

民族資料とシソーラス

栗田 靖之 (国立民族学博物館)

守屋 祐子 (トータルメディア開発研究所)

国立民族学博物館では1976年から、研究者が記述する自然語による民族資料データをコンピュータに蓄積してきた。今回そのデータのなかから標本資料(モノ)の一般的な名称(検索名)を選び、それについて同義、類義、上下、協働使用、全体-部分関係を整理し、シソーラスを作成した。そしてこのシソーラスを組み込んで、標本資料データを専門知識がなくても簡単にのれなく検索できるような、標本資料検索システムを完成した。その作成過程とシステムの紹介をおこない、名称とモノとの関係、世界のさまざまな文化のなかで生きているモノを日本語で命名すること、そしてそれを検索することに関する問題をいくつか提起してみる。

ETHNOLOGICAL MATERIALS AND THESAURUS

Yasuyuki KURITA (National Museum of Ethnology)

Yuko MORIYA (Totalmedia Development Institute)

At the National Museum of Ethnology, the information on ethnological materials has been computerized. As the information is described by the members of research staff using their own natural languages, the database scarcely serves general users. Thus we have prepared a thesaurus by classifying the artifacts' names by synonyms and hierarchical, whole-part, as well as associative relationships. During the preparation of such thesaurus, we launched on a next step to develop a retrieval system for artifacts of the Museum. In this article, our work aforementioned will be described. Also several problems will be analyzed about searching for ethnological artifacts with only their names as the clues, especially artifacts belonging to various cultures in the world with the Japanese names.

1. はじめに

国立民族学博物館（以下、民博と記す）では1976年から、民族資料のデータの蓄積、処理、検索を機械化してきた。データの記述は民博の研究者がおこなったのであるが、そのさいに使用用語や表記の統制は行われなかった。自然科学や理工学系とちがい、人文社会科学において、用語の制限をおこなうことは、研究者の自由な発想をさまたげることであり、と考えられたからであった。

データベースの利用者は記入者本人である教官や館内のスタッフであったため、資料の名称や収集地、現地名などの見当があり、お目当ての資料にたどりつくことはそれほど困難ではなかった。

民博ではIBMのSTAIRSを導入したので、入力された単語のヨミ部がすべて検索の対象となり、コンピュータ検索のための特別なデータ加工は不要であった。そのうえ、トランケーション検索（前方一致、後方一致）も可能であったため、とくに検索用のツールがなくてもある程度の検索結果を得ることができた。

やがて、多数のデータが蓄積され、研究者の数も増加し、データベースの公開も考えられるようになった。公開された場合、従来のような研究者だけではなく、民族学に興味のある一般市民の利用も予想される。むしろ、後者の方が多いかもしれない。また、民博が提供する情報も民族学研究ばかりでなく、一般向けの出版物や映像資料、展示会など、さまざまな企画に利用されるだろう。こうした時、民族学や物質文化（いわゆるモノ）の専門家でない検索者が、研究者により長年にわたって蓄積された、世界各地の膨大な情報をコンピュータで十分に検索できるだろうか。たとえば腕につける装身具を検索する場合を仮定しよう。「腕輪」や「ブレスレット」のような言葉は入力できるかもしれない。が、実際に入力されている「腕飾り、うでわ、手首かざり、手首輪、手頸輪、肘飾り、足および腕飾り、手首につける輪」はどうだろうか。あるいは世界各地のそれらの名前を知っているだろうか。

もしその時、「検索語と同じ、あるいは近い意

味の言葉（例：腕輪とブレスレット、わんとカップ）」がわかる「類語集」と、「検索語の下位の言葉（例：わんに対しては茶わん、茶わんに対しては湯飲み茶わん）」が整理されている「シソーラス」があれば助けになるのではないだろうか。さらにコンピュータ検索で下位語は自動的に検索されればなお便利であろう。そこで、私たちは「常識的な名称から、簡単に、もれなく」標本資料データベースを検索するためのシソーラスとそれを利用した検索システムを開発した。ここにデータ作成の現場からその標本資料学術情報検索システムの作成過程を述べ、紹介をおこないたい。

なおここでは「シソーラス」という言葉を使用しているが、この用語集は階層性よりも言葉と言葉の関連性を重視し、また階層についても正確なツリー構造ではないので、厳密な意味でのシソーラスではないことをおことわりしておきたい。

2. 標本名について

データベースを検索するにはそれがどのようなものかを知っておく必要がある。そこでまずはじめに民博の標本資料のデータベースを紹介しよう。

現在民博には、博物館の研究者の手によって世界各地から収集された約19万点の民族資料（館内では標本資料と称される、いわゆる「モノ」）が収蔵されている。これらの民族資料は、すべて一件ずつ事務管理用のデータ（管理カード）と民族学研究用のデータ（情報カード）が作成される。後者は「学術研究情報データベース」を構成するもので、その資料についての詳しい情報が研究者自身によって記されて、民族学研究や博物館活動のために利用される。

では標本資料の学術研究情報はどのように作成されるのだろうか。昭和51年資料管理委員会作成の『情報カード記入について』には、記入責任者名からはじまって、現地名・訳名、標本名、収集年月日、収集地、使用地、使用民族、用途・使用法、製作年代、製作法、製作状況、製作者、変遷・分布などの22項目とそれらの記入のしかた

が細かく定められている。通常、検索のアクセスキーとなるのは標本名なので、ここに標本名の記入要項を抜粋してみよう。

3.3. 標本名

海外標本の場合は、その品物に対応すると思われる日本語の一般的名称を、国内標本の場合は、たとえば現地名「すりつけぎ」にたいしての一般的名称である「マッチ」のように記入します。

なにが一般的名称であるかについては、さまざまな問題が生じるところです。将来は、本館の収蔵品を統一的に分類する名称体系の整理をおこなわねばならないことになるでしょうが、それを待つわけにもいきません。現状ではアチック・ミュージアムの『民具蒐集調査要目』にあげられた名称などを参考にしながら、記入していただき、記入者の常識的な判断に、標本名の命名をおまかせするほかありません。一つの前よりどころとしては、国語辞典に出てくる名称をなるべく採用していただくことです。また、ばくぜんとした一般名称ではなく、用途、材質など、その標本の特徴がわかるような名称を書きこんでください。たとえば、

桶ではなく 鉢桶、水桶

籠ではなく 背負籠、魚籠

短刀ではなく 銀鞘短刀

パイプではなく 陶製パイプのように

以上のようなルールのもとに与えられた標本名は約一万一千種である。そのなかには「釣り針」「菜切り包丁」「腕輪」「ブローチ」といった一般的で、具体的なモノのイメージが浮かぶものも多数ある。一方「手首につける輪」「胸飾り」「手首ナイフ」「殺魚棒」「瓶栓機」「筒杖楽器」「たる作りの用具」「指ではじく楽器」「熊乗り金太郎」といった、モノの「一般名称」とは考えにくいようなものや「晴れ着用ワンピースの下に着るワンピース」「木製くりぬき背負い瓶」といった、その標本の特徴を記したために修飾部分が長く冗漫なものもふくまれるようになった。

そして『記入について』では、表記の方法については触れられていないため、情報カードで使われている漢字、かなづかい、送りがな、外来語の表記法、翻字法などもさまざまであった。たとえば、「殻ざお」に対する標本名の表記は「カラサオ、から棹、唐竿、殻竿、唐棹」、読みは「カラサオ、カラザオ」があり、「殻ざお」をくまなく検索するには漢字検索（平成5年より導入されたJAIRS）で五種類、カナ検索（STAIRS）で二種類の検索語を入力しなければならなくなった。

3. 検索システム用データ作成作業について

3-1. 標本資料のO C M分類

標本資料をO C Mコードによって分類し、同じような標本名を収集することからシソーラス作成作業を開始した。

ここでO C Mについて述べておくと、O C Mとは、『文化項目分類 Outline of Cultural Materials』の略称である。これは Human Relation Area Files, Inc. が考案した、すべての人間社会に共通する「文化」の分類体系である。そこでは「20 コミュニケーション 201 身ぶり」と記号「202 メッセージの伝達 203 ニュースや情報の流布 204 出版物 205 郵便制度・・・」、「75 病気 751 予防医学 752 外傷 753 病気の理論 754 邪術 755 呪術的療法と精神療法・・・」というように、民族や地域にかかわらず普遍的にみられる人間行動や社会生活、それに関する現象が二桁の大区分と三桁の小区分によって637項目にわたって分類されている。

情報カードの用途・使用法、使用者、変遷・分布などのデータにもとづいて資料を分類すると、同じような概念をもつ標本名がO C Mコードのもとに収れんしてきた。たとえば「415 用具」の下位区分である「食べ物と飲物の容器」には「木製碗」「白樺皮製碗」「輪島塗黒碗」「籬吸物碗」「ラマ僧用茶入れ碗」「ふた付き茶漬け碗」「茶飲み茶わん」「椰子の碗」「民族碗」「三足わん」「弁当箱」「弁当籠」「タジン」「重箱」「もち桶」などである。

こうしてあつまった標本名のなかで、個別の標本の特徴を述べているような修飾語などを削除し、なるべく一般的で、かつその資料の個性を示すようにモノの名称を整理した。

3-2. 検索名の選定

上記の作業を経ても、「椀」「碗」「漆器椀」「吸物椀」「茶入れ碗」「飯わん」「木わん」「茶漬け碗」「スープ茶わん」「茶飲み茶わん」というようにまだ自然語といってよいものであった。

そこでこれらの名称のなかから、新明解国語辞典の見出し語になっている単語を選んで「検索名」とした(約1700語)。新明解国語辞典を典拠とした理由は『記入について』に記されていたように、名称は百科事典ではなく国語辞典に拠るものであること、現代の「常識」という基準を重視したこと、そしてモノに関する収録語数が多いこと、である。歴史的、あるいは専門的な語や意味は避けるようにした。そこで現地名、用途を示す名称、技法を示す名称、専門的な名称、仕掛けを示す名称、汎称などは検索名とならなかった。

検索名の定義は新明解国語辞典だけでなく、広辞苑やその他の事典、研究書を参照し、言葉の意味が一面的になりすぎないように注意して、検索語辞書を作成した。(表1)

検索名の表記についても「送り仮名の付け方」「現代仮名遣い」「常用漢字表」「外来語の表記」によって表記法を統一した。

3-3. 検索名の付与

以上のようにして選定した検索名を学術研究情報の標本名に照合しなおし、あらためてデータベースの個々のドキュメントに付与した。

検索名はシソーラス上の関係を考慮せずに、その標本に最もふさわしいものだけを付与するようにした。たとえば標本名の「釣針(擬餌針)」には「針」、「釣針」でなく「擬餌鉤[ぎじばり]」である。上下関係、同義関係は機械的にシソーラスでたどられて検索キーとなり、検索される。これは将来、よりふさわしい名称が検索名になった場合や、シソーラスで他の語との関係が変更になった場合のメンテナンスを考慮したためである。

この作業では、具体的な「モノ」である標本名と、抽象的な「言葉」である検索名とが直面することになったためさまざまな問題が生じたが、そのうちのいくつかをあげておきたい。

- 1) 「料理用権」 モノが本来の用途以外に用いられている場合、
- 2) 「パトカー(玩具)」 モノの概念だけが生かされている場合、
- 3) 「腰みの(日常着)」 モノとしては類似しているが日本と用途が異なる場合(注:日本の腰みのは防水着であるが、オセアニアのある地域ではその意味はなく日常の衣類として用いられている)
- 4) 「機(綿繰り用)」 同じ単語であるが現地での意味と国語辞典での意味が異なる場合、(注:通常、機は布を織る道具である)
- 5) 「ふろしき」「刀」「トロイカ」 名称に

表1 検索名辞書

1 アイシャドー	まぶたに塗る、青・灰色などの化粧品。
2 アイスボックス	氷を使う冷蔵庫。特に携帯用の冷蔵容器。
3 アイゼン	登山靴の底につける、滑り止めの鉄製のつま。
4 アイロン	衣服や布のしわをのばし、形を整える器具。金属の部分を熱して用いる。
5 あぐり網	巻網の一種。長方形にまわして魚を囲み、網の下方から繰り上げるもの。
6 あごひも	帽子がぬげないようにあごにかけるひも。
7 足かせ	足にはめる刑具。
8 足駄	高い歯の下駄。高下駄。
10 アズサ弓	アズサの木で作った丸木の弓。
11 汗取り[肌着]	直接肌につけて汗を吸い取らせる肌着。

それがうまれた文化のイメージが大変つよく残っている場合

上記の1)、2)、5)にはモノの「概念が生かされている」と判断して該当する検索名を付与し、3)、4)は「意味が異なっている」ので付与しなかった。

モノはそれぞれ固有の形態・機能・材質・用途をもっている。いったい、何が同じであれば同じ名前がつけられるのか。たとえ、それがすべて同じであっても言葉には暗に示すイメージというものがある。(ふろしき、刀という時に何か日本のなものを思い浮かべはしないだろうか) 一般の辞書にはこのような細かい記述はないが、もしこうした微妙なニュアンスを検索名の定義に採用すると、検索名は日本的な色彩の強いものになるだろう。すると、日本以外のモノには該当しにくくなり、検索名のないものが増える。そして、検索できるものは日本文化中心になる。多様な世界文化を日本語でとらえる難しさがここにある。

一方、ひとつの標本名のなかに複数のモノがある場合(「臼と杵」)やひとつのモノを複数の名称でいいかえている場合(「大盆または盤」)など、単純に解決できるケースもあり、最終的には約80パーセントのドキュメントに検索名を付与することができた。

ところで情報カードは記入責任者の研究成果なので、その内容を本人以外が書き換えることはできない。そこで私たちは学研究情報のデータはそのままにし、あらたにエリアを設けて検索名を付与した。その結果、検索者は研究者自身の自由な名称からでも、シソーラス辞書による限定的な名称からでも検索できるようになった。

3-4. 検索名間の関係づけ

検索名の意味を考えながら、以下のように同義関係、類義関係(協働使用関係、全体一部分関係を含む)を整理した。同じ概念のモノを検索するという当初の目的からは、協働使用関係は不必要である。しかし「学研究情報」は研究者ばかりでなく、博物館の展示や、一般からの質問の調査に利用されることも多い。そのような場合、モノ

だけでなく、それを取りまく状況や他との関連を調べる可能性もあることを考慮して関係づけることにした。また全体一部分関係は検索対象が博物館の収蔵資料であることを配慮した結果である。

同義関係(D) あらゆる場面において語の意味が同じで、置き換え可能な関係

例: 首飾り⇔ネックレス、杯⇔酒杯

類義関係(S) 語の意味は非常に近いが、同義関係にはならない関係

例: 縄⇔綱⇔ロープ⇔ザイル、カップ⇔コップ

上位関係(B) その語より大きな概念や、その語を含んだ全体像をしめす関係

例: 茶わん⇒わん、かんかん帽⇒帽子

下位関係(N) その語より小さな概念や種類をしめす関係

例: わん⇒茶わん、帽子⇒かんかん帽

協働使用関係(C) 実際に使用される場合、同時に用いられる可能性が非常に高い関係

例: 釣針⇔釣ざお、うす⇔きね、銃⇔弾丸

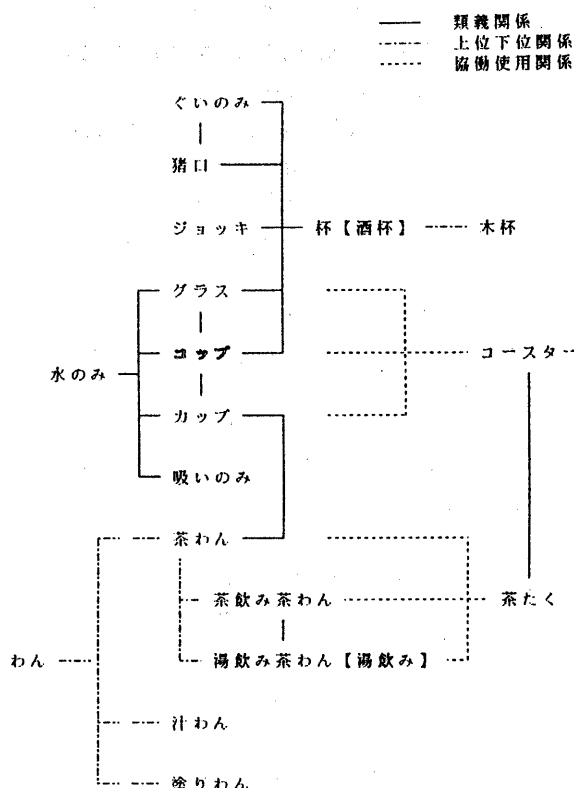
全体一部分関係(A) 「モノ」全体とそれに特有の部分との関係

例: きせる⇔がん首、織機⇔おさ

コンピュータ検索の場合、検索語の同義語は画面上に表示されて、検索語と同様の処理がおこなわれる。「首飾り」が検索語の場合は「ネックレス」も自動的に検索される。つぎに上位語と下位語は画面に表示されない。ただし、検索語の下位語は検索語と同様の処理がおこなわれる。たとえば、「わん」を検索した場合、その下位語である「茶わん」「汁わん」「塗りわん」、そして「茶わん」の下位語である「茶飲み茶わん」「湯飲み茶わん」とその同義語「湯飲み」という六語は自動的に検索される。また、類義語と協働使用関係語、全体一部分関係語とは入力した検索語に追加できるようになっている。

コップやわんの類義関係を例に図示するが(表2)作業用のファイルなどはこのようになっていくわけではない。

表2 コップを中心にした類義関係



今後、検索名が増加したり変更される場合にそなえて、語と語の関係の意味によってグループを作った。そして各グループと検索語とのむすびつきの強さの度合いによって順位づけを行った。この順位は類義語がコンピュータの画面に表示されるときに表示順として役立つことになった。

- ひも S1 糸、緒（幅のない細長いもののグループ）
 S2 バンド、ベルト（幅の多少ある細長いもののグループ）
 S3 帯締め、帯止め（和装の帯を押さえる細長いもののグループ）
 S0 テープ、リボン（その他）

4. 検索システムについて

では具体的にシステムを利用して「コップ」を検索してみよう。

標本資料検索システム

このシステムは標本資料検索用語集を用いたものです。標本資料の検索を始めるには次の方法があります。

- 1 標本資料の標準的な名称で検索する。
- 2 文化項目分類（OCMコード）で検索する。

① メニュー画面

- 1) ①メニュー画面で「1. 標本資料の標準的な名称で検索する」を選ぶ。

標本資料検索システム【検索語入力】

検索語: コップ

② 検索語入力画面

- 2) ②検索語入力画面で「コップ」と入力する。

標本資料検索システム【類義語選択】

検索語 コップ には次のような類義語があります。

類義語【同義語】一覧

グラス
 カップ
 水飲み
 酒杯【杯】
 コースター

③ 類義語選択画面

- 3) ③類義語選択画面では「コップ」の類義語（グラス、カップ、水飲み、酒杯【杯】・酒杯と杯は同義語）と協働使用関係語（コースター）が表示される。入力語に追加して検索したい語があればクリックする。

標本資料検索システム [OCM選択]

検索語 コップ、グラス、水飲みは次のような
OCMコードと関連があります。

OCM 説 明
264 食物消費：食事行動
271 飲み物、薬物、嗜好品：水と湯き
272 飲み物、薬物、嗜好品：非アルコール飲料
273 飲み物、薬物、嗜好品：アルコール飲料
276 飲み物、薬物、嗜好品：幻覚剤と興奮剤
322 原材料の加工：木工
323 原材料の加工：陶磁器産業

④ OCM 選択画面

- 4) ④OCM選択画面では「コップ」と3)であ
らたに検索語に追加した「グラス、水飲み」に
関係するOCMコードが表示される。用途や材
質などを限定して検索したい場合はOCMで指
定する。OCM本文は⑤OCM説明画面 でみ
ることができる。

シソーラス検索システム [OCM説明画面]

OCMコード：264 食物消費：食事行動

- .01 食物摂取の規則性
- .02 変則的な食物摂取 (例、間食、働きながら、遊びながらの食物摂取)
- .03 食事 (例、時間、1日の回数、参加の仕方、構成)
- .04 食べ方
- .05 食事作法
- .06 食後の問題 (例、食べかすや食べ残しの処理、食器類の後始末)
- .07 関連する信仰や慣習
- .99 その他

⑤ OCM 説明画面

標本資料検索システム [検索結果]

検索結果： 20 点の標本資料があります。

検索語： コップ、グラス、水飲み

OCM： 272

⑥ 検索結果画面

- 5) ⑥検索結果画面では検索条件の表示と該当する資料数が表示される。

標本資料検索システム [標本名一覧]

標本名	使用地
コップ (Glass)	フィンランド共和国
コップ (Glass)	フィンランド共和国
得度式用コップ	タイ王国 BANGKOK
得度式用コップ	タイ王国 BANGKOK
得度式用コップ	タイ王国 BANGKOK
バイナッブル酒用コップ	コロンビア共和国 VAU
コップ (真鍮製)	
煉物コップ	コロンビア共和国
コップ (木製)	ネパール王国 PHIJER
コップ (木製)	ネパール王国 TUKCHE
コップ (木製)	ネパール王国 TUKCHE
コップ (木製)	ネパール王国 TUKCHE

⑦ 標本名一覧画面

- 6) ⑦標本名一覧画面では該当する標本資料の資料名、使用地などの一覧が表示される。各資料の詳細情報 (学術研究情報) は⑧詳細情報画面でみることができる。

標本資料学術研究情報データベース

標本番号： H0002917
カード番号： 01
OWC： E01
OCM： 323.07 415
収蔵場所： 03-1G-30-02
シート1
標本の種類： 実物 (現物)
現地名： ㊦㊧㊨
LASEJA (フィンコ)
標本名： コップ (GLASS)
コップ (Glass)
収集年月日： 1975.08
収集地： ㊦㊧㊨㊩㊪㊫㊬㊭㊮㊯㊰㊱㊲㊳㊴㊵㊶㊷㊸㊹㊺㊻㊼㊽㊾㊿
現認 フィンランド共和国

⑧ 詳細情報画面

標本資料検索システム [同音異義語選択]

検索語： かま

入力された語には同音異義語があります。

鎌 [かま]

釜 [かま]

⑨ 同音異義語選択画面

- 7) ⑨同音異義語選択画面。もし2)で入力された表記に異義語がある場合は、ここでどちらかを選択する。

今度は「運搬に使用するモノ」を検索してみよう。

標本資料検索システム

このシステムは標本資料検索用語集を用いたものです。標本資料の検索を始めるには次の方法があります。

- 1 標本資料の標準的な名称で検索する。
- 2 文化項目分類(OCMコード)で検索する。

⑩ メニュー画面

- 1) ⑩メニュー画面で「2. 文化項目分類(OCM)で検索する」を選ぶ。

標本資料検索システム [OCM入力]

OCMコード: 482

⑪ OCM入力画面

- 2) ⑪OCM入力画面で「482」と入力する。
運搬に関するOCMコードはの「482 荷物の運搬」である。

標本資料検索システム [検索語選択]

OCM 482 荷物の運搬 は次のような言葉と関連があります。

関連する検索語一覧

網
板
笥 [おい]
おうご
おか持ち
おけ
箆
鉤 [かぎ]
籠 [かご]

⑫ 検索語選択画面

- 3) ⑫検索語選択画面では「482」に関係ある検索名が表示される。モノを指定して検索したい場合はそれをクリックする。

- 4) 以下は「コップ」の場合と同様である。

5. おわりに

このシステムは民博の教官からは好評を得ることができたが、より活用するために、オペレーションの簡便化、地域や民族による検索とそのシソーラスの作成、検索名とシソーラスの画面表示と拡充、画像情報の表示などの要望もよせられた。

データ作成の立場からは、3-3. 検索名の付与で述べた問題のほかに、検索名にならなかった現地名や汎称、「石貨」「割れ目太鼓」「穂づみ具」などの民族学用語のあつかいや、またそれらを検索名にした場合、非専門家をその言葉へ導く方法、などがあげられよう。

実際にモノを探す場合には名前だけでなく、用途や使用場面から探すことが多い。現在の検索システムではその機能をOCMコードが果たしているわけだが、OCMは本来民族学の専門家のために開発されたものである。それゆえ一般人が予備知識なしに、自分の望む内容をしめすコードに到達することは難しい。またOCMは本来文献研究のために開発されたものなので、モノそのものではなくモノに関する情報(用途、使用法、使用者など)を分類するのに適している。モノを名称だけで分類するには不適当である。たいてい、モノの名称と用途は結びついているので前述の「482 荷物の運搬」を例にすると「かばん、リュックサック」という言葉が画面に表示されても違和感はない。が、「網、板」はどうだろう。これは、この名称をもった標本資料が「運搬に使用される」という情報をもっている、という意味であって網や板が運搬具なのではない。モノの名称とOCMコードを直接結びつけるのには無理があるのではないだろうか。このように考えるとOCMにかわって、用途や使用法をあらわす言葉のシソーラスの開発がつぎの問題になると思われる。

「一般常識」から「学術的情報」へ、簡単に、的確にたどりつくにはどうすればよいか。それがシソーラスをふくめて、私たちのなおの課題である。