

# 花粉データバンク

## —石油地質学からみて—

齊藤 隆\*

Palynological Date Bank : A Petroleum Geologist's View

Takasi SAITO \*

筆者は、石油井戸 (oil well) のデータを取り扱うデータバンクである "Well Data System" について、この約1年のあいだ調査をおこなってきた。最新の文献の収集と、米国・カナダにおけるこの種のデータバンクの現況視察とが調査の早道と考え、それを実行した。調査旅行は、石油会社の研究所および情報部門をはじめ、大学、地質調査所、資源保護局、情報サービス会社、コンピューター会社などを対象とした。その結果、石油開発公団において、昭和47年度に1つの小型データバンクを試作するはこびとなり、その準備もほぼととのった。

この1年間の調査の内容は、もちろん大部分が石油井戸のデータ処理に関するものであったが、あわせて古生物データの貯蔵・検索などについても多少の知識を得ることができた。近年、花粉学の分野においてもコンピューターをベースにした情報システムの問題がとりざたされている模様であるが、このたび、上野花粉学会長とこの問題について話しあう機会があり、これについて執筆をすすめられた。筆者は、花粉については盲同然であるが、筆者の小さな経験が花粉研究者の方々にお役にたてばと思い筆をとった。

### 1. データバンクの2つの型

先人の業績が、それが誤ったものや架空のものでない限り、絶対的な先取権をもつ「博物学」や「特許」などの分野においては、情報の収集が研究活動に優先することが多い。営利を目的とした研究活動の場合は、情報の収集のための時間（費用）と、それによってもたらされるメリットとがバランスするところで情報収集はうち切られるであろうが、分類学などの場合には、最古のものから最新の情報まで完全に収集されることが強く期待される。ここに国際的な分類学データバンクの必要性が主張される1つの大きな理由がある。

一方、このような「カタログ」としてのデータバンクに対して、個々の種 (species) の地理的・空間的、あるいは時代的 (層位的) な分布の定量的な観測結果、ならびにそれと他の現象との相関についての観察結果を大量に集めたデータバンクがある。例えば、ある盆地を代表する一定の空間の中の花の組成と頻度の季節的な変化についての20年間の観測結果、ないしはそれとその盆地内の小中学生の種々の花粉病の患者数との関係の観察データは、1つのデータバンクを必要とするだけの情報量であろう。また、ある堆積盆に掘られた500本の石油井戸における、各花粉化石種ごと各深度ごとの産出頻度のデータも同様である。これら、いわば応用例を集積したデータバンクは、もちろん情報量が多ければ多いほど価値は高いが、カタログのデータバンクのような完璧さは欠いても、それによって重大な誤りが生じることは少なく、十分実用にはたえうる。

### 2. カタログのデータバンク

カタログとしてのデータバンクには、一般に次のような項目のデータが貯蔵されているのが正しいと考える。

標本の名称：種名、亜種名など

採集データ：産地、採集者、採集日など

管理データ：模式標本の保管・移動など

記載：形態的特徴

地理的・時代的分布範囲

生態

花粉病、石油層などとの関係

その他

文献データ：その標本が紹介された文献名、著者、  
図版番号など

これらの項目のなかで、もっとも詳細に記録されるの

\* 石油開発公団情報室 東京都港区赤坂一丁目九番一三号 (三会堂ビル)  
Information Management Department, Japan Petroleum Development Corp.

は「形態的特徴の記載」である。この型のデータバンクとして、(1)オランダの国立地質鉱物博物館 (*Rijksmuseum van Geologie en Mineralogie*)による“花粉・胞子の記載のための数字コードシステム”、(2)米国のスミソニアン研究所 (Smithsonian Institution) による“FNA (Flora North America) 情報システム” (3)コロラド大学 (University of Colorado) による“TAXIR (Taxonomic Information Retrieval) システム”、(4)コロラド鉱山学校 (Colorado School of Mines) による“白亜紀有孔虫の古生物学的・書誌的データの総括プロジェクト”などが知られている。

(1) オランダの国立博物館による数字コードシステムは、花粉・胞子のみを対象としたものであり、次の項目のデータがインプットされる。

《登録データ》

研究所コード、標本スライド番号、産地コード、地層コード、時代コード、命名者コード、記載登録年月日、改訂命名者コード、改訂年月日、新しくおきかえられた標本番号

《分類上の名称》

種名 (*species name*)、科名 (*family name*)

《記載》

形態コード、大きさコード、模様コード、などこれらデータは、1ないし8桁の数字コードに置きかえられて記録するようになっている。このような徹底したコード化は、一見合理的のように思われがちだが実はインプットをするにもアウトプットを読むにも“暗号解読表”なしには作業がまったく進められないという大きな欠点をもつ。最近のデータバンクでは、必要以上のコードは用いず、また用いる場合には“暗号表”なしである程度意味が通るよう工夫される。数字コードは、「小容量コンピューター」時代の名残りともいえる。

一方、あらかじめ定められた検索用語によるさまざまな角度からの検索結果が、アウトプットとして期待できる。しかし、このシステムにおいて、その例として示されているのは、次のような検索のみである。

《記載のとり出し》

希望する種 (*specils*) についての記載を引き出すこと

《種のリストの作成》

いくつかの条件に合致する種のリストを作成する。

これらは、計算をまったくもなわない単純な検索であり、いわば書架にある図書の内容を複製したり、図書リストを作ったりするのと変りない。それだけの検索をコンピューターにやらせるために、多くの複雑なコードを作り、データを「作られた言語」に翻訳し、多くの人手と時間と費用をかけて、彼らはデータの貯蔵をおこない、またおこなおうとしているのであろうか。それとも、進化や堆積の場のシミュレーションなど、高度の応用を考えているのであろうか。データバンクは、貯蔵が目的ではない。研究者の視野をひろげ研究の速度を高めるばかりでなく、その学問分野に質的な進歩をもたらすという確かな展望がなければ、データバンクづくりは徒勞に終ろう。この点について、同博物館から更に情報を入手する必要がある。

(2) スミソニアン研究所のFNA情報システムは、1965年ごろから着手され、昨年までにほぼ完成している植物分類学のデータバンクである。これは北米のシステムであり、これに対応するソ連のシステムは1964年にすでに完成している。また、“Flora Europea”も完成に近づいている。

FNAシステムに花粉データが含まれるかどうかは不明だが、システムとしてはよくできたものなので研究してみる価値があろう。FNAシステムの利点は、同定などの作業においてオンラインでアクセスできることである。コンピューターを使うと使わないにかかわらず、「カタログ」は常に手もとにあって、必要が生じた時に直ちに情報を提供できるものでなければ、カタログとしての意味が薄くなる。CRTのようなディスプレイ装置を通して、コンピューターと直接対話できるものが、利用者本位の「カタログ」データバンクであるといえよう。

(3) TAXIRは、コロラド大学のDavis J. Rogersのグループによって開発された米国で最初の植物分類学データバンクであるといわれ、構造も比較的単純なようである。詳しい情報は得ていないが、研究の対象となりうるものとしてあげておく。

(4) コロラド鉱山学校のシステムは、白亜紀の有孔虫のカタログであり、花粉とは関係ないが、同システムの開発を指揮したHarry C. Kent教授から有効な情報を入手したので紹介する。

インプットは、次の項目についてなされる。

《書誌》

著者、出版年、論文名、書(誌)名、巻冊、ページ、図版

《種名》

次のように、11桁の文字と数字で表現する。

(例) *Ammobaculites alexanderi* Cushman,  
1933 → 11桁コード AMMALE17C33

AMMは属名の頭3文字

ALEは種名の頭3文字

17は属名と種名の文字数の合計から上記の6を引いた数

Cは命名者の姓の頭1文字

33は命名年の下2桁

《模式標本》

種類 (holotype, paratype, hypotypeなどの別)

参考図版、登録された博物館名、博物館におけるスライド番号

《シノニム》

著者によってシノニムとされた種名とその文献名

《層位上の位置》

《地理的な位置》

第1図参照

Western Interior Foraminiferal Project

Stratigraphic and Geographic Location:

December, 1968

Data Code	Reference Code	Locality	AAPG Code	Formation Name, Member, Lentil, Tongue
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80				
060101				
060102				
060103				
060104				
060105				
			Age (and Biostratigraphic) Code	Zone Code
060201				Geologic Age - Biostratigraphic Classification (Verbal)
060202				
060203				
060204				
060205				
			Verbal Description of Stratigraphic Position	
060301				
060302				
060303				
060304				
060305				
			Political Area Code	Latitude and Longitude *
070101				
			Verbal Description of Geographic Location	
070201				
070202				
070203				
070204				
070205				
070206				

\* Precision Code

第1図 コロラド鉱山学校の白亜紀有孔虫カタログ・データベースのインプット書式の例

ここでは、形態の記載が省かれており、また意味のあるコードもおろこんで、システムをコンパクトにまとめている。この学校では、タイムシェアリング専用

のPDP-10というコンピューターを使用し、各研究室や廊下などに多数の端末機が設置されているので、研究者たちは手軽にこのシステムを利用できる。この

ように、検策のみを目的とし、他の応用を一切考えない「カタログ」データバンクであれば、データ項目をこの程度に制限し、さらに詳しい情報は、文献に速やかにかつ確実に到達することによって得るのが賢明なやり方かもしれない。

### 3. 応用のデータバンク

次に述べるのは、医学でいえば「症例のデータバンク」とでも言えるものである。博物学以外の分野で一般にデータバンクと呼ばれているのは、実はこの種のものである。この型のものが、花粉学の分野ですでに開発されているかどうか筆者は知らないのですが、ここでは、ある石油会社の有孔虫 (Foraminifera) の例を示す。

データは、各試料ごとにデータシートに記録される。石油井戸の場合、日本では 10~20 m に試料をとるから井戸の深さが 4000 m として、200 ないし 400 枚のデータシートが作成される。また、データをコンピューターに入れるためのデータカードは、各シートの各種 (species) ごとにパンチされるから、その数は、井戸 1 本で 1 万枚以上に及ぶ。

このように、インプットの手間が大変だがパンチ作業をパンチャーにやらせれば、古生物技師の労働は軽減される。また、データをいったん貯蔵すれば、化石種の垂直分布図などはもとより、biofaeies maps や古環境図の作成も自動化される。データバンクを作ることによって、それにつづく作業 (応用作業) の著しい省力化、迅速化、ないし質的な向上が確実に見込まれる場合は、インプットのためにかなりの労力が投入される。

石油会社では、一般に、化石の種名は学名で示さず番号を用い、例えば "Bulimina 1" のように示すが、この略語の作り方は、1967 年に AAPG (American Association of Petroleum Geologists) によって定められ、手順は次のごとくである。

① 次の順序で文字を次々に省いていき、4 文字になるまでくりかえす。

1. A   2. E   3. I   4. O   5. U   6. W   7. H  
8. Y   9. 2 重文字のうちの 1 文字   10. T   11. N  
12. S   13. R   14. L   15. D   16. C   17. M  
18. F   19. G   20. P   21. K   22. B   23. V  
24. X   25. J   26. Q   27. Z

② 語の冒頭の文字は省かない。

③ 文字の除去は右から左に向っておこなう。

④ 2 重文字はそのうちの 1 文字だけを省く。

この方法によると "Bulimina" は次のように省略される。

BULIMINA  
BULIMIN  
BULIMN  
BULMN  
BLMN

これに類するデータバンクは、有孔虫の分野では、規模の大小を問わなければ、米国のいくつかの大学や研究所などでも保有しているらしい。特に、海底試料データのデータは、種の多様化度 (species diversity) など、定量的な環境解析に有効に利用され、一般にデータの量も相当多いので、データバンクを作って貯蔵しているようである。

### 4. IPDBC

本誌前号で、川崎次男氏によって IPDBC (International Palynological Data Bank Consortium) について紹介された。これは、アリゾナ大学の G.O.W.Kremp 教授を中心とした KPCRP (Kremp Palynologic Computer Research Project) を母体とし、コンピューターをベースとした花粉データの国際的な伝達の促進を目的とするものようである。これが、単に文献の検索を目的とするものか、あるいはそれ以外の特徴をもつものか、その方向づけはこれからの問題なのであろうが、もう少し明確なビジョンを示してほしいと思う。人間相互の冗長なる伝達手段である「言語」をもって書かれた文献の内容を、細大もらさずコンピューターに移しこむことは、不可能ないしは非効率なことである。「すべての文献が語学のハンディなしに国際規約のテーブルをたよりに読解できるという結構なもの」をつくるためには、想像をこえるような労力がインプットのために注ぎ込まなければならないのではあるまいか。また、花粉学の文献の 50% はロシア語で書かれているといわれるのに、国際センターを米国にのみ置くのはなぜか。海洋学のように、「世界データセンター A」を米国に、また同じく B をソ連に置くというようなやり方をなぜとらないのか。それが International と銘うっているだけに筆者にはいぶかしく思われる。

## 5. 問題点

地質学の分野では、地質データを、コンピューターで処理できるような形で、国際的な規模の貯蔵・提供・交換をおこなうことを目的として、1967年にCOGE-ODATA委員会が設置され、1871年現在で15ヶ国がこれに参加しているが、いまだに統一された書式の一部すら発表されていない。ここに、国際的なデータバンク確立の困難さを、筆者は強く感じさせられる。また、カナダでは、国内の地質データを“ナショナル・システム”のもとに集約することについて、1965年から1967年までの2年間、国の代表委員会によって徹底的に検討されたが、その結果、国家的なデータバンクは中央集権的であり作るべきではないと答申している。

一方、かりに大規模かつ詳細なデータバンクが完成したとして、必ずしもすべての研究者がその恩恵に浴せるとはかぎらない。第一に、かなり大型のコンピューターを自由に利用できる立場にならなければならない。第二に、費用の問題である。かりに、データバンクのためのハードウェアおよびインプットの費用が、国費や寄附でまかなわれたとしても、磁気テープの中のデータをとり出すための費用は利用者が負担せねばなるまい。たとえば、たった1つの文献を検索するために、3000円もあるいはそれ以上もの金を支払わねばならないとしたらそのデータを使える人の数は著しく制限されよう。第三に、コンピューターによらず、印刷物やマイクロフィルム の形で提供される方が便利なることも決して少なくないことを忘れてはならない。

他人の行動を自分のやり方にならわせようとするのは大変なことである。大勢の人々の行動を標準化しようとするのは、なおさら大仕事である。データバンクの建設のために、何人もの有能な花粉学者が数年間の貴重な研究時間を割き、そしてその結果が、経済的に恵まれた一部の研究者にしか利用できないようなものとなることには、決してしてはならない。

## 6. むすび

筆者は、花粉学のデータバンクが、何らかの形で実現することを期待する。そして、日本の花粉学者も、この間蔵にもう少し関心をはらってしかるべきだと考える。データバンクの開発をすでに経験した、他の多くの分野の人々と交流をもち、自分たちのデータ伝達の体系はいかにあるべきかを真剣に考えるべきである。そのうえで、データバンクの目的をどこにおき、規模をどの程度にするかを決め、次いで、システム開発の具体的な戦略をたてるのがよいと思う。

筆者はまた、日本の多くの大学や企業が使用している中型のコンピューターにのるようなデータバンクを、日本の花粉学者が、日本のデータをもとにして試作されることをすすめたい。それと同時に、国際的な大型のデータバンクが完成した場合の、日本側の受け入れ態勢を、十分に検討すべきである。

以上、部外者のこととて、的を射ていない記述はお赦していただきたい。石油地質家の立場から、花粉学の石油産業へのよりいっそうの進出を望んでやまない。

## 文 献

- (1) Germeraad, J.H. and Muller, J. (1971): A computer-based numerical coding system for the description of pollen grains and spores. Vol. 1 and 2, Rijksmuseum van Geologie on Mineralogie, Leiden-the Netherlands.
- (2) Shetler, S.G. et al. (1971): Pilot data processing systems for floristic information. in Data Processing in Biology and Geology, J.L. Cutbill, ed., Academic Press, London, p. 275-310.
- (3) Cutbill, J.L. (1971): A format for the machine exchange of museum data. in ditto, p. 255-274.
- (4) Kent, H.C. (1970): Final report NSF Project GN-785; Compilation of paleontologic and bibliographic data for Cretaceous Foraminiferida. (Colorado School of Mines), typewritten.
- (5) Kremp, G.O.W. (1970): Towards computerization of palynology-paleobotany: a progress report on a fact-finding trip. Interim research report No. 1, typewritten, 27 p.
- (6) Cross, A.T. (ed.) (1964): Palynology in oil exploration; a symposium (San Francisco, 1962). SEPM Special Publication No. 11, 200 p. (includes 15 articles).
- (7) Committee on Standard Stratigraphic Code (1967): Standard stratigraphic code adopted by AAPG. Amer. Assoc. Petrol. Geol. Bull., vol. 51, p. 2146-2151.
- (8) 川崎次男 (1971): I.P.D.B.C. について、日本花粉学会会誌 No. 8, p. 37.

## 新刊書紹介

長田武正著 日本帰化植物図鑑

A 5版245ページ 北隆館 1972年3月発行 3,000円

著者は先に“帰化植物図譜”第一学習社、(1967)を発行し、私もおおいに利用していたが、なにぶん福岡県内の帰化種を対象としたものであるため、187種が記載されているだけであつた。著者がこの図鑑の序分の中でいっているように“戦後、帰化植物の数が著しくふえている。しかしそれらの大部分は今日までに日本で刊行された植物図鑑の類にはのっていないものであつて、なかなか名を引き出し難く”ために各方面から“日本帰化植物図鑑”の出版がのぞまれていた。そして今回まちのぞんでいたこの図鑑が発行されたのである。帰化植物の中には花粉病をおこすなどの有害な種も多く含まれており、帰化種の同定は花粉を研究するものにとつてもたいせつなことである。

この図鑑にはレンゲソウなどのように従来どの図鑑にも載つていて、他の種と見誤るおそれのない種は含まれておらず、それ以外の532種(内471種が図示されている)が1ページに2種ずつ収録されている。そして、各種ごとに全形図と種の検出の鍵となる細部の拡大図が豊富につけられている。これらの図は日本産スギゴケ科の研究で学位をとつた著者のきめこまかな目でとらえられているため、明解であるうえたいへん正確である。また種の記載文も近似種との比較に重点がおかれ、そのうえ近似種が多く同定の難しいグループには検索表がつけられていて、あくまでも種の検出ということに主をおいてつかいやすい図鑑である。

(田中肇 東京都練馬区羽沢2-28)