計  
姫路工大研究報告 No.28-A 17-27