

ランカスターのメドラーズ実績評価を通して見たメドラーズの索引法
MEDLARS Indexing, As Observed Through Lancaster's Evaluation
of the Operating Efficiency of MEDLARS*

松村多美子
Tamiko Matsumura

Résumé

MEDLARS, an acronym of Medical Literature Analysis and Retrieval System, is a computer-based information system developed at the National Library of Medicine of the United States. It has been in operation since 1964, providing bibliographic services to the biomedical communities all over the world. It is in fact one of the largest and the oldest systems of this kind. However, its true significance lies in the fact that the system was subjected to a large-scaled evaluation study conducted in the period from 1966 through 1967 under the directorship of F.W. Lancaster, then the Information Systems Specialist of the National Library of Medicine. The principal objectives of the project were to study the demand search requirements of MEDLARS users and how effectively and efficiently the service was meeting these requirements, and also to recognize factors adversely affecting the performance of MEDLARS in hopes of providing suggestions for means by which the requirements of users can be satisfied more efficiently and economically as well. The test was designed to determine the operating efficiency of the specific aspects of MEDLARS: overall performance, coverage and processing, indexing, index language, searching and input and computer processing. The results of the test and analysis of failures are full of implications.

However, the present paper is concerned only with indexing, the intellectual factors which significantly affect the performance of all information retrieval systems. The MEDLARS indexing is practised by a group of indexers, some with sufficient experience and the others with little of it. As Robert Fairthorne stated, "indexing is the basic problem, as well as the costliest bottleneck of information retrieval."¹⁾ This is particularly true when it is entirely depended upon the intellectual efforts of human indexers. With her own experience as a MEDLARS indexing instructor and reviser, the writer discusses MEDLARS indexing policies and practices as seen through the eyes of Lancaster in his evaluation test.

(Kitasato Memorial Medical Library)

* Lancaster, F.W. Evaluation of the Operating Efficiency of MEDLARS. Bethesda, Md., National Library of Medicine, January 1968. 276 p.

目 次

- I. はじめに
- II. メドラーズ
- III. 評価テストの概要
- IV. メドラーズの索引法
- V. 評価テストに反映した索引法
- VI. おわりに

I. はじめに

機械化の急速な進歩にともなう、電子計算機を利用した情報システムが数多く開発されている。それぞれのシステムによって機械化の程度が異なり、特に情報の入力・出力の形態はきわめて多様化している。メドラーズ (Medical Literature Analysis and Retrieval System) は米国立医学図書館で開発された医学生物学の情報検索システムで、1964 年以来、米国内および世界各地にコミュニケーションのネット・ワークをひろげて活動が続けていることは既によく知られている。しかしこのシステムの意義は、現在稼働中の情報システムとして長い歴史と実績をもつことのほかに、1966 年から 67 年にかけて大規模な評価テストが実施されたことである。評価はシステムの入力・出力の主要な部門について実施され、その結果はその後の改良へ大きな示唆を与えたことはいふまでもない。本稿では、入力部門のうちでも特に索引づけに焦点をあて、この評価テストを通してみたメドラーズの索引法について、筆者自身の米国立医学図書館における索引者としての体験も加味して考察を加えてみたい。

II. メドラーズ

メドラーズ (MEDLARS) については、既に数多くの文献が出されているのでここでは改めて詳しい説明は加えないが、この評価プロジェクトに関連する点だけをとり上げて簡単に解説しておこう。

メドラーズの目的は、(1) 医学生物学関係の雑誌記事索引誌である *Index Medicus* の作成、(2) さまざまな主題分野について定期的に発行される書誌類 (recurring bibliographies) の出版、および (3) 個人を対象とし特定の要求に対して機械検索により書誌を作成するサービス (demand search service) にあるが、今回の評価テストの重点はこの中でも (3) の部分、すなわち特定の要求に応じて過去の一定期間にさかのぼって情報を検索する機

械検索サービスにおかれた。この逆及的検索 (retrospective search) の対象となるデータ・ベースは、50 万をこえる医学生物学関係の雑誌記事論文の文献で 1964 年 1 月以降システムに入力されたものである。このデータ・ベースは評価テストが実施された当時、1966 年から 1967 年にかけては年間 175,000 件の割合で増加していた。現在では年間増加数は 200,000 文献に達し 1969 年 10 月には蓄積文献の総数は 100 万を超えるようになった²⁾。

システムに蓄積される情報の源になる雑誌記事論文は、世界中で発行されている医学生物学の分野の代表的な雑誌約 2,400 タイトルに収録されたもので、その中の約 45 パーセントは英語以外の言語によって書かれたものである。メドラーズ収録誌のリストである *List of Journals indexed in Index Medicus* は毎年 1 月に *Index Medicus* 1 月号の別冊第 2 部として発行されている。

雑誌記事論文の索引は、メドラーズのシソーラスである *Medical Subject Headings* (以下 MeSH と略す) から文献の主題を最も明確に表現する件名標目を選んで与える作業と要約できるが、一文献あたり平均 6.7 個の件名標目が索引語として与えられる。ただし、メドラーズの索引の対象となる雑誌の約 3 分の 1 は “depth journal” に指定されていて、これらの雑誌に収録されている論文・記事はすべて深く完全に索引する。すなわち depth indexing を行なうわけで、この場合には一文献あたり平均 10 個の件名標目を与えることになる。残りの 3 分の 2 の雑誌は “non-depth journal” でその収録文献は前者よりも浅い索引づけ non-depth indexing が行なわれ、一文献あたりの平均索引語数は 4 ないし 6 程度である。

シソーラスである MeSH は、約 7,500 の件名標目とこれをやや上回る数の相互参照から成り、それらをアルファベット順に排列したアルファベティカル・リストと互いに関連ある件名標目を集めてグループにわけたカテゴリー・リストの 2 つのリストから成っている。

カテゴリー・リストでは件名標目は2段階の階層構造に排列されているが、このほかに件名標目をさらに細かく4段階の階層構造に排列した *MeSH Tree Structure* と呼ばれるリストがあり、機械検索の際に使用される。

MeSH には件名標目と相互参照のほかに副標目 (sub-headings) とよばれるものが含まれる。これは biosynthesis (生合成) とか pathology (病理), physiology (生理機能) など医学生物学の主題の一般的な概念を表す語で、件名標目と組合せて使用され件名標目の指示する主題概念をより明確に規定する機能を果すものである。それぞれの副標目は特定の意味内容を持ち、特定のカテゴリーに属する件名標目としか組合せて使用することが出来ない。例えば abnormalities (先天性奇型を表す副標目) はカテゴリーAに属する件名標目 (主として体の部分、器官などを表す解剖学上の名称からなる) とだけ組合せることができるが、一方、congenital (“先天性”の意を表す副標目) はカテゴリーCに属する各種の疾患を表す件名標目としか組合せられない、というようにそれぞれの副標目について規定が与えられている。

Demand search の機械検索は、Honeywell 800 の電子計算機を使用し 50 万以上の文献情報が蓄積されている磁気テープの索引語プロフィールを連続的に検索する方法をとっている。これは根本的にはマッチングの過程であり、雑誌記事論文の文献の索引語のプロフィールに検索式をマッチさせる作業である。

検索式は特定の主題についての利用者からの要求をメドラスのシソーラスの用語に翻訳し、一定の論理式の形式で表現したものである。Demand search の依頼は主として手紙の形式あるいは正式の申込用紙を使用し、郵便によって米国国立医学図書館に直接おくられてくるが、また米国各地に散在するメドラス地域センターで処理される探索要求の大部分は利用者が直接センターを訪れて申込みを行なう場合が多い。今回の評価テストが実施された当時、米国内ではカリフォルニア大学、コロラド大学、ハーバード大学および Natinal Institutes of Health にセンターがおかれていた。検索式は一定のトレーニングを受けた検索者によって MeSH の件名標目および副標目と照合事項を含む各種の索引項目 (tags) をブレー代数式の型式にまとめたものとして作られる。いくつかの概念を包括するような generic な検索はいわゆる “Explosion” とよばれているが、これは件名標目の階層構造表 *MeSH Tree Structure* を使って行なう。例えば、検索式中の “A 9.44.44 Explosion” という部

分は、*Tree Structure* の中で A 9.44.44 のコード番号に相当する件名標目 RETINA (網膜) の下に包含されるすべての件名標目、すなわち、FUNDUS OCULI (眼底), MACULA LUTEA (黄斑), RODS AND CONES (杆状体および錐状体) をもうらする検索が行なわれることを指示している。メドラスの機械検索では、サブ・ソートによって検索された文献のうちで要求内容に最も適合したものからその周辺領域のものへと3段階の異なった適合度をもつグループにわけることが出来る。書誌が最終的に電子計算機から打ち出される時は、紙面あるいはカードのいずれの型式に於ても、最も適合度の高い文献のグループが最初に印刷され、次いで2番目のグループそして最後に3番目のグループという順序をとる。各グループの中では、文献は著者名のアルファベット順に排列されるのが普通の方法であるが、特に利用者の希望によって文献の収録誌名の順による排列や文献の発行年月日による排列方法などをとる場合もある。

このサブ・ソートの手法はすべての要求に対して無差別に適用されるものではなく、要求質問の内容に応じて最も効果的な検索ができるように、探索式をたてる際に検索者が判断していくつのグループにわけられるか決定される。Demand search の書誌でそれぞれの文献について得られる情報は、論文・記事の著者名、標題、収録誌の巻、号数、発行年月日および頁数などの書誌的事項のほかに、その文献の索引づけに使用されたすべての索引語の記録 (tracing) が含まれる。これには MeSH の件名標目、およびそれに組合せられる副標目ばかりでなく照合事項 (check tags) も含まれ、また件名標目の中ではそのいずれが *Index Medicus* に見出し語として使用されたかが * 印によって指示されている。

II. 評価プロジェクトの概要

1. 目的

ランカスターは評価テストの目的を次のように要約している。

- (1) メドラスの機械検索サービスへの要求を調査する。
- (2) これらの要求に対して現在のメドラスがどの程度効果的にまた効率よくサービスを行なっているか調査する。
- (3) システムの実績に逆効果を与えている因子を洗い出す。
- (4) メドラスの利用者の要求がより効果的になおか

ランカスターのメドラーズ実績評価を通してみたメドラーズの索引法

つ経済的に充たされるような方法を見出す。特に次代の電子計算機のプログラムの開発や備品の選定が最も効果的に行なわれるような方法を提示する。

特定の情報システムの評価においては、そのシステムに関連するいろいろな要素が評価の対象として考えられるが、メドラーズの評価テストでは利用者の要求に最も関連あるものとして、次のような要素が取り上げられた。すなわち、

- (1) メドラーズの収録範囲……特定の主題に関してシステムに蓄積されている有効情報の量で、これには情報がシステムに入力される迄に要する時間、すなわち索引作業に必要とする時間の制約がある。
- (2) 再現性 (recall) ……有効な情報の検索能力で、このテストでは特に利用者の特定の要求内容に対して有効な価値を有する情報であることはいふ迄もない。
- (3) 適合性 (precision) ……有効でない情報が検索されることを防ぐ能力。
- (4) 反応時間……メドラーズ・センターで要求を受理してから利用者に完成した書誌が渡されるまでに要する時間。
- (5) 検索の結果として作成された書誌の形式。
- (6) 利用者がメドラーズから満足すべきサービスをうけるために、どの程度の個人的な努力を支払わねばならなかったか。

などの点である。評価テストはこれらの諸点に関して利用者の要求と満足度を設定できるように設計されたが、その上、システムのいろいろな部門で特定の問題について解答がえられるように考慮が払われた。

索引部門では、次のような事柄に焦点があてられた。

- (1) 索引者の間に作業上の大きな差異がありうるかどうか。
- (2) もしあるとしたら、それが索引の経験と“リバイズ” (revise) の程度にどれほど関係があるか。
- (3) 索引者はさまざまな利用者グループのそれぞれにとって興味ある特定の主題を認識するかどうか。
- (4) 現在の索引に関する方針はシステムの実績にどのような影響を与えているか。特に索引の対象となる雑誌を depth indexing と non-depth indexing の2つの異なった深さの索引づけのグループに區別していることが、検索結果に大きな相違をもたらしてはいないだろうか。また、*Index Medicus* の見出し語に使用される件名標目だけで検索した結果はど

うであろうか。

2. 結 果

1966年8月から1967年7月までの1年間に米国国立医学図書館およびコロラド大学、ハーバード大学、カリフォルニア大学ならびに National Institutes of Health の各センターで実際に処理された探索要求の中から302件のテスト・リクエストが選ばれた。このうち適合率 (precision ratio) と再現率 (recall ration) の両者がえられたもの299件について最終的な結果をまとめたところ、メドラーズは平均57.5パーセントの再現率、50.4パーセントの適合率で運営されていることが明らかにされた。

更に302件のテスト・リクエストについて分析を行なったところ、238件において検索もれを生じたことがわかり、文献数にして797件の有効文献が何らかの理由で検索されなかった。また適合率については、278件の検索例でもれが起り、3,038件の文献が有効でないと判断された。これらの結果をさらに詳細に分析したデータが再現率については表1に適合率に関しては表2にそれぞれ示されている。再現率ならびに適合率の低下に関連ある種々の要因として、これらの表にも指示されているように、索引用語、検索法、索引法、システムと利用者間の相互作用および電子計算機による処理の過程があげられる。

表 1 再現上のもれの原因

302 検索例のうちの 238 例に生じた 797 件の検索もれの原因

原 因	検索もれの文献数	総ての再現上のもれの比率	検索件数	238 検索例の比率
索 引 用 語	81	10.2	29	12.2
検 索 法	279	35.0	133	55.9
索 引 法	298	37.4	203	85.3
specificity の不足	46	5.8	31	13.0
exhaustivity の不足	162	20.3	100	42.0
exhaustivity の過剰	5	0.6	4	1.7
索 引 も れ	78	9.8	61	25.6
不適当な索引語の使用	7	0.9	7	2.9
コンピュータ処理	11	1.4	7	2.9
システムと利用者の相互作用	199	25.0	70	29.4

表 2 適合上のもれの原因
302 検索例のうちの 278 例に生じた 3,038 件
の適合上のもれの原因

原 因	不適合文 献 の 数	総ての適 合上のも れの比率	検索件数	278 検 索 例の比率
索 引 用 語	1,094	36.0	255	91.7
検 索 法	983	32.4	186	67.0
索 引 法	393	12.9	167	60.1
exhaustivity の過剰	350	11.5	137	49.3
exhaustivity の不足	5	0.2	2	0.7
specificity の 不足	1	0	1	0.4
索 引 も れ	1	0	1	0.4
不適当な索引 語の使用	36	1.2	26	9.4
コンピュータ処 理	3	0.1	3	1.1
システムと利用 者の相互作用	503	16.6	111	39.9
価 値 判 断	71	2.3	40	14.4
不適合文献のさ けられない検索	4	0.1	4	1.4

本稿では、この中で特に索引法をとり上げて考察を加えてゆくが、本論に入る前にメドラスの索引法の原理と実際について解説しておく必要があろう。

VI. メドラスの索引法

1. 概念整合 (Concept coordination)

メドラスの索引法は、概念整合 (concept coordination) の原理にもとづいた対応づけの索引法 (coordinate indexing) であるが、その具体的な方法としては次の 4 種類がある。

(1) 2 つ以上の件名標目を整合させる

例えば、“喫煙の統計学的研究”

(喫煙 (SMOKING)

統計 (STATISTICS)

“寒天ゲル電気泳動法”

(電気泳動法 (ELECTROPHORESIS)

寒天 (AGAR)

ゲル (GEL)

(2) 件名標目と副標目

例えば、“関節炎患者の肝臓代謝作用”

(関節炎+代謝

(ARTHRITIS *metabolism)

肝臓+代謝

(LIVER *metabolism)

* 印以下の語が副標目

副標目は、前にもふれたように、MeSH の件名標目によって表わされる概念の特定な面を指示する機能をもつものである。メドラスが稼動を開始した 1964 年には使用されていなかったが、1966 年に導入されその年には 42 の副標目が、そして 1967 年には 53 の副標目が MeSH に含まれていた。1970 年現在ではその数が 60 に増加したが、これも今回の評価プロジェクトの結果、これら副標目がロール・インディケーターと同様な機能を果し、索引用語としての重要性が認識されたからに外ならない。MeSH の副標目の特徴は、それぞれが固有の意味内容をもっているばかりでなく一定のカテゴリーに属する件名標目とだけ組合せて使用することが出来る点にある。上例中の metabolism は、A.B.C.D.E.F. の各カテゴリーに属する件名標目と組合せられるが、ARTHRITIS はカテゴリー C に、LIVER はカテゴリー A にそれぞれ属するので、いずれの場合も正しい組合せといえる。MeSH の件名標目はすべて一つ或いはそれ以上のカテゴリーに属していて、このことは MeSH のアルファベタイカル・リストのそれぞれの件名標目の後に表示されている。

(3) 件名標目と照合事項 (Check tags)

例えば、“児童における肝炎”

肝炎 (HEPATITIS)……件名標目

児童 (CHILD)……照合事項

人間 (HUMAN)……照合事項

“犬に於ける実験的肝炎”

肝炎 (HEPATITIS)……件名標目

犬 (DOGS)……照合事項

動物実験 (ANIMAL EXPERIMENTS)……

照合事項

照合事項とは、一言でいえば、索引の対象となる文献の一つ一つについて必ず照合されるべき事項で、メドラスの索引作業に使用される所定のワーク・シート (data form と呼ばれる) に予め印刷されている。索引者は文献の内容を分析し、主題を選択する際に照合事項のそれぞれの項目について検討し、がい当するものに指定のマークをつける。照合事項に含まれるものは、(1) Review (総説および引用文献数)、(2) English abstract (英文抄録の有無)、(3) Pregnancy (妊娠に関連あるか否か)、

(4) Age groups (文献中に扱われている対象の年齢層), (5) Experimental animals (実験動物の種類), (6) History of medicine and chronological tags (医史学および歴史的著述の時代区分), (7) miscellaneous tags (その他, 研究のタイプなどを表す項目) などである。

(4) 複合標目 (Pre-coordinated headings)

対応づけの方法の (1) として二つ以上の件名標目によるものをあげたが, この方法で組み合わせて同時に使われる頻度の高い二つの件名標目は, 前もって整合された precoordinated heading として MeSH の件名標目になっている。この場合, LIVER GLYCOGEN (肝グリコーゲン) のように LIVER (肝臓) と GLYCOGEN (グリコーゲン) の二つの件名標目をそのままの形で整合したものや, HEPATIC COMA (肝性昏睡) の場合のように LIVER と COMA のうちの LIVER が HEPATIC という形容詞形をとるものもあり, さらに TUBERCULOSIS, PULMONARY (肺結核) のように LUNG と TUBERCULOSIS の転置形をとる場合もある。

2. 索引の対象

メドラスが対象とする主題分野は医学生物学であり, 索引の対象となる文献は世界中で発行されている医学生物系の雑誌約 2,400 タイトルに収録されている論文記事である。それぞれの雑誌については, もうら的に索引するという原則があるが, 実際には次のようなガイドラインが適用される。³⁾ すなわち,

- (1) 原著論文はその長さや形態のいかんにかかわらず索引する
- (2) 医学生物学関係の人物の伝記および死亡記事で, 新聞のニュース以上の実質的な内容をもつもの
- (3) 臨床・病理討論会の記事
- (4) 学会で発表された科学的資料の報告類
- (5) 事実上, 論文とみなされるような論説記事
- (6) 実際には論文と考えられるような実質的な内容をもつ“読者の声”(letters to the editors)。特に薬品の副作用についてはこのような欄で貴重な報告が見出される場合が多いからもちろらず索引する
- (7) 学会の住所録や会員名簿
- (8) シンポジウム, パネル, 円卓会議などで明確な主題をもつものは公式, 非公式の別なく索引する

3. 主題の選択と件名の選定

Index Medicus を始めとして各種の recurring bibliographies や demand search の書誌類を含めてメドラスは主題から情報を探索するシステムである。主題索

引は, Vickery がいっているように, “ある特定の主題に関する凡ゆる情報の存在とその存在場所を示すものであり……情報の探索者と求めている情報との間の連絡を正確にもうら的になおかつ迅速に確立する機能をもつものである。”⁴⁾ このような主題索引の作成過程を Bernier は次の諸段階にわけている。⁵⁾

- (1) 著者が報告している主題をドキュメント中から選択する
- (2) 抽出された主題を, 著者の表現をそのままかたり或いは云いかえたりして明確に表現しなおす
- (3) 主題に対してガイドとなる語句をえらぶ
- (4) 必要とあればこれらのガイドとなる語句を基準として使用している索引用語に翻訳する

文献の内容を分析し主題を選び出す作業は, 索引作成の重要な要素であるにもかかわらず幾多の問題点を含んでいる。メドラスの索引に於てもまた一般的にもいわれている原則は, 情報の生産者すなわち著者が報告している主題を, 現在および将来おこりうるであろう要求を考慮して分析する。しかしながら, “主題”とは何かという大きな問題点が核心にあり, また情報に対する要求は特定の利用者の特定な時点に於ける特定な状況によって異なるものであるから不特定多数の利用者を対象にした場合はその多様性は無限に大きくなる。メドラスの索引法で索引者が主題の選択に当たって念頭におくべき事柄としてあげられている項目は,⁶⁾

- (1) その論文中のどんな概念が *Index Medicus* およびメドラスの利用者にとって有益であるか
- (2) どんな概念が将来利用されうるであろうか
- (3) 文献に含まれる事実あるいはそのどんな部分が文献の主眼点ではないにしても関連ある研究にとって価値あるものであろうか
- (4) どんなタイプの研究者がその文献に価値を見出すであろうか
- (5) 主題が既存の索引用語で適切に表現できるであろうか

などである。これらの事項を念頭において文献の内容を検討するが, この場合に文献の全文をよんでいたのでは時間がかかりもとと非常に費用のかかる索引作業にさらにコストを加算する結果になる。従って文献の全文をよまずにそれと同じ効果をあげる目的で, 索引者はまず文献の標題を注意深くよむ。ついで序文から著者の意図するところを理解し, 標題とてらし合せて文献の内容を確認する。次に内容の項目見出しや図表などがあればそ

れらに目を通し、最後に結論をよんで著者の目的が達せられたか否かを判断する。著者自身あるいは第三者の書いた抄録あるいは要旨が与えられていれば標題の検討以外の過程はある程度省略ないしは短縮できる。いずれの場合にも、主題はもうら的に索引するというのが根本原則である。

以上のような方法で主題が把握されたらそれを構成要素に分析し、ガイドとなる語句をえらびさらにそれらのガイドを MeSH の件名標目あるいは照合事項の適当な tag におきかえる作業が行なわれる。この場合の原則は最も specific な索引語を使用することであり、この目的のために MeSH の階層構造や“より特殊な……をも見よ”(see also specific……)の相互参照が与えられている。ところで、いくつ位のガイドを使用すれば一つの文献中の主題を完全に表現することができるかということとは、(1) 主題をどの程度もうら的に云い表わすかということと、(2) 主題を表現するのに(或いは主題のそれぞれの構成要素を表現するのに)いった方が適当かも知れない)いくつのガイドを必要とするかという二つの条件によって決ってくる。この主題選択のもうら性 exhaustivity と索引語の specificity は今回の評価プロジェクトに於て索引に関連する部分の考察の焦点になっている。

4. Depth indexing と non-depth indexing

主題の選択のもうら性に関連して、メドラス索引法の方針の一つである depth indexing と non-depth indexing について説明しておく必要があろう。索引の対象となる雑誌が指定されていることは既にのべたが、これら約 2,400 タイトルの雑誌はさらに depth indexing を行なう depth journals と non-depth indexing を行なう non-depth journals の 2 つのグループに分けられている。それぞれのグループに指定された雑誌に収録されている論文記事は、その長短や内容のいかんにかかわらず depth indexing か或いは non-depth indexing の方針に従った索引づけを行なう。

Depth indexing は、文字通り深く索引する方法で文献中の主要な主題ばかりでなく完全にもうら的に索引する。具体的には、一つの文献について平均 15 から 20 の索引語を与える。Non-depth indexing は、これに反して索引の対象を主要な主題に限定するもので、一文献あたり平均 3 から 5 の索引語を与えるのが普通であり、depth indexing に比較してはるかに少ない。しかしいずれの索引法においても、IM 標目の方針は変わらず適用

されることは注意すべきである。IM 標目については次項で解説するが、このことは索引誌 *Index Medicus* を使用して探索する場合には同一文献が depth indexing で索引されていても 或いは non-depth で索引されていてもそのことに関係なく探索されうることを意味している。しかし demand search の機械検索に於てはこの索引法の区別が検索結果に影響を与えることになる。

5. IM (*Index Medicus*) 標目と NIM (*Non Index Medicus*) 標目

メドラスの索引作業によってシステムに蓄積された情報は、索引誌 *Index Medicus* のほかに demand search service のための機械検索ならびに各種の recurring bibliographies のデータ・ベースとして使用される。従ってこれらの目的の違いは索引作業の実際にさまざまな形で反映しているが、その一つが IM 標目と NIM 標目である。文献の内容の評価・分析の結果いくつかの件名標目が MeSH から選定されると、これらの件名標目の一つ一つについて評価が加えられ *Index Medicus* の見出し語として使用するかどうか、すなわちその件名標目の下に文献のサイテーションが印刷されるか、或いは索引語として磁気テープに蓄積され機械検索に使用するかかの決定がなされる。*Index Medicus* に見出し語として使用される件名標目を IM 標目 (*Index Medicus* headings) とよび、一方テープに蓄積される件名標目を NIM 標目 (*Non Index Medicus* headings) と呼ぶ。ただし一つの文献に与えられた件名標目は IM 標目も NIM 標目も含めてすべて磁気テープに蓄積されることには変りない。

そこで IM 標目を決定する基準であるが、文献の主題のポイントを指示する件名標目ということになる。これに反して NIM 標目は主題の二次的あるいは三次的な要素を指示する件名標目で主題のポイントを規制したりあるいは特定の観点を表わすなどより特殊なきめのこまかい情報検索に適した手がかりとなるものである。一つの文献についての IM 標目の数は depth indexing あるいは non-depth indexing の区別なく平均 2.6 個の件名標目である。この IM 標目は、いわば weighted term の機能をもつもので機械検索上の手法の一つとして使用することができる。

V. 評価テストに反映した索引法

1. Exhaustivity と Specificity の定義

評価テストの結果の考察に入る前に、二つの重要な語

exhaustivity と specificity について今回のプロジェクトに於けるこれらの語の意味内容を明確にしておく必要があろう。

Exhaustivity と specificity は共に索引づけと検索方式の両方に適用されるが、索引づけにおける exhaustivity とは主題分析の段階で文献中に含まれる主題をどこまで索引づけの対象にするかという範囲をさす。例えば、抗生物質を使用して5種類の皮膚疾患を治療した臨床報告の文献で、もしこの中の一種類の皮膚疾患を索引しなかったならば、この索引づけは exhaustive であるとはいえない。これに反して、例えば20人の患者の症例報告の文献で、その中で扱われているすべての疾患、診断法、治療法ならびに使用された薬品などを索引した場合には文献中の主題は exhaustively に索引づけされていると考えられる。索引づけの exhaustivity が高ければ検索時の再現率は高くなる傾向にあるが、その反面適合率は低くなる。すなわち、文献中に含まれる主題の最も重要なポイントのみにしぼった exhaustivity の低い索引づけを行なうと、適合率は高くなるが再現率は低くなるといえる。

索引づけの exhaustivity は、メドラスのシステム運営上の方針の一つとして決定されるもので、具体的には一つの文献に与えられる索引語の数が指示されていて原則的な exhaustivity のレベルを設定している。この指示の範囲内で、索引者は文献中の重要と思われる主題を表現する件名標目を選択するわけである。ある特定の概念が検索されないという再現上のもれは、exhaustivity の欠除により、その概念が索引づけされていないことによる。同様にして不必要な非有効文献が検索されるのは、exhaustivity の高い索引づけの結果としてあまり重要でない二次的な概念が索引されているからである。いかなるタイプの文献にも共通してあてはまるような絶対的な exhaustivity のレベルなどというものがないことは明らかである。しかしながら、特定のシステムとして望まれる最適のレベルは当然考えられるものである。

Specificity は、主題の構成要素を索引づけするレベルを指すものといえる。例えば、“とらふぐのトキシン”は“とらふぐトキシン”という特定な一語で表現できるばかりでなく、“トキシン”と“とらふぐ”という二つの語の組合せとしても表現できるし、さらに一般的なレベルでは“トキシン”と“魚”の組合せによって索引づけることも可能である。また更に広義のレベルでは、“動物の

トキシン”或いは“トキシン”と“動物”の組合せで索引することになり、さらに進んでより一般的な“トキシン”で表現することも可能である。このように特定な概念を索引づけするのにさまざまなレベルが考えられる。

索引づけの specificity が高ければ検索結果の適合率が高くなるのは当然のことであり、その反面再現率が低下することもまた事実である。索引づけの specificity に起因する再現率の低下を防ぐためには、検索式をたてる際に何らかの手段を講じることが必要になる。すなわち、検索の対象をより広義なクラスにまで拡大し有効情報を含む可能性のある関連文献をもらさず検索することである。これに反して、specificity の充分でない索引づけの欠点は検索上のいかなる手法をもってしても埋め合わせが出来ないことに注意する必要がある。

索引づけの exhaustivity は一般的にいてシステム運営上の方針による決定と個々の索引者の判断決定との二つの要素によってきまるものであるが、一方、索引づけの specificity は主として索引用語に左右される。前例の“とらふぐのトキシン”をとり上げてみても、もし“とらふぐトキシン”という件名標目が存在すれば問題はないが、それが不存在の場合には“とらふぐ”と“トキシン”の組合せで表現することになり、それもないければ、より広義な“魚のトキシン”さらに“動物のトキシン”から単に“トキシン”へとより一般的な件名標目を使用せざるを得ないことは明らかである。しかし実際には索引用語の問題とは別に、索引者がいくつかの概念を索引する場合に個々の概念をそれぞれの件名標目で表現するかわりに、それらを包括するようなより広義な概念を指示する件名標目で代表させてしまうことがありうるし、また索引用語に充分なれていない未経験な索引者が specific な件名標目の存在に気づかず誤ってより一般的な件名標目を使ってしまうこともしばしば起こる。従って今回の評価テストでは、索引用語に起因する specificity の不足と索引づけに於けるそれとは明確に区別されており、本稿でとり上げるのは後者の場合である。

2. 索引づけの失敗

表1および表2から明らかなように、索引づけに起因する再現上のもれは37パーセントで、他のいかなる因子よりも高い比率を示しているが、一方適合上のもれは13パーセントとはるかに低い。ランカスターは索引づけの失敗を(1)索引者の誤りによるものと(2)一文献あたりいくつの索引語を与えるかについての索引方針によるものの二つのタイプに大別している。

索引者の誤りには第一に重要な主題の索引もれがあげられる。この場合の索引もれとは、主題の重要な要素であるにもかかわらず索引によって表現されなかった場合をさし、non-depth indexing に於ても当然索引されるべき重要な概念がもれているといった場合である。索引者の誤りによる索引もれは、索引法の方針にもとづく exhaustivity の不足と明確に区別さるべきものである。両者ともに再現率の低下の原因となりうるが、exhaustivity の不足は主題の末梢的な要素であるために一文献あたりの平均索引語数の制約をうけたり、或いは一文献を平均 15 分で完全に索引する（記述索引の部分も含めて）という非常に限られた時間などの理由により索引づけから除外された結果である。

索引者の起こすその他の誤りとしては、不適当な件名標目の使用があげられる。今回のテストではこのタイプの誤りの結果は再現上のもれならびに適合上のもれの両者の約 1 パーセントにしかすぎないが、これに反して索引もれは再現上のもれの全体の約 10 パーセント近くを占め、再現もれの生じた検索例のうちの 25 パーセントに於て最低限一つの検索もれ文献を起す原因になっている。文献中の重要な概念が索引もれになると、その文献に含まれる主題が適合するような検索ですべて検索もれになり、システムに死蔵されてしまう結果になる。例えば、このプロジェクトで“水晶体”に関する検索要求が 2 件あったが、そのいずれに於ても *Science* に収録されているサンショウ魚の水晶体に関する文献が検索されなかった。その理由は、“水晶体”(LENS, CRYSTALLINE)が索引されていなかったという簡単な理由に外ならない。索引もれは不適当な索引語の使用の場合と異なり、リバイザーの制度をもってしても誤りがそのまま見逃されてしまいやすいことも現在のシステムの問題点の一つといえよう。索引もれは通常、再現率の低下をもたらすが不適当な索引語の使用は適合率と再現率の両者の低下をきたす可能性がある。

索引もれの代表的な例として“ねずみの視覚皮質の発育に及ぼす視力奪取の影響”の索引づけがある。この主題は、“大脳皮質”(CEREBRAL CORTEX)、“視覚”(VISION)、“暗闇”(DARKNESS)、“ねずみ”(MICE)の索引語を使用して索引されているが、このほかに“発育”(GROWTH)と“感覚奪取”(SENSORY DEPRIVATION)の索引語も加えられるべきである。MeSH には“視覚皮質”という特殊な件名標目がないのでそれを包含するより一般的な件名標目として“大脳皮質”(CERE-

BRAL CORTEX)を使用するのは正しいが、“視力奪取”の部分は“感覚奪取”(SENSOR DEPRIVATION)と“視覚”(VISION)を整合させて始めて現存する索引用語の範囲内でこの概念を充分に表現することができたといえる。また“発育”が全く索引づけからもれていたのは検索の際に致命的な影響を与えたと考えられる。また“メドロキシプロジェステロンの使用中止後に起った一時的な不妊症”に関する文献では、主題の最も重要な構成要素の一つである“不妊症”を表わす件名標目 STERILITY, FEMALE が全く使用されていなかったために“不妊症”に関するすべての検索例でこの文献は検索されなかった。

これらの索引もれの起こる原因としては、非常にきびしい索引作業スケジュールが考えられる。米国立医学図書館の索引課では、各索引者はそれぞれに統計用紙を渡され、毎日いくつの文献を索引したかを記録し報告することを義務づけられている。講義および演習を含む正規のトレーニングを受けた索引者はそれぞれリバイザーがつき、その指導のもとに実際に索引作業を行なうが、数か月の経験をへると一文献あたり平均 15 分のスピードで索引することが要求される。これには主題を分析しシソーラスから適当な索引語を選定する主題索引ばかりでなく、著者名その他の書誌的事項の記述と英語以外の言語による文献の場合には標題を英訳する作業が含まれるので、この平均速度をたえず保ち続けて作業をつづけるのは想像以上にきびしい仕事であることは事実である。従ってこのようなプレッシャーが原因となって索引もれを生じることは容易に考えられるが、それとは別に、がい当する件名標目がなく、その上すでにある件名標目のどれを使って索引したらよいのかという指示が何も与えられていない場合が多く、これもまた索引もれの大きな原因になる。このような場合に、索引者はその概念を全く索引しないか或いはしたとしても極めて一般的な大まかなレベルでしか行なわない。例えば、“サイロキシンの光によるイオン消失”(photodeionization)という主題を表現するのに“光によるイオン消失”という件名標目は MeSH に含まれていないし、“イオン消失”(deionization)を表す件名標目も与えられていない。そこでこの概念を表現する既存の件名標目としては“ヨウ素”(IODINE)が最も適当であるが、この場合に“光によるイオン消失”ならびに“イオン消失”などの概念はどの件名標目の下に索引するという指示が何らかの形ですべての索引者に行きわたっていることが必要である。さ

もなければ、索引者はこの概念を完全に無視して索引づけを行わないか、或いは各自がそれぞれ独自の判断に基づいて異なった件名標目を使用する結果になり、索引者の間ばかりでなく同一の索引者に於ても一貫性が保たれない結果になる。MeSHに適切な件名標目がない場合にどれを使って索引するかという指示はある程度までは典拠カードのファイルによって与えられている。しかしこの典拠ファイルはもうらめではなく、また一貫性をもってその維持に特別の注意が払われているわけではない。もともと米国国立医学図書館は各種の書誌的活動に長い歴史と伝統をもち、索引作業の中心となっているいくつかの索引者はいずれも10年以上の経験を有しMeSHの意味内容を殆んどすべて体得しており、これらの人々がリバイザーとして経験の浅い索引者を指導し、MeSHの定義を伝授し、上記のような場合に特定の概念をどう索引するかを指導する。しかし前述したようなきびしい生産スケジュールもあり、未経験な索引者は疑問が起こるたびごとにリバイザーをわざわざ呼ぶことをちゅうちょする傾向があり、一方典拠ファイルは完全ではないところから索引もれが生じる結果になる。

索引もれに比較して“誤った索引づけ”が少く再現上のもれおよび適合上のもれのいずれに於てもわずか1パーセントにすぎないという事実は、先にもふれたリバイザーの制度によるものであらうとランカスターは指摘している。索引者は一定期間の公式な訓練を受けた後は一本立ちになる迄の間は経験の長い索引者がリバイザーとして指定され、索引するすべての論文についてひとつひとつ照合し誤りを訂正し指導を行なう制度になっている。

リバイズの対照となる項目には、著者名のとり方、頁数の表記法その他の記述索引の部分から件名標目および副標目の選定、照合事項 (check tags) の正否に到るまで索引者の責任において記入されるべきデータ・シート上の項目はすべて含まれる。従って、このリバイズという過程で誤った索引語は比較的目立ちやすく、それだけに訂正される可能性が大きい。その反面、索引もれは発見がむづかしくそのまま見過されてしまう場合が多いといえる。

3. 索引づけの exhaustivity

索引づけの exhaustivity が高くなればなるほどそれに比例して再現率は高くなるが、同時に適合率は逆に低くなることはこの両者が逆比例の関係にあることから明らかである。

いかなる情報システムに於てもこの原理が程度の差こそあれ当てはまることは既に幾多の研究から明らかなことである。

今回のテストの結果、再現上のもれの 20 パーセントは exhaustivity の不十分な索引づけに原因し、また適合上のもれの 1.5 パーセントは主として exhaustivity の高すぎる索引づけによるものであることが明らかにされた。ここで索引づけの exhaustivity に関連して depth indexing と non-depth indexing の索引方針についてもう一度考えてみる必要がある。索引づけの exhaustivity とは前にも述べたように文献中の主題をどこまで索引するかということであり、具体的には一文献につきいくつの索引語を与えるかということである。メドラスの収録誌はすべて depth indexing か non-depth indexing のいずれかに指定されている。雑誌総数の約 3 分の 1 を占める depth journal では収録文献はすべて深く完全に十分な exhaustivity をもって索引され、索引語の数は一文献あたり平均 10 個となっている。一方、non-depth journal では一つの文献について平均 4 個の索引語が与えられる。メドラス全体としての平均索引語数は一つの文献について 6.7 個である。収録誌をこのように索引づけの実際が異なる二つのグループに区別することは、当然の結果として与えられる索引語の数に大きな差をもたらすことになる。従ってこの区別はメドラスの実績を検討する場合に除外することのできない要素である。

評価テストの期間中の depth indexing と non-depth indexing の文献数の割合を推定する目的で、ランカスターは小規模な実験を行なっている。その結果は表 3 に示されているが、明らかに両者の比は変化し depth indexing の占める割合が年と共に高くなってきている。また、depth と non-depth の両方の索引づけを通じて一文献あたりの索引語の数が増加している。すなわち索引づけの exhaustivity が高くなってきていることもこの表から明らかにされた事実である。1966 年 8 月から

表 3. Depth indexing と Non-depth indexing の比率

年	Depth と Non-depth の文献数の比	1 文献あたりの 平均索引語数	
		D	N D
1964	42 : 58	7.0	5.9
1965	54 : 46	7.6	4.2
1966	58 : 42	8.3	5.9

1967 年 7 月にかけての評価テストの期間中の depth indexing と non-depth indexing の文献の割合は 55:45 と考えるのが妥当であろうとランカスターはのべている。

3.1. Exhaustivity の不足による再現上のもれ

この場合の典型的な例として“麻薬中毒者に於けるウイルス性肝炎”に関する文献がある。この文献中の肝炎は注射器具の汚染によって起こされたものであるにもかかわらず、その感染原因が索引づけられていなかったために“血液あるいは静脈穿刺以外の非経口的接種によるウイルス性肝炎の伝播”に関する検索例では最も有効な文献の一つであるのに検索されなかった。“非経口的注入”を表す件名標目 INFUSIONS, PARENTERAL が索引されているならばこの文献は正しく検索されたはずである。このように明快な例は別として、単に exhaustivity の不足とはいっても通常の索引づけではやむを得ぬ場合が多くあることもまた事実である。例えば、“高血圧と言語障害の関連性”についての唯一の有効文献が検索もれになったが、これは 25 頁の音声行為に関する総説中に含まれたたった一節であった。このような場合に、総説中のすべての文節を索引づけするような極端に exhaustivity の高い索引法をとらない限り再現もれを生じることが当然の結果であるし、このような索引方法は通常のメドラス索引作業では考えられない。しかし今回の評価テストではこのような場合も含めて画一的に索引づけの exhaustivity の不足と見做されている点は、数値の解釈の上で留意しておくべきであろう。

3.2. Exhaustivity の過剰と適合上のもれ

Exhaustivity の高い索引づけは、適合率の低下に大した影響を与えない場合と、また時として大きな原因となりうる今回のテストから明らかにされた。例えば、“乳癌あるいは前立腺癌の患者における血中および尿中ステロイド”に関する検索例では、過剰な索引づけの結果として次の二つのタイプの不適合文献が検索され、適合率の低下をきたした。すなわち、(1)“癌”という概念が殆んど含まれていない文献で、例えば血中アンドロゲンについての文献で多数の対象患者のうちのたった一人がたまたま乳癌であったというだけで、その文献の主要な主題とは全く考えられないのに索引づけされたもの、(2)“尿”の概念が重要でない文献、例えば“前立腺”と“尿”の二つの概念の論理積の結果として検索された文献の中には、前立腺癌の患者の尿中ホルモンの定析に関するものが含まれていた。

過剰な索引づけが適合上のもれに決定的な影響を与えた例としては、“クロランフェニコールの作用”に関する検索例がある。ここでは検索された文献の 50 パーセント以上で、クロランフェニコールがたしかに言及されている等、文献の主題とは考えられないものが大部分を占めていた。通常の索引法では、文献の主題が培養基あるいは培養の手技、手法に関するもの、或いは著者の特別の意図がおかれていない限り、日常ごく普通に行なわれる検査については索引の対象としないのである。同様な理由から適合率がわずかに 15 パーセントという低率を示した検索例は“尿管のクレアチニン分泌”に関する検索例で、非有効文献の大部分で言及されているクレアチニンは例えば常用の肝機能検査でえられるクレアチニン値についてであった。尤もこのような場合には、検索式の中で CREATINE AND CREATININE の件名標目を IM 標目として指定することにより再現率を保ちながら同時に適合率をあげることが可能である。IM 標目は *Index Medicus* に見出し語として使用される件名標目であり従って文献中に扱われている最も重要な主題を指示しているわけで、従ってクレアチニンが主題のポイントである文献に限定することが出来る。

Exhaustivity の高い索引づけのへい害の顕著な例としてあげられるのが、“コンピューター”に関する索引づけの方法である。文献中にコンピューターが言及されている場合はすべて、COMPUTERS 或いはその specific な件名標目を使用して索引されている。従って、例えば“コンピューターによる細胞の認識”という検索では、検索された文献の半数以上が単なる計算の手段としてのコンピューターの使用に関するもので、もともとの要求内容からはずれた非適合文献であった。

地名を表す索引語もまた限度をこえて使用されていることが指摘された。地名は MeSH の正式な件名標目ではないので *Index Medicus* に見出し語として使用されることはないが、demand search の機械検索の対象となるデータ・ベースには索引語として含まれている。

この地名が索引に使用されるのは、医学生物学に関連ある人物の伝記、死亡記事などの場合にその人物の出身国を表すために索引するほかに、特定の地域や国に於ける疾病の疫学的研究、公衆衛生や保健行政などに関する文献の索引に使用されることが規定されている。しかし実際には、地域に特別関係ない主題についても使用され、その結果たとえば“ミネソタに於ける医学の現況”

に関する探索で検索された文献の約 20 パーセントがミネソタに特に関連あるものではなく、例えば Mayo Clinic（ミネソタにある）で行なわれた症例研究などに MINESOTA の索引語がふられているといったような場合であった。

3.3. Exhaustivity の程度とその影響

以上、考察してきたように exhaustivity の不足は再現率に影響を与え、また過剰な索引づけは適合率を左右することが実証されたが、ここでランカスターはさらに 73 の検索例について分析を行ない、exhaustivity の程度がもたらす影響について調査を行なっている。その結果、39 例で exhaustivity の不足による再現もれが起こり、また 69 例で過剰な索引づけによる適合率の低下が生じた。再現もれの起った 39 例の文献の平均索引語数は 7.6 でこれに反して適合上のもれの生じた 69 例の平均索引語数は 11.8 であった。

索引の exhaustivity は前述のように depth indexing と non-depth indexing の方針に深い関連をもつ。そこでこの二つの索引法の相違が与える影響について無作為標本により調査を行なった結果は、depth indexing による適合率は 51.1 パーセント、non-depth indexing では 43.7 パーセントであり、再現率は depth indexing で 70.1 パーセント、non-depth indexing で 54.2 パーセントであったとランカスターは報告している。これらの数値から明らかなように、depth indexing すなわち exhaustivity の高い索引づけが、低いものよりも再現率、適合率ともに高い結果になっている。

さらに索引の exhaustivity を IM 標目だけのレベル、すなわち一文献につき平均 2.6 の件名標目を与える程度の索引づけを行なった場合に、適格的検索にどのような影響を与えるか 111 の検索例について分析が加えられた。機械検索により作成される demand search の書誌では、一つの文献に与えられた索引語のレコードは件名標目、件名標目と副標目のくみ合せ、照合事項をも含めてすべてトレーシングとして印刷される。そのうえ、IM 標目に指定された件名標目は*印によってそのことが指示されている。従って検索式に使用された件名標目にトレーシング中の IM 標目を照合させることによって比較的簡単に分析が行なわれる。しかしこの方法による比較は、23 の検索例で全く無意味であった。というのも、demand search 検索式では使用できる NIM の標目や照合事項 (check tags) は決して *Index Medicus* に印刷されることはないから、従って IM 標目だけに頼った

のでは一つも文献が検索されなかったり、或いは検索されてもごく少数の文献に限られてしまったりする結果になる。そこで、このような例を除いて残りの 88 例について分析を行なった結果は、完全な索引づけによる再現率は 60 パーセントで一方 IM 目標だけによるものは 44 パーセントにすぎなかった。適合率は、IM 標目だけによる検索では 60 パーセントであるのに反して完全な索引づけからの検索では非適合文献の検索により 52 パーセントに低下した。これらの結果を総合したものが次のような表に示されている。

表 4. 索引の exhaustivity と再現率
ならびに適合率

Exhaustivity の程度	再 現 率	適 合 率
Depth indexing (平均件名標目数 10)	% 70	% 51
Non-depth indexing (平均件名標目数 4)	54	44
IM 標目 (平均件名標目数 2.6)	44	60

この表で再現率と適合率の関係をとり上げてみると、depth indexing と IM 標目の間ではこの両者は互いに逆比例の関係にあることが明白になる。しかし non-depth indexing では例外で、再現率、適合率ともに depth indexing より低くなっている。これは non-depth indexing では、与えられる索引語の数が少ないばかりでなく通常より一般的な件名標目を使用して索引することが多いからである。

ここでランカスターは、さらに二つの小規模な実験を行なっている。すなわち一つの実験は、non-depth indexing のために索引もれになった文献をもう一度 depth indexing することによって再現率を高めることが出来るか否かを調べる目的で 10 の検索もれ文献を対象に再索引を行なった。再索引で与えられた件名標目の平均数は 11.2 で初回の索引づけの 5 個に比較してはるかに多くなっている。この実験の結果、depth indexing による exhaustivity の増加で non-depth の索引法では検索もれになった文献の 30 ないし 40 パーセントが検索されることが明らかになった。

2 番目の実験は、depth indexing の文献でありながら exhaustivity の不足のために検索もれになった 13 の文献について再索引を行ないその結果が検索にいかなる影響を与えるかをみる目的で行なわれた。対象となる文

献は検索もれになった文献中から無作為にえらばれたが、もともと depth indexing されていて平均 7.2 個の件名標目がすでに与えられているから、索引もれになった概念は比較的小さな末梢的なものであり再索引したところで再現率の改善に大した影響は与えないと予想された。再索引は一人の非常に経験の長い索引者によって行なわれ一文献あたり平均 9.1 個の件名標目が与えられたが、結果は予測通りで、新たに何らかの概念が索引づけされた文献数は 13 のうち僅か 2 文献にすぎなかった。

4. Specificity の不足

評価テストを通じて、specific な件名標目を使用したことに起因する適合上のもれはたった一件にすぎなかった。しかし再現上のもれに関しては、その中の 5.8 パーセントが索引づけの specificity の不足によるものであることが指摘された。メドラスの索引法では、前述したように specificity と exhaustivity の両者に於てその不足は索引づけの深さの程度、すなわち具体的には一つの文献に対して与えられる索引語の数に関する方針に密接に関連している。Non-depth indexing では、いくつかの関連ある概念を索引するのにそれらのひとつひとつを表す件名標目を使用する代りに全体を包含するようなより一般的な件名標目を使って索引する場合が多い。

例えば“二分脊椎”と“無脳体”に関する検索例で多数の non-depth journal に収録されている有効文献が検索もれになったが、これはいずれもそれぞれに該当する件名標目 SPINA BIFIDA あるいは ANENCEPHALUS ではなくてそれらを包括するより一般的な件名標目“奇型”ABNORMALITIES の下に索引されていたからである。Depth indexing ではこれらは SPINA BIFIDA あるいは ANENCEPHALUS などの件名標目を使用して specificity の高い索引づけがなされている筈である。また、“大腸菌とリボ多糖類”に関する検索で大腸菌のリボ多糖類の化学についての文献が検索もれになったのは、この文献が depth indexing の雑誌の収録文献であるにもかかわらず、文献中に言及されている多糖類の特殊な名称の件名標目ではなくより一般的な“多糖類”POLYSACCHARIDES の件名標目の下に索引されていたからである。また“細菌性多糖類”POLYSACCHARIDES, BACTERIAL の件名標目が使用されている場合も発見されたが、これもより specificity の高い“リボポリ多糖類”LIPOPOLYSACCHARIDES の件名標目を使用するのが最も specific な件名標目を使用して索引するというメドラス索引法の根本原理にてらし合せて正しいの

である。

このほかに、索引の specificity の不足は索引用語である MeSH の件名標目それ自体に大きな原因があるが、この問題についてはまた稿を改めて考察を加えたい。

5. まとめ

5.1. 索引における exhaustivity はもちろん相対的な事柄であり、また索引者がいかに客観性を保とうと努力してもある程度の主観的価値判断が入りこむ必然性をもっている。また情報に対する要求の種類、観点の相違などさまざまな条件によって“適合性”の基準も異なってくる。例えば、“家畜の眼球の肥肝細胞とそのヒスタミンについて”という主題を索引するのにその構成概念の一つである“水晶体”を LENS, CRYSTALLINE の件名標目を使用して十分に exhaustive な索引づけをすると、“脊椎動物の水晶体”についての検索でこの文献は検索されることになり、当然のことながらこの検索の依頼者はこれを不適合文献と判断し、検索の適合率を下げる原因となりうる。しかしながら、もし“水晶体のヒスタミン濃度”に関する検索要求があればこの文献はさしずめ適合文献の筆頭に数えられるであろうし、そうすればこの索引づけの exhaustivity の程度は全く正当づけられることになる。

5.2. 一般的にいって exhaustivity の高い索引づけを行なっておく方がむしろ妥当であることが明らかに指摘された。特定の概念が索引されていないれば検索しようにも全く方法がないのに反して、メドラスの機械検索では限られた範囲内ではあるが、exhaustivity を制限調整する機能をもっている。例えば、前に述べた IM 標目の索引法を利用して検索語を IM 標目だけに限定すれば多少の再現率の低下を伴うとしても適合率を高めることが出来る。また同様に、検索式中の重要な索引語だけを IM 標目に指定し、その他の対応づけを行なう索引語に関してはこのような制約を与えないという、いわば term weighting の方法を利用することによって再現率に大きな影響を与えることなく適合率を高めることが可能になる。

5.3. 索引の対象になる雑誌を人工的に depth journals あるいは non-depth journals にわけるとは、このテストの結果の分析からみて再現率ならびに適合率の両者の低下をもたらす原因になると考えられる。もちろん non-depth 索引のグループに属する雑誌に収録されている論文記事の多くは内容が乏しく見かけ倒しのものが多いが、また一方に於ては非常に充実した内容

をもっているものがあることも事実である。しかしこのような価値ある情報を含んだ文献も、depth indexing と non-depth indexing の区別を明確にしている方針のためにやむをえず不当に浅い不十分な索引づけをされている。

一方、例えば *Lancet* は depth journal に指定されているという理由から、その中の4分の1頁にもみたくない短い記事に対して15から20の件名標目が与えられ、検索されても要求内容に対して少しも価値の認められない情報源に終わってしまうことが起る。従って、depth journal と non-depth journal というように雑誌単位で機械的に区別するのではなく、それぞれの文献を単位にその内容価値によって索引の exhaustivity を決定することがメドラス索引法の変則性をたたくために必要であることが明白にされている。

5.4. Review (総説記事)に関する索引方針は特に疑問が多いと指摘されている。総説はすべて non-depth で索引されているが、これは評論されているもとの文献は多分 depth indexing されているであろうという仮説にもとづいている。しかしながらこの推定は多くの問題点を含んでいる。例えば、評論の対象になっている文献のあるものはメドラスの始まる以前にさかのぼるので、現在のメドラス・テープには蓄積されていない。また良い総説はもとの文献には含まれていないような新しい見地からのデータを提供している場合がある。さらに総説では、他に扱われることが比較的少ない珍しい主題について相当に内容のある討論を加えていることがありうる。例えば、今回のテストでも、“眼窩のアスペルギルス病”に関する既存の文献は非常に少なく、適合文献の中の主なものの一つは“眼の真菌性疾患”についての総説記事のうちの8頁にわたる部分であった。しかしこの文献は探索要求に適した検索式では検索されなかった。というのも“眼病” EYE DISEASES と“真菌症” MYCOSIS の2つの一般的な件名標目の下に索引されていたので、検索式をアスペルギルス病から不必要なまでに広げて exhaustive なものにしない限りこの文献は検索されることがない。MeSH にはアスペルギルス症を表す ASPERGILLOSIS の件名標目があるのでそれと、“眼窩”を表現する EYE DISEASES の件名標目と対応づけてこの主題を索引することが出来る筈である。

同様なことがこのテストの検索例で数多く起っていることが明らかにされている。

例えば、宇宙飛行における体位の限界とその長期病臥

における対比では無重力状態の医学上の問題点についての総説記事などは価値のあるものである。にもかかわらずこのタイプの文献が実際に検索されなかったのは、WEIGHTLESSNESS (無重力), HUMAN, REVIEW の3つの索引語しか使用されていなかったためで、しかもこの中の HUMAN と REVIEW はいわゆる照合事項 (check tags) で MeSH の件名標目としては WEIGHTLESSNESS だけであった。また、“肺の浸潤性疾患”についての検索で“播種性間質性肺の疾患”の主題をもつ文献が検索もれになったのは、実際に文献中では肺腺維症、組織球増殖症、鞏皮症、サルコイドーシス、肺塵埃症などを扱っているにもかかわらず索引語としては LUNG DISEASES * diagnosis (肺疾患の診断) HUMAN, REVIEW, THORACIC RADIOGRAPHY (胸部造影) しか与えられていなかった。この文献は、PULMONARY FIBROSIS (肺腺維症), SCLERODERMA (鞏皮症), SARCOIDOSIS (サルコイドーシス), PNEUMOCONIOSIS (肺塵埃症) のそれぞれの件名標目が MeSH に与えられているから、これらの specific な件名を使用して完全に索引されることが出来るものである。

5.5. 機械検索の立場からみて、non-depth indexing で、文献中に扱われている主題を表現するのに実際よりも一般的な件名標目を使用するという方針は弁解の余地なく誤っているとランカスターは指摘している。例としてあげられているのは、Poultry Science という non-depth 誌に収録された論文で“ひよこに於ける抗生物質の生長に対する大便連鎖球菌の役割”に関するものであった。この文献は *Index Medicus* を使用したマニュアルな検索では検索されたが、demand search の機械検索では検索もれになったが、その理由は POULTRY (家禽), INTESTINAL MICROORGANISMS (腸内細菌) と EXPERIMENTAL LAB STUDY (実験研究) の3つの索引語を使って索引されていることに起因する。ここでこの主題をもう一度分析して件名標目を検討してみると、ひよこは当時の MeSH の件名標目にはないので、その上位概念として POULTRY を使用するのはやむをえないとしても、大便連鎖球菌は MeSH の件名標目として STREPTOCOCCUS FAECALIS があるので non-depth indexing であるからという理由でより一般的な INTESTINAL MICROORGANISMS (腸内細菌) を使用することは正當づけられない。このような方針で索引を行なっている限り、この論文は例えば“ひよこに於ける大便連鎖球菌”や“大便連鎖球菌に対するペニシリン

の効果”とか、或いは“ひよこに於ける抗生物質の生長効果”といったようないずれも適合する探索要求に対して検索されることは望めない結果になる。事実、このままの索引づけではこの文献が検索されるのは外ならぬ“ひよこに於ける腸内細菌”に関する探索であり、しかもそのような要求内容に対してはこの文献は価値のあるものとは考えられない可能性が大きい。

最も specific な件名標目を用いて索引するという原理は正しいのであって、特定の概念を表現する適格な件名標目があるにもかかわらず索引語数に制約されてより一般的な件名標目を使用することはへい害が大きしばかりで利点は何もないといってよい。上記の例に於ても、もし STREPTOCOCCUS FAECALIS という件名標目を使用していればこの特定の細菌に関連するいかなる情報検索に於ても効果的に検索することが出来るはずである。一方、機械検索では“腸内細菌”という一般的な主題に関する検索に於てもこの文献は効果的にもれなく検索されうる方法がある。これは“Explosion”とよばれる検索方法で、MeSH の階層構造表である Tree Structune を使用してすべての細菌に関連ある件名標目を包含する generic な検索を行なうことができる。従って索引づけでは最も specific な件名標目を使用しない理由はどこにも見当たらないといえる。上記の例の文献は, POULTRY, PENICILLIN (文献中で実際に言及されている抗生物質はペニシリンである), STREPTOCOCCUS FAECALIS, GROWTH (生長) および EXPERIMENTAL LAB STUDY の 5 つの索引語を使用して十分に明確な索引づけを行なうことができる。現在のように一般的な件名標目を使用して索引していたのでは、適合する要求内容の検索では検索もれになり、また検索される場合といえは要求内容に対してはるかに一般的で価値の低い情報である場合が多く生じる結果になる。

VI. お わ り に

今回の評価テストでは、メドラスのファイルに蓄積されている索引語はすべて索引者の与えたものを忠実に反映しているという前提にたってデータの分析が行なわれた。しかしメドラスの索引では、索引者はデータ・フォーム(Data Form)とよばれる所定のワーク・シートを使用しそこに索引の結果を記録する。ワーク・シートはコンピューター入力部門に送られそこで flexowriter によって穿孔テープが作られる。従って索引者がワーク・シート上に記録した索引語が少しの誤りもなくその

まま正確にファイルに蓄積されているかどうか疑問の余地が残る。例えば flexowriter の過程でタイプもれを生じたり或いはファイルの更新維持の時に何らかの手違いが起りうる。通常は電子計算機に入力されると同時に索引のデータ・フォームと flexowriter のハード・コピーは処分されてしまうのでこの両者を比較検討することは不可能である。しかしランカスターは特に1966年度に入力された 164,000 文献のデータ・フォームと flexowriter のハード・コピーを残して索引者の誤りと電子計算機への入力の段階での誤りを区別するための小規模なテストを行なっている。その結果、索引もれの約 25 パーセントは実際の索引作業の後に起ったものであることが判明した。すなわち索引もれに起因する再現上のもれは 9.8 パーセントではなくて実際には 7 パーセント程度であろうと結論を下している。これに反して、exhaustivity の不足に関連して flexowriter のタイピングやファイルの更新維持の過程に起因する要素があることは考えられず、これは明らかに索引の方針に根本的な原因があると考えられる。従って 20 パーセントの再現上のもれは他ならぬ索引方針に起因するといえる。

今回の評価プロジェクトの結果、メドラスの実績に否定的な影響を与えているいくつかの事実が明るみに出され、それに基づいて改善が行なわれた。索引部門に関してランカスターが指摘した最も大きな影響因子は exhaustivity の不足であるが、これは具体的には索引の対象となる雑誌を depth journal と non-depth journal の 2 種類のグループに機械的にわけそれぞれの索引法について一文献あたりの索引語の数を規定した索引方針と、それに起因する exhaustivity の不足にある。そこで depth indexing と non-depth indexing の区別は 1969 年 7 月をもって廃止されることになり、⁷⁾ それ以来メドラスに入力される論文記事はすべてそれぞれの内容に応じてその主題を完全に表現するに必要であると索引者が判断した数の索引語を与える方針に変更された。従って同一の雑誌のなかでも個々の文献の内容によって索引の exhaustivity が異なるわけである。従来の depth journal, non-depth journal の区別にかわって、すべての収録誌は PRIORITY 1, 2, および 3 のグループにわけられるがこれはいわば雑誌の特徴、重要性によるグループわけで、PRIORITY 1 の雑誌は一般的に重要な利用度の高い雑誌なので従って PRIORITY 3 の雑誌に比較して exhaustivity の高い索引づけが期待されている。しかしながらこれはあくまで原則であって、最終的な判

断の基盤はそれぞれの文献にあることはいうまでもない。1969年度の統計では、一文献につき与えられる索引語は平均9.4と増加を示している。⁸⁾ また総説記事に関しても non-depth indexing の方針は変更され、すべて内容に応じて扱われている主題を完全に索引することになった。

“システムが生き残るためには、自らを監視し実績を評価しそして可能な限り改善してゆかなければならない”とはランカスターの冒頭の言葉であるが、この評価テストの結果の分析に基いて、索引部門のみならず検索部門、索引用語、demand search の申込用紙の形式、利用者とシステムの相互作用などいろいろの部門で改善策が実施された。メドラーズの第二世代への移行も目前に迫っており、そのためのシステム設計にこの評価の成果が充分生かされているであろうことは容易に想像できる。次代のメドラーズへの期待もさることながら、この評価テストの意義と次回の評価プロジェクトへの期待はさらに大きなものがあるといえよう。

(北里記念医学図書館)

引用文献

- 1) Fairthorne, R. *Computer journal*, vol. 1, 1958, p. 36-42.
- 2) *MEDLARS/NETWORK technical bulletin*, vol. 4, August 1969, p. 9.
- 3) Charen, T. *MEDLARS indexing manual*. National Library of Medicine. 1 vol. Unpaged.
- 4) Vickery, B.C. *Journal of documentation*, vol. 6, September 1950, p. 140.
- 5) Bernier, C.L. *American documentation*, vol. 16, 1965, p. 323-8.
- 6) Charen, *op. cit.*
- 7) *MEDLARS/NETWORK technical bulletin*, vol. 3, July 1969, p. 4.
- 8) *LIBRARY NETWORK/MEDLARS technical bulletin*, vol. 9, January 1970, p. 5.

参考文献

- 1) *List of journals indexed in Index medicus*. National Library of Medicine.
- 2) *Medical subject headings*. National Library of Medicine.
- 3) *MeSH Tree structure*. National Library of Medicine.