

日本地質文献データベース(GEOLIS)の紹介

—GEOLIS、FD版の公開にあたって—

本 庄 時 江・菅 原 義 明 (地質情報センター)
Tokie HONSHO・Yoshiaki SUGAWARA

1. はじめに

地球科学分野での世界の文献データベースとしては GEOREF GEOARCHIVE GEODE などがあります。中でも GEOREF は Dialog のファイルの一つとなっており商用ベースで利用できますし 年間7万~8万件を収録する膨大な情報量を持っております。しかし日本に関する文献は数パーセント(2千件前後)です。一方 GEOLIS は年間8千件ほどの情報量ですが 地球科学分野では日本で唯一の文献データベースです。

世界の情報は GEOREF で そして日本に関する情報は GEOLIS から を合言葉に 私達はデータベースの構築に臨んでいます。

地質学分野での日本で初めての文献目録は1個人によって作られました。元東京教育大学教授藤本治義氏による「日本地質文献目録」で 昭和17年に地人書館より出版されました。その後改訂されて 昭和31年に「日本地質文献目録 1873-1955」として刊行されました。この目録には 論文数約8,000件 雑誌タイトル数174件が収録されています。

藤本先生の目録改版作業とはほぼ時を同じくして 地質調査所が創立75周年記念出版物として目録作りに取り組み 「地質文献目録 1945-1955」が 昭和32年に出版されました。これが ここで紹介する GEOLIS の前身です。

「地質文献目録」はこの後 年刊で刊行が続けられ「1981年版」までの刊行を終了しました。1986年からデータベース化を開始し 同時に「日本地質文献目録」として年刊発行してきましたので 未充分は1982~1985までの4年分となりました。

地質調査所発行の「地質文献目録」には 1945年版から データベース化された3年分約2万1千件まで加えると約10万件にもおよぶ論文が収録されています。

藤本先生の目録は80年分で約8,000件が収録され GEOLIS では 1988年の1年分で8,259件でした。近年の情報量の多さには驚くばかりです。「GEOLIS 1988」は「日本地質文献目録1988」として冊子印刷されました。B5判904ページ 本の厚さは4cmにもなりました。

GEOLIS は 所内ではオンライン検索が出来ますが 外部の方々はこの分厚い冊子目録を一年毎に探さなければなりません。私たち 作る側からみても 資料を受け入れてから 多くの研究者に論文の選択をしてもらい キーボードをたたいて入力をする といったたくさんの労力をかけてせっかく作り上げたものが 所内だけの利用では 非常に残念であり 何とか外部の方々にも有効利用して頂けないものかと常々望んでいましたところ この程 情報解析課で GEOLIS の FD 化利用を開発する運びになり このたび試作版が完成しました。

利用等の詳細については 17ページの野呂・村田の記事に譲って ここでは GEOLIS の概略について紹介します。

2. 収録の対象

地質調査所地質情報センター資料情報課(通称 資料室)が収集し受け入れた資料(単行本 雑誌 報告書 地図類等)に限定して収録の対象としています。従って GEOLIS で検索した情報源は必ず地質調査所資料室に所蔵しており 資料室で確実に文献に出逢うことが利用者にとって利点になっています。

地質調査所資料室での年間の延べ受入資料数は 表1のように 1万冊を越す雑誌 報告類と数千枚の地図類(国内地形図を含む)となっています。しかも 表が示しているように 購入によるよりも 寄贈や文献交換による入手数の方が 特に国内資料でははるかに上回っています。このことは 一般には入手しにくい資料 購入では入手できない資料 例えば 政府 地方自治体出版物 大学関係 学会・協会関連出版物 等が多く収集されていることを示しています。

地質調査所資料室が入手した資料に限っての収録となっていますので、私たちの収集能力が GEOLIS の質を左右します。このため 私たちは 収集の範囲を広げるため 研究者への情報提供の依頼 資料収集のための予算要求 国会図書館等発行の情報誌の利用など できる限りの努力を続けています。

例えば 文部省が研究補助金を出して推進している研究報告書の類 地方都市の史誌などがあります。また

学会の講演要旨集は14学会（表2）を収録しています。

表1 資料受入数（昭和63年度）

単位：冊・枚

種別	入手方法		購入	寄贈・交換	計
	国内	国外			
単行本	国内		168	163	331
	国外		168	65	233
		計	336	228	564
雑誌	国内		1,299	4,352	5,651
	国外		2,501	2,535	5,036
		計	3,800	6,887	10,687
地図類	国内		106	2,241	2,347
	国外		59	38	97
		計	165	2,279	2,444
別刷等	国内		240	14	254
	国外		48	2	50
		計	288	16	304

表2 収録学会講演要旨集一覧

日本地質学会	三鉱学会
日本鉱物学会	日本火山学会
日本地球化学会	物理探査学会
日本第四紀学会	日本地熱学会
地震学会	日本海洋学会
日本地理学会	地球電磁気・地球惑星圏学会
地学団体研究会	海洋工学講演会論文集

3. 収録の範囲

学問分野としては、地球科学および地下資源に関する情報を収録の範囲としています。具体的には例えば国際十進分類法（U. D. C.）の表現を借りると分析化学（の一部）結晶学 鉱物学 地球科学（地球物理学の一部 気象学・気候学のほとんど全部を除く）古生物学 鉱床の探査と調査に該当する分野としています。そして更に ①日本および日本周辺海域を扱った ②日本人の著者による ③日本で出版された 以上3つの条件のいずれかに該当する論文を収録の範囲としています。しかし「日本人の著者による」という条件は不明確な場合も含まれていることをお断わりしておきます。

1989年8月号

というのは、日本人の名前には日本人か、他国籍の二世かの判断がつかかねる場合もあるからです。

4. 論文の選択

データベース構築開始と同時に所内に所長の諮問機関として「文献データベース委員会」が置かれました。委員長は研究部長があたり、各専門分野の研究者で構成され、事務局は資料情報課に置かれています。委員の最大の業務が GEOLIS 作成のための文献選択です。

次に掲げる7つの専門分野から10名～15名の研究者が選ばれて文献の選択に当たっています。

- 1) 地質学一般 構造地質学 古生物学・燃料鉱床学
- 2) 鉱物学 鉱床学・鉱山地質学 岩石学（無機地球化学の一部）
- 3) 堆積学 海洋地質学
- 4) 第四紀地質 地形学 環境地質（土质地質を含む）岩石力学
- 5) 地球物理学 物理探査
- 6) 火山 地熱 リモートセンシング 情報処理（画像情報処理 情報地質を含む）

文献選択をしている研究者は各自の研究活動のあい間をぬって、資料情報課に足を運び文献選択にあたっています。

選択され入力されたものは8千件前後ですが、実際に目を通す量は、例えば目次だけと限ってみても膨大な労力です。年間12,000冊受け入れると、月間約1,000冊です。論文数にすると月1万件、10人に分担したとしても1人1千件という計算になります。こうした研究者の協力によって、GEOLISは支えられています。

5. 項目の記述

現在 GEOLIS の本体は40の項目から構成されていますが、GEOLIS の公開版（FD版）では利用に不必要な項目は省いています（表3参照）。

各項目の内容定義は項目名通りですので詳しい説明は省略します。GEOLIS 公開版にある項目および使用するにあたって必要と思われる記述内容の説明だけにとどめます。

・文献番号

8桁で以下の形式で記述されています。

Y Y M X X X X X

Y Y 入力した年を西暦年の下2桁

表 3 GEOLIS 入力項目等

△印：部分収録

番 号	入 力 項 目	略 称	属 性	本 体 インバーテット ファイル	F D 版
1	文献番号	NO	英 語	○	○
2	論文題名 (日本語)	TJ	日 本 語		○
3	" (ローマ字)	TR	英 語		
4	" (英語)	TI	"		○
5	著者名 (日本語)	AJ	日 本 語	○	○
6	" (ローマ字)	AR	英 語	○	△
7	" (英語)	AU	"	○	○
8	団体著者 (日本語)	CJ	日 本 語	○	○
9	" (ローマ字)	CR	英 語	○	△
10	" (英語)	CA	"	○	○
11	編集・翻訳者 (日本語)	EJ	日 本 語	○	○
12	" (ローマ字)	ER	英 語	○	△
13	" (英語)	EE	"	○	○
14	資料名 (日本語)	JJ	日 本 語		○
15	" (ローマ字)	JR	英 語		
16	" (英語)	JA	"		○
17	シリーズ名 (日本語)	SJ	日 本 語		
18	" (ローマ字)	SR	英 語		
19	" (英語)	SD	"		
20	巻	VV	"		○
21	号	VI	"		○
22	ページ	VP	"		○
23	発行者 (日本語)	PJ	日 本 語		△
24	" (ローマ字)	PR	英 語		△
25	" (英語)	PB	"		△
26	論文言語区分	LN	"		
27	アブストラクトの言語区分	LS	"		
28	図版	PL	"		
29	発行年	VY	"	○	○
30	資料の種類	CT	"		
31	書誌レベル	BL	"		
32	I S S N	IS	"		
33	I S B N	IB	"		
34	インデックス (日本語)	IJ	日 本 語	○	○
35	" (ローマ字)	IR	英 語	○	△
36	" (英語)	IN	"	○	○
37	発行国	CC	"		
38	図幅番号	QD	"		
39	図書台帳 No.	RG	"		
40	分類番号	LB	"		
41	目録コード	BK	"		

M 入力した月を1～9 A B C

XXXXX 月毎の1連番号

・著者名-日本語

姓名の順でスペースをはさんで記述されています。

例 富士山 昇

・著者名ローマ字

読みはヘボン式ローマ字で記述されています。但し‘ン’を表わす‘N’と‘母音字’または‘Y’を切り離す場合に‘N’の後ろに‘;’〔セミコロンの〕を入れていきます。

例 KIN; EN (禁煙) KINEN (記念)

・著者名-英語

姓名の順で記載され 姓はすべて大文字表記されています。スペルは著者の表記どおりです。

例 FUJIYAMA Noboru

・編集者 翻訳者-日本語

著者名に同じです。

・編集者 翻訳者-ローマ字

著者名に同じです。

・編集者 翻訳者-英語

著者名に同じです。

・団体著者-日本語

論文記述どおりの記載をしています。次の場合は変更を加えています。

a 官庁の場合

省庁名の記述を省略する。但し内部部局の場合のみ そのすぐ上位の省庁名をつける

b 地方庁の場合

都道府県および市町村の名称がついている機関名はそのまま記述する。名称がついていない機関名で 判別に困るものは名称を最初に補記する。

c 大学の場合

大学の付置機関で大学名称が記載されていない時は 大学名を最初に補記する。

d 部課名しかない団体著者名

所属機関名を最初に補記する。

・団体著者-英語

日本語と同じ基準で入力しています。

・論文標題-日本語

論文記述どおりを原則としています。以下の場合は補記をしています。

a 講演要旨の場合

論文名のあとに(演旨)という語を補記する

・論文標題-英語

論文記述どおりを原則としています。以下の場合は補記をしています。

a 講演要旨の場合

論文名のあとに(abs.)という語を補記する

・資料名-日本語

雑誌名 図書名 報告書名を記述しています。図書のシリーズなども含まれています。また雑誌の特集号などに 記載されている 特集名や会議名なども含まれています。記述は収録資料の記述どおりとしています。次の場合は補記または修正をしています。

a 研究報告 研究紀要だけの資料名の場合は 機関名を最初に補記しています。

b 昭和一年度………の形式の資料名

昭和一年度の記述を資料名のあとに(昭和一年度)の形で記述する

・資料名-英語

記述内容は 資料名-英語に同じです。

英文の大小文字の記載は 以下の基準で行っています。

前置詞 冠詞 接続詞を除いた単語の最初の文字を大文字にし 他はすべて小文字にする

・発行者(社)-日本語

・発行者(社)-英語

発行者(社)および発行地を記述しています。

発行者(社)と発行地は、[コンマ]で区切っています。

例 平凡社, 東京

・発行年

発行年を西暦年で記述しています。

・巻

以下のものを巻として数字のみで記述しています。

a Vol. Tom. 巻等に表示されているもの

b 西暦年が区切りとなっている西暦年

c 会議等の開催数

d 図書のシリーズ番号

・号

以下のものを号として数字または英文で記述する。

a no. nor. 号等に表示されているもの

b 巻の記述以外で数字のみの表示で 資料を区分している番号

c Special Issue

・ページ数

ページ数を以下の形式で記述しています。

- a 論文単位の場合
始めと終わりのページを— [ハイフン] でつな
ぎ記述。収録号独自のページと通巻ページが
ある場合は後者を採用する

例 16-36

- b 資料一冊単位の場合
総ページ数を記述

例 286

- c 地図等の場合
シート数を記述する

例 2 sheets

・インデックス-日本語-

・インデックス-英語-

基本は論文題名および資料名より抽出しています
以下の例の場合は補記しています。

- a 地域性があるものは都道府県名を記述
日本語の場合 山形県
英語の場合 YAMAGATA PREFECTURE
- b 著者が付与しているキーワード

6. キーワード付与の問題点

GEOREF も JICST も独自のソーラスを作成して
おり この中からキーワードを付与していますので キ
ーワードの数は極めて多いのです。これはソーラス
から主題の上位・下位語 関連語を自動的に追加できる
仕組みとなっているからです。

しかし GEOLISは残念ながらソーラスを持ちませ
ん。多大な労力と費用を必要とするため作成を諦めざ
るを得ませんでした。

キーワード(インデックス)付与の殆んどは 私たち
図書館員が入力しながら決めていきます。私達は地質
学にはシロウトですので 関連語を補うことができませ
ん。

改善の策として キーワードの付与は 前項で述べま
したように論文題名および資料名から抽出しています
(但し キーワード指定のある場合は指定も加えています)。
はなはだ不完全とは認識しながらも現状では最善の策で
す。

もう一つの問題点は キーワードの定義付けが出来て
いないことです。例えば 論文題名「飛驒変成帯中央
岩体北部岩石学的研究」からキーワードを抽出する場合
“飛驒変成帯中央岩体”か“飛驒変成帯”とするか“変
成帯”だけで別にとる必要があるのかといった問題で日

夜悩んでいるのです。

上記のような問題が解決していませんので 現状のイ
ンデックスは入力者によってキーワードの決め方に差が
生じていますがご容赦下さい。

7. おわりに

3年間 2万件を越す文献の入力を行った今 まだま
だたくさん解決すべき問題を抱えています。最大は
キーワード付与ですが その他に 資料名の表記の統一
化や姓名のヨミや表記の統一化があります。

公開版の日本語の氏名には 英文がない時は ローマ
字読みが入っています。ローマ字付けが違ってしま
うと同一人物が別人と扱われてしまいます。例えば「須
藤」を ストウと読む方もスドウと読む方もいます。
同じく「清野」セイノ キヨノなど。同一人物の泣き
別れを防ぐために人名作業ファイルを作って対処してい
ますが 新人の現れ方の方が早くてファイルの修正が追
いつかないのが実情です。

所内の研究者からは 資料が閲覧室に出てくるのが遅
い といつも苦情が来ています。その通りなので 私
達はいつも「すみません」と謝っているのですが これ
は データベースを作っているのだから収録が終了するま
では閲覧に出せないことから起こってくる問題です。

受け入れ手続きを済ませて 委員の研究者に文献の選
択をお願いするのですが 多忙をきわめておられる方々
が多くて受け入れ資料はすぐに山と積み上げられてしま
うのです。研究者の選択が終了するとやっと入力とな
るのですが 私達も人手不足で 多種の仕事をこなしま
がら入力です。ここでもまた数日かかってしまいま
す。こんな訳で 所内の研究者にお目見えするまで随
分長いこと資料は事務室にとどまっているのが現状で
す。

情報解析課の全面的な支援のもとに ともかくデー
タベースは着々と構築され始めましたが 一番残念なのが
利用が所内に限られていたことでした。しかし これ
も同課の村田泰章氏 野呂春文氏の開発によって解決さ
れることになりました。また GEOLIS 初期より御
協力いただいた同課の佐藤岱生氏の御支援も 頂きました。
心から感謝しております。構築の様々な苦労や
困難は 多くの方々が利用されることによって報われま
す。

所内・外の方々の御協力やアドバイスを今後とも願
いする次第です。