

資 料

55(083.6) : 651.8 : 681.177.2 : (73)55 : 061.6

米国地質調査所における試料の記載、分析データの 保存とその再利用の方式*

T. G. Lovering & D. F. Davidson

高 橋 清 訳

1. 緒 言

米国地質調査所では、いわゆる地質試料（岩石、鉱物、土壌、植物など）の物理的性質の記載と化学分析値の保存とその再利用の方法としてパンチカード（punched card）方式をとりうまく運営している。

この方式は約 50,000 試料の種々の記載と分析値を保存・整理してある“mechanical library”と考えてよい。当然のことながら、米国地質調査所の全研究職員がこの“mechanical library”を利用することができるが、おのおのの試料についてのこの方式の運営に必要な記載の義務をもたされている。

このデータ保存と再利用の方式が運営されてからは、よく利用されているが、とくに統計的な研究計画の立案、実施に役立つている。

現在、米国地質調査所の各研究室では毎年約 20,000 個の岩石、鉱物、土壌、植物などの試料の化学分析が行なわれており、1 試料について 1 ないし、ときには 100 以上の成分の分析が実施されている。

1961 年にこのデータ保存、再利用の方式は確立されたが、それ以前は、試料の分析結果は試料採取者と分析依頼者だけに報告されていたので、この試料の分析値は大多数の研究職員の目にふれることなくほとんど役に立っていなかった。たとえこの分析値が論文の中に掲載されるなどして、多くの興味をもっている研究者の目に触れやすい状態にあつたとしても、試料の産地、産状、性質などの記載がまちまちであることが多かつたため、その有用性はそれほど大きくなかつた。

2. この方式のおいたち

現在運営しているこの方式は、地質調査所の各研究室で行なわれている物理的性質の解析や化学分析値を、どのようにすれば全研究職員が十分に活用することができるか、地質や化学関係の研究職員からなる小委員会で充分討議のすえ生れたものである。小委員会では、所内の地質職員、データ解析者、研究室の管理職員などの、また会社や大学の専門家の意見をとり入れ検討を重ねた。この小委員会の目的は、データ保存とその再利用の方式を確立し、それに関連した組織作りをすることにあつた。

討議の結果、現在使用されている機械処理によるパンチカード保存（machine-based punched card storage）方式を採用した。これは 1961 年当時の地質調査所の電子計算機は、多量のデータ保存、再利用計画の運営には適当ではなく、“エツデパンチ”によるパンチカード法ではデータの整理、とりだしがきわめて遅く適当ではないと考えられたからである。

* T. G. Lovering and D. F. Davidson (1966) : A system for storage and retrieval of descriptions and analyses of samples.

この原稿は日本の地質調査所のために米国地質調査所長の発表許可を得て寄稿されたものである。

3. この方式の2, 3の特長

機械により容易にカードを類別することができ、また最小の使用面に最大量の情報を記録することができるこのパンチカード方式では、ある方法で情報を符号化 (code) する必要がある。米国地質調査所方式はアルファベットと数字の組合せ符号を採用している。

この方式で、50項目以上の記載情報と、8種類の異なつた分析法で得られた1ないし100成分以上の分析結果を保存している。それゆゑ試料の記載についてのものと、それぞれ異なつた分析法によつて得られた分析結果についてのものと、別々のカード箱 (card deck) を必要とする。個々の試料には、それぞれ独自の試料番号が付けられ、その試料の情報の入つていないすべてのカードに付けられた試料番号が書きこまれる。

個々のシリーズのカードは、その探しだしと再利用をやりよくするために順序よく並べることができるように連続番号が付けられるのである。

調査研究職員に配布される書式 (form) には、試料の記載についてのものと、分析依頼のためのものと二通りあり、分析を依頼したい研究職員は、“便覧” (manual) に基づいて、これら二通りの書式にきちんと書きこむ。

データ保存、再利用計画の職員は、これらの書式と、書式に基づいた指令を、それぞれの分析研究室に配布する。その指令というのは、これらの書式が適当に拡張されていることについての注意とか、試料番号を割当てることとか、分析の数や型に無関係に割当てられた試料番号を用いることなどである。

その結果、この方式は試料提出者と分析研究室の試料取扱い管理者との密接な共同体勢をつくり上げる。

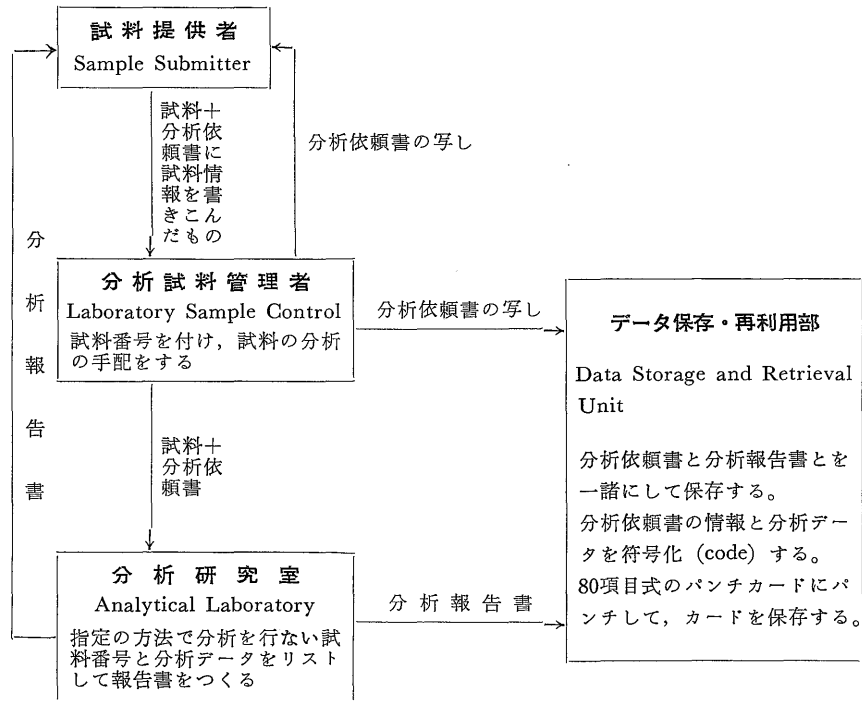
試料提出者はこの方式の運営が行なわれる前に較べると、現在はかなり多くの記載情報を準備しなければならないし、さらにこの記載情報を“便覧” (manual) を使つて符号化しなければならない。われわれが考案したこの符号化の方法 (coding system) は、個々の試料提出者が容易にかつ迅速に用意できる試料情報の量と、情報の再利用者が後に求める情報の量とのバランスがとれるように工夫してある。その結果として、“提出者のための試料符号化指令便覧”の目録に入つている記載語句の大部分は広範囲な一般的なものである。その結果として、一般的な数多くの情報よりも、特殊の情報だけが知りたいという研究職員には他の情報はまったく余分のものとなる。このような研究職員のためには、データ保存、再利用計画へ要求さえあれば、他の余分の情報 (noise) を消して報告することもできるのである。

4. 保存方法への受入れ

はぎ取り式の試料提出に必要な書式と、それを書き込むために必要な指図書とは、すべての試料提出者はだれでも分析研究室で貰うことができる。書式を書きこむときに不明な点があれば、データ保存・再利用計画の職員は試料提出者に適時助言している。分析研究室の試料管理職員は、個々の試料に試料番号を割り当て、書式にその番号を記入する。個々の書式一枚の写しは、データ保存、再利用計画へ送られ、そこでは分析研究室からのその試料の分析結果報告書の写しを受けとるまで保管している。試料の記載は、試料提出書式 (submission form) から一連のカードへキーパンチされ、分析研究室からの分析結果の情報は別の一連のカードへキーパンチされる。両方の系列のカードいずれにも分析研究室でつけた同じ試料番号をつけておく。もし一方法以上の分析法での分析が依頼された場合には、全分析結果がでてくるまで“未決ファイル” (active file) として保存する。

試料提出書式とすべての分析が終つた分析結果報告書の写しは、“既決ファイル” (inactive file) として保存する。

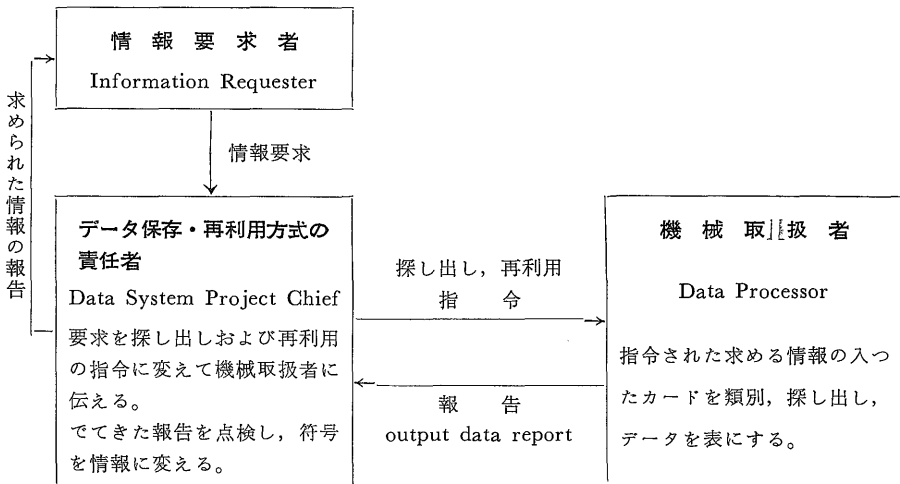
第1図は試料の記載と分析データをデータ保存、再利用方式に受け入れる概念図である。



第 1 図 データ保存，再利用方式への受入れ概念図

5. この方式からの再利用

この方式の中に保存されている情報やデータを集めたいという要求があれば，保存再利用計画の地質職員によつて“探し出しと再利用”指令（search and retrieval order）に翻訳されて，機械オペレーターにまわされる。その指令というのは，機械オペレーターがカード系列を探し出すパラメーターとなるコード符号は何であるという種類のものである。保存カードはこれらの指示されたパラメーターによつて，カード選別機で類別され，類別された情報およびデータカードが正しいかどうか照合される。もし情報を要求した職員の要望さえあれば，類別



第 2 図 データ保存，再利用方式からの再利用の概念図

された情報およびデータは印刷されるか、あるいは新しくその情報データのためだけのカード箱を作ってパンチされ保存される。もし再利用された情報とデータが印刷される場合には、計画の職員は符号化された情報の翻訳を行なう。

第2図はこの方式に保存されたデータの受けだしを説明する概念図である。

6. この方式の利用

この保存、再利用方式に保存されているデータの量が増えるにつれて、これらのデータ情報を再利用する需要も増えてきた。現在保存ファイルには約50,000試料についての分析データと記載情報とが保存されている。1965年には11,700試料のデータを使つて、21の研究グループのために50件の再利用を行なつており、また、1963年この年は再利用を行ない始めた最初の年であるが、3,300試料のデータを使つて8研究グループのために22件の再利用を行なつている。

7. この方式の問題点

このデータ保存、再利用計画の運営遂行にあつてもち上つた問題点は大きく二種類に分けられる。すなわち (1) データの取扱い (processing) に関する問題と (2) 他の部課との連絡保持の難かしいことからおこる問題とである。

はじめの(1)の問題としては、書式からカードへデータを移すときの間違いを最小限に止めることとか、もつとも要求度の高い情報をもつとも効果的に使えるようにまた新しい分析法で得られたデータを保存するためにカード形式を新しく変えたり、改訂したりすることなどが挙げられる。

第2の問題としては、試料提供者が書式を便覧 (manual) で符号化して書きこむときにおこる間違いを正すこと、そして試料提供者にこのような間違いをおこさないよう注意することとか、この方式の情報再利用の能力容量を他部課へ流しよく知つておいてもらうこととか、つけられた試料番号を統一して使い、また各分析研究室の分析データの報告も統一した方法で行なうこととか、将来電子計算機の利用のさいに、現在この方式のファイル中に保存されているパンチカードのデータが両立して使うことができるようにすることなどである。

8. 将来の問題

1961年以来、このデータ保存、再利用方式をパンチカード方式で運営しているが、あわせて米国地質調査所の電子計算機利用も進歩してきた。

科学的問題をとくための、また現在のカードファイル中に保存されているような数多くのデータの統計的解釈を行なうための電子計算機のプログラム作製も進歩し続けている。

近い将来には、これまでパンチカードのファイルに保存してきた情報データの大部分を、磁気テープに記憶させ保存し電子計算機によつて直接再利用が行なわれるようになるものと思われる。しかしこのようなカード保存から電子計算機保存への変革のためには、これまでの試料情報と分析データを記録する方法もいくつか根本的に変えなければならない。

〔後 記〕

米国地質調査所の Branch of Geochemical Census では、Chief の D. F. Davidson 氏、および Tom. G. Lovering 氏の努力により、データ保存、再利用方式を確立し現在運営している。1964年に I. U. G. S. の小委員会として data coding の委員会がおかれ上記の二氏が中心となつて、この問題を世界的な規模にしようと努力している。この原稿も希望により日本の地質調査所のために好意的に寄稿されたものである。

具体的な試料の記載の方法とか、書式に書きこむさいに参考にする「便覧」(manual) などについては「地質ニュース」に別に紹介するつもりである。(高橋 清)