

情報検索の基礎

＊ ＊ 情報リテラシーのスキルアップのために ＊ ＊



2005.11

産業医科大学図書館

はじめに

学習および研究活動における必要不可欠な行動の一つとして、情報の収集とその利用・活用があげられます。こうした活動の重要さは、インターネットで代表されるようなネットワーク環境においては、必要な情報をいかに迅速・適切に入手できるかが、ますます重要になっています。

これは情報を効果的・効率的に検索する手段、システム、技術に対して理解が必要なことを示しています。

目 次

1. 情報検索 (information retrieval) に関する用語集.....	3
2. 情報リテラシー.....	5
3. データベース.....	6
4. 検索戦略の立て方.....	9
5. 検索の流れ.....	9
6. 索引作業.....	10
7. 検索フィールド.....	11
8. キーワードの選択	
8-1 統制語(ディスクリプタ)とフリーキーワード.....	12
8-2 シソーラスの役割.....	13
8-4 キーワード選びのコツ.....	15
8-5 ストップワード(不要語).....	16
9. 検索機能	
9-1-1 論理演算機能とは.....	17
9-1-2 論理演算の優先順位.....	18
9-2 トランケーション機能(ワイルドカード)	
9-2-1 トランケーション機能とは/9-2-2 前方一致検索/9-2-3 中間一致検索.....	19
9-2-4 両端一致検索/9-2-5 後方一致検索.....	19
9-2-6 トランケーションとワイルドカード/9-2-7 トランケーション機能使用の注意点...	20
9-3 横断検索機能/9-4 EXPAND 機能(エクスパンド).....	21
9-5 近接演算機能/9-6 完全一致検索機能/9-7 範囲検索機能.....	22
9-8 検索式一覧機能(検索履歴の表示).....	23
9-9 検索式保存機能.....	23
10. 検索結果の絞り方.....	23
11. アラート機能.....	24
12. ドキュメント・デリバリー・サービス.....	24

1. 情報検索 (information retrieval) に関する用語集

アイ・エス・エス・エヌ (ISSN)

International Standard Serial Number のことで「国際標準逐次刊行物番号」と訳されます。1つの雑誌に付与された一意の番号表記のこと。

アイ・エス・ビー・エヌ (ISBN)

International Standard Book Number のことで「国際標準図書番号」と訳されます。1つの図書に付与された一意の番号表記のこと。

インターネット (internet)

世界各地のコンピュータネットワークを相互に接続した世界最大のコンピュータネットワークの総称。

エンドユーザー (end user)

情報の最終的な利用者のこと。

OPAC (online public access catalog)

オンライン閲覧目録のことであり「オパック」と発音する。

書誌 (書名、著者名等) 情報を蓄積したデータベースに端末装置やネットワークに接続されたコンピュータからアクセスし、資料の検索を行うこと。

図書館では、「SALY OPAC」<サリーオパック>と名称づけ

図書館ホームページからアクセスします。

かいご (下位語: narrower terms) 上位語

見出し語の概念に含まれる用語のこと。

キーワード (Keyword)

検索の手がかりとなる語。文献の主題内容を表わす“鍵”としての役割を持つ語のこと。

自然語---文献の標題や本文中に用いられていることばの中から抽出されるもの。

統制語---あらかじめ選定・統制されている用語リスト (件名標目表、シソーラス等)

から文献内容に合った用語 (件名、ディスクリプタ) を割り当てて付与されたもの。

けんさくモレ (検索モレ) ノイズ

検索条件に合致しているにもかかわらず検索されなかった情報のこと。

けんめい (件名: subject)

資料の主題や概念を表わす語のこと。

コマンド (command)

コンピュータで各種の処理を行わせる命令のこと。

さくいん / インデックス (索引: index)

索引語を一定の順序に並べ、かつその出所 (= その索引語が付与された文書) が分かるようにしたもの。各索引語とその出所を表すデータは転置ファイルによって関係付けられる。

さくいんさぎょう (索引作業: indexing)

登録するそれぞれの文書の主題を分析し、その内容を的確に表す言葉 (キーワード) やコードを付与すること。

シソーラス (thesaurus)

情報検索や索引作業を行う際に使用頻度の高い用語を集め、用語間の相互関係や用語の適用範囲が分かるように記載されている用語集のこと。

しゅだいぶんせき (主題分析 : subject analysis)

文献の内容を分析して、そこに含まれている中心となる重要な概念を捉えること。

じょういご (上位語 : broader terms) 下位語

見出し語の概念より広い概念を持つ用語のこと。

じょうほうけんさく (情報検索 : information retrieval)

データベースなどに収録されている情報の中から、必要とする情報を探し求めること。

しよしじこう (書誌事項 : bibliographic data)

文献を識別するための指標となる事項のこと。

例えば、図書の標題、著者名、発行年等

そきゅうけんさく (遡及検索 : retrospective search)

過去に溯って検索すること。

タイムラグ (time lag)

発生した情報が利用者の手元に届くまでの時間のこと。

てきごうりつ (適合率 : precision ratio)

検索条件に適合した情報がどれくらいの率で検索されたかを測るための指標のこと。

データベース (database : DB)

大量のデータを系統的に整理し、コンピュータで扱えるようにしたファイルのこと。

どうぎご (同義語 : synonym)

ほかの用語と全く同じ意味又はほとんど同じ意味をもつが表現が異なっている語のこと。

とうせいご (統制語)

シソーラスとして統制された用語のこと。ディスクリプタともいう。

ノイズ (noise) 検索モレ

検索条件に合致していないのに検索されてしまった情報のこと。

ピーディーエフ (PDF : Portable Document Format)

デジタル書類の内容 (文字、画像、レイアウト情報など) が表示・出力できるファイル形式。

ひょうだい (標題 : title)

図書の書名や論文・雑誌記事などの論題名のこと。

フィールド (field)

レコードの内容を構成するデータの項目のこと。

ぶんけんけんさく (文献検索 : literature search)

図書、新聞記事、論文などの文献を対象に検索を行うこと。

レコード (record)

データベースで扱う1件分のデータのこと。

2. 情報リテラシーとは

1990年代後半のインターネットの出現と普及が社会にもたらした影響は大きいものがあります。インターネット上には膨大な量の情報が流れていて、机上のパソコンから世界中のデジタル情報資源へのアクセスを可能しました。

このため、インターネットを使える人と使えない人の間には、“情報の格差”が生まれています。インターネットが使えないと必要な情報を十分に得られなくなりつつあります。

リテラシー（literacy）とは読み書き能力をさします。

情報リテラシーとは、コンピュータなどを利用して幅広く情報を収集・活用し、それを分析、判断する基礎的な情報活用能力のことをいいます。



2-2 情報リテラシーの指針

アメリカの情報リテラシー指針（10か条）は以下のとおりです。

- 情報の必要性を認識する
- 正確で完全な情報が意思決定の基盤になる
- 情報ニーズに基づいて質問を定式化する
- 利用可能な情報源を選択する
- 効果的な検索のための戦略をたてる
- コンピュータを利用して情報源にアクセスする
- 情報源を評価する
- 情報を組織する
- 既存の知識体系の中に情報を統合する
- 批判的思考と問題解決において情報を利用する

3 データベース

3-1 データベースの誕生

データベースという言葉は、1950年代に米国国防総省が当時世界中に展開していた兵員、武器の管理を手作業で行うことが困難となり、コンピュータ技術を駆使したライブラリーを開発し、これを「データの基地(ベース)(data base)」と呼んだのが始まりだといわれています。

3-2 データベースの定義

データベースの定義の主なものは次のとおりです。

(1)著作権法の定義(著作権法第2条10の3)

「論文、数値、図形その他の情報の集合物であって、それらの情報を電子計算機を用いて検索することが出来るように体系的に構成したものをいう。」

(2)日本工業規格の定義

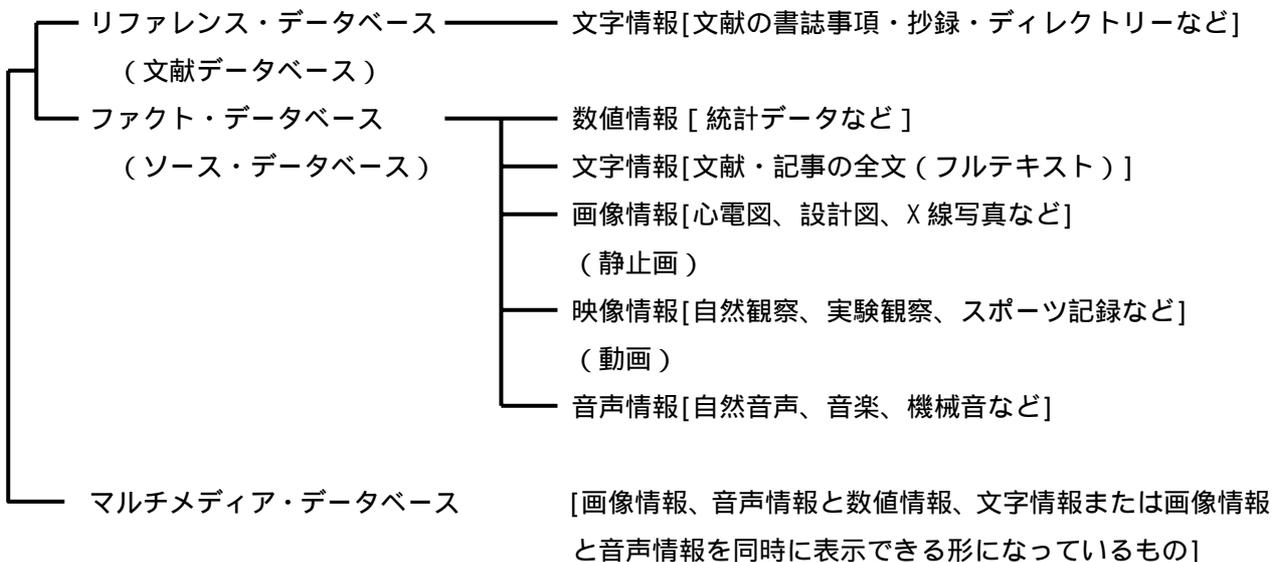
「適用業務分野で使用するデータの集まりであって、データの特性とそれに対応する実体との関係とを記述した概念的な構造によって編成されたもの」

(3)データベース白書2003の用語解説での定義

「相互に関連のあるデータの集まり。複数のユーザーが種々の目的で利用できるよう、検索や更新の効率化が図られたもの。一般には情報を統合化し、コンピュータ処理が可能なファイルの形で蓄積する」

これらの定義から明らかなように、データベースとはさまざまな情報がコンピュータ処理できるように体系的に整理され、統合化・構造化されて蓄積・保存されており、その中から必要な情報だけを部分的に取り出せるもので、情報検索とデータベース検索は同義といえます。

3-3 データベースの分類



データベースは収録されるデータの形態により、リファレンス・データベース(reference database)とファクト・データベース(fact database)に大別できます。

(1) リファレンス・データベースとは・・・

雑誌論文や新聞記事などの書誌事項(標題、著者名、雑誌名、巻号数、掲載ページ、発行年など)にキーワードや抄録などの二次情報を文字情報で収録したもので、索引誌(Index Medicus など)や抄録誌(医学中央雑誌)のような二次資料がデータベース化されたものです。

探し求めている情報が、どのような図書、雑誌、新聞などの一次資料に収録されているかを案内してくれるデータベースであり、文献データベース、書誌データベース、案内データベースとも呼ばれています。

(2) ファクト・データベースとは・・・

数値データベース、全文データベース(フルテキスト・データベース)、画像データベース、映像データベース、音声データベースの5種類に細分できます。

探し求めている情報、すなわち一次情報そのものが検索でき、一次資料がデータベース化されたもので、事実型データベース、ソース・データベースとも呼ばれています。

ソース(source)とは、情報源という意味であり、一次情報を示しています。

(3) マルチメディア・データベースとは・・・

CD-ROM やインターネットの普及に伴って増えてきたのが、マルチメディア・データベースです。

文字、数値、画像、映像、音声などを様々な組合わせて同時に表示できる形になっているものがマルチメディア・データベースです。

3-4 データベースの選択

最近では、ネットワーク(特にインターネット)上に有償または無償で多数の情報資源が存在しています。その中から自分の検索テーマに合致するデータベースを選択することが重要です。

選択するポイント

データベースを提供している検索システム(どういう検索ソフトなのか)

データベースの内容(どういう主題、分野のものか)

データの出所(オリジナルデータは何か、どこが作成し提供しているか)

更新頻度(即時、毎日、毎週、隔週、毎月、毎年、不定期など)

収録期間(最新のみ、過去1年間、1971年以降など)

データの形態(文字情報、数値情報、画像情報など)

言語(日本語、英語、フランス語、ドイツ語など)

3-4-1 主な商用データベース

分野	データベース名	編集機関	内容
医学	医中誌 Web	医学中央雑誌刊行会	日本の医学雑誌論文の文献情報
	メディカルオンライン	メテオインターゲート	日本の医学雑誌論文の全文情報
科学技術、 医学	J-Dream	科学技術振興機構	科学技術、医学に関する、学術論文や解説的記事などの文献情報
医薬品	JAPIDOC	日本医薬情報センター	医薬品の有効性、安全性に関する国内雑誌に収録された文献情報
新聞記事	G-Search	ジー・サーチ	新聞記事、企業概要、財務、医薬品、法律情報など記事情報
	朝日新聞記事データベース	朝日新聞社	新聞記事全文情報
	毎日新聞記事データベース	毎日新聞社	
	日経新聞記事データベース	日本経済新聞社	
	読売新聞記事データベース	読売新聞社	
人物	NICHIGAI WEB	日外アソシエーツ	人物のプロフィール情報および文献情報

分野	データベース名	編集機関	内容
医学	MEDLINE	National Library of Medicine	世界中の医学、歯学、看護学分野の雑誌論文の文献情報
	EMBASE	Elsevier Science Publishers B. V.	世界中の生物医学、薬学分野の雑誌論文の文献情報
	CINAHL	CINAHL Information System	世界中の看護学分野の雑誌論文の文献情報
	Biomedical Reference Collection	EBSCO	雑誌論文の全文(フルテキスト)情報
全分野	inside Web	British Library	雑誌論文情報とドキュメントデリバリーサービス
生物科学	Biosis Reviews	BIOSIS	世界中の生物および生物医学分野の文献情報
化学	CA Search	Chemical Abstracts Service	世界中の化学および化学工学分野の雑誌論文、特許等の文献情報
物理	INSPECT	The Institution of Electrical Engineers	世界中の物理学、電気工学、電子工学、コンピュータ等の文献情報
工学	Ei COMPENDEX PLUS	Engineering Information Inc.	世界中の工学、技術分野の雑誌論文、技術報告書、単行本等の文献情報

4 検索戦略の立て方

有効な検索結果を得るためには、検索を始める前に周到な計画(検索戦略)を立てることが重要です。

検索戦略とは、実際に検索をどのような順序あるいは、どのような内容で進めるか検討することです。

① 検索主題を明確に

a) 疾病名	b) 疾病の発生個所	c) 疾病の対象臓器
d) 診断方法、治療方法	e) 使用する医薬品、化学物質	f) 疾病の因果関係
g) 研究対象	h) 研究の種類	

② 検索主題をどのようなキーワードで示すか

③ 検索主題をどのような検索式(キーワードと集合演算子の組み合わせ)で示すか

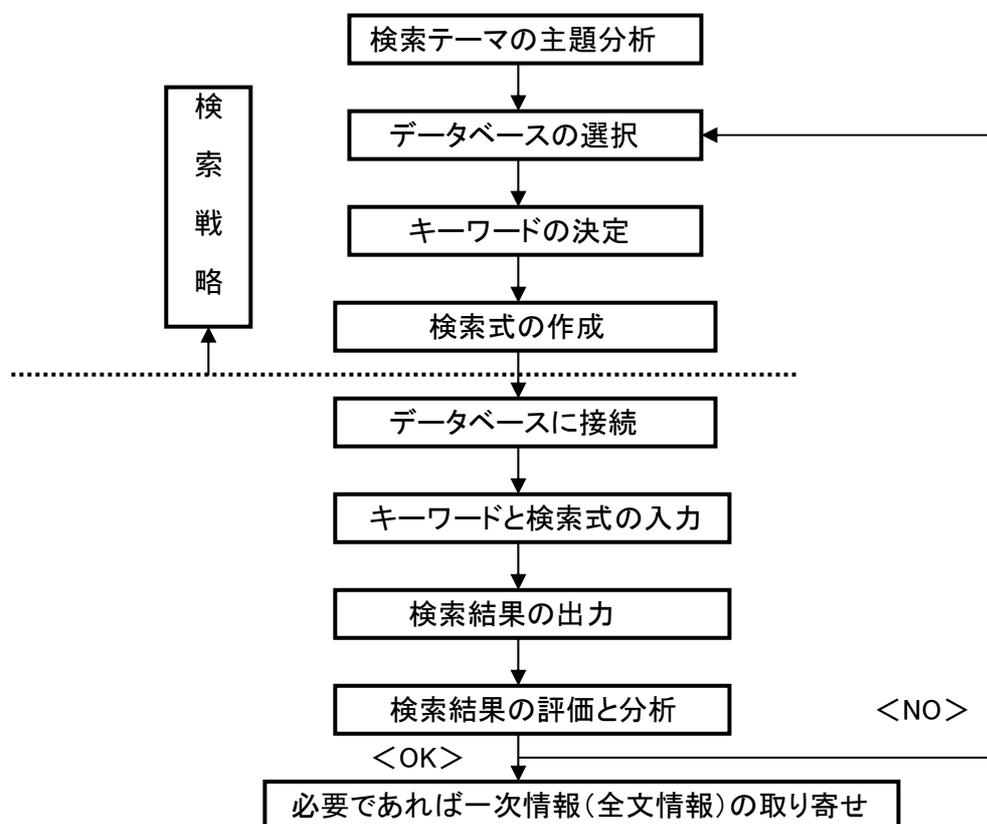
④ 検索システムのどの機能を使うのか

⑤ 情報が多い時の絞り方あるいは、下位概念のキーワード(下位語)の検討

⑥ 情報が少ないかあるいは、ない時の上位概念のキーワード(上位語)の検討

5 検索の流れ

情報検索には、データベースを使用した実際の検索の実行に加えて、検索を始める前の準備から検索結果の報告まで多くの作業が含まれます。



6 索引作業

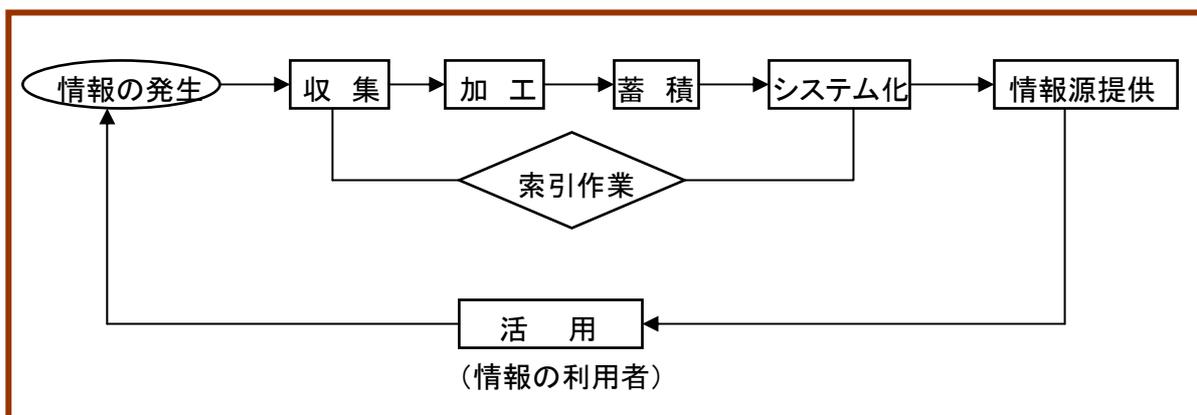
索引(index)とは、情報の中の主要な語句や事項を抽出して、アルファベット順などの一定の順序に配列し、その所在元を容易に探し出せるように作成されるものです。

索引作業は、索引作業の対象となる個々の情報について、ある特定のアクセスポイントの検索を可能にするためなどに行われます。

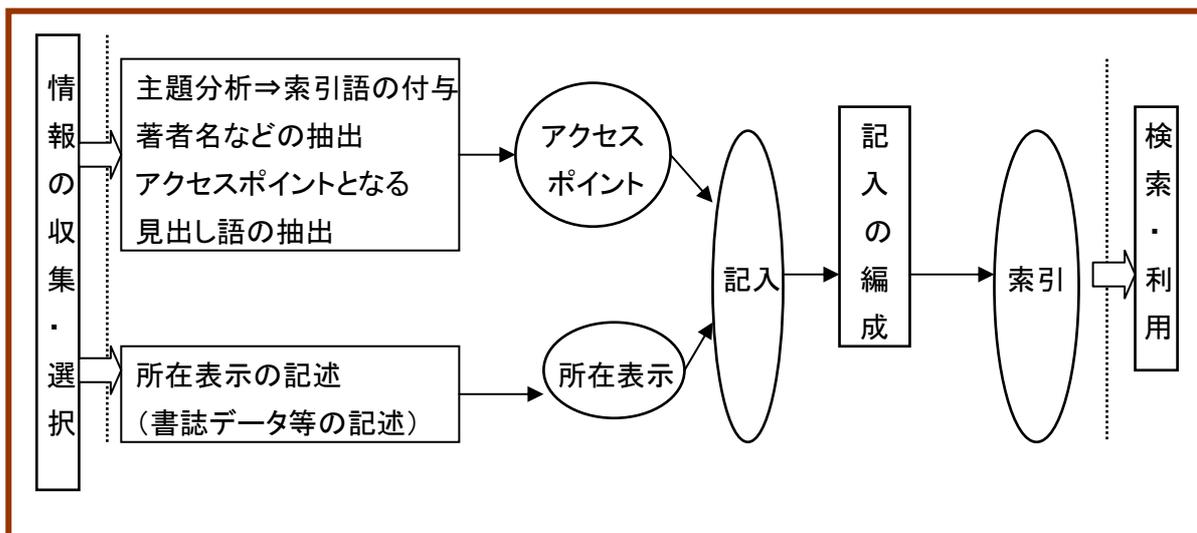
索引者は、主題分析を行い、その結果に基づいて、適切な索引語(キーワード)を付与します。

また、アクセスポイントの抽出と並行して、所在表示の記述(入力)作業を行います。

<情報のサイクル>



<索引作業の流れ>



7 検索フィールド

情報検索は、データベース全体の中からキーワードや著者名などを入力して、キーワードに合致する情報を検索します。

データベースを構成する個々の単位をレコードとよび、各レコードは検索フィールドから構成されています。

- | | |
|--------------|--------|
| ①レコード番号 | ⑥本文の言語 |
| ②タイトル | ⑦抄録 |
| ③著者名および所属機関名 | ⑧統制語 |
| ④情報源(雑誌名など) | ⑨準統制語 |
| ⑤資料種類 | |

-
- ①AN 2004019512
- ②TI F.ナイチンゲールにおける「人間の発達過程に関する見方」真理を理解し受け入れる能力とその発達過程について
- ③AU 和住淑子(千葉大学 看護学部)、青木好美
- ④SO 総合看護
ISSN 0038-0660
VOL.38 NO.3;PAGE 17-30(200308)
- ⑤DT 原著論文
- ⑥LA Japanese
- ⑦AB 「思索への示唆」は、F.ナイチンゲールの思想の形成期から執筆されているとされ、彼女の見方・考え方を知る上で重要な著作であるといわれている。ナイチンゲールの「人間の発達過程に関する見方」を探るため。。。
- ⑧CT 看護理論;人間の発達;医学書誌
- ⑨ST ナイチンゲール、F

7-1 検索フィールドの限定

データは、レコード番号順に格納されており、各フィールドの中の語を対象に検索してきます。検索を行う際に、キーワードの入力のほかに、検索フィールドを指定することもできます。

(1) 検索フィールドがあらかじめ指定されている場合

検索したい検索フィールドにカーソルを合わせてキーワードや検索式を入力して検索します。

(2) 検索フィールドが指定されていない場合

原則としてキーワードの前に検索フィールド(検索タグ)を付けたり、検索語の後にスラッシュ(/)で検索フィールドを限定し検索します。

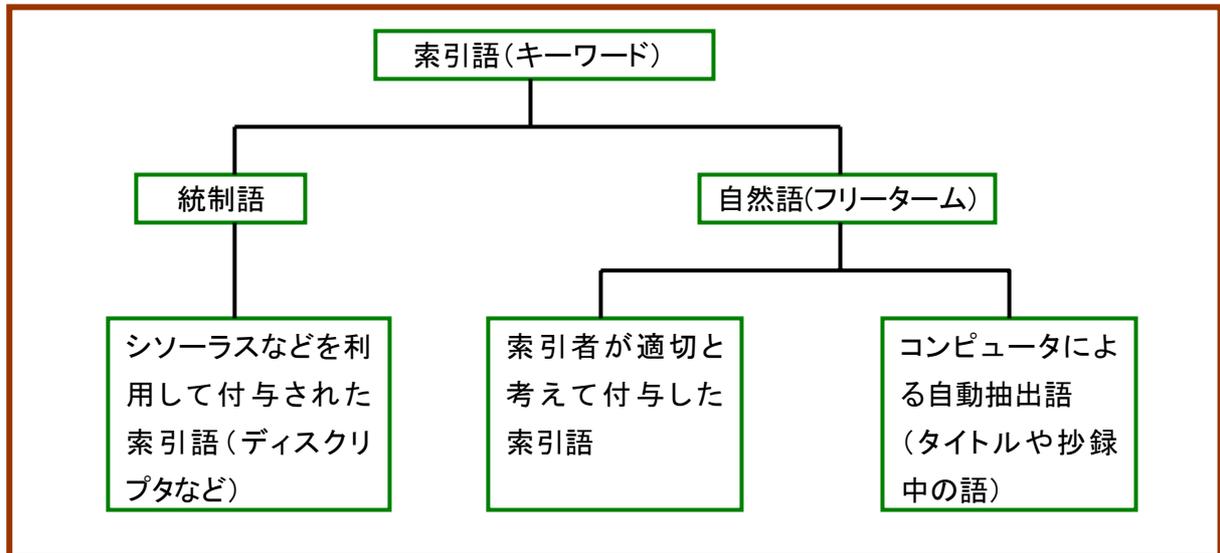
例】 TI cancer and YR 2002

例】 田中一郎/AU

8 索引語（キーワード）の選択

8-1 統制語（ディスクリプタ）とフリーキーワード

主題分析の結果、文献に付与される索引語（キーワード）には次のような種類があります。



「統制語」---例えば、アルコール中毒に関するデータを検索するには、「アルコール中毒」ではなく「アルコール症」という言葉を使いましょう、といった約束ごとに基づいて使用されるキーワードのことです。

これを用いることにより、Aという人は「アルコール中毒」で検索し、

Bという人は、「アルコール症」で検索したために結果が異なるという弊害が防げます。

但し、統制語を用いて検索する際には、シソーラスなどのキーワード集が必要です。

注意

キーワード集が作成された以降に登場した新しい言葉で検索する場合は、統制語での検索は、比較的困難です。

「フリーキーワード」---統制語のような言葉に対する統制を加えていない、全く思い付いた言葉で検索するためのキーワードのことです。

「自由語」「フリーターム」「フリーワード」とも言います。

注意

新語・造語・固有名詞に対応しやすいといえますが、前述の「アルコール中毒」と「アルコール症」の両方を検索する必要があります。

このように情報を漏れなく検索する場合には考えられる限りの関連語・同義語を列挙しなければなりません。

8-2 シソーラスの役割

シソーラス(thesaurus)の語源は、「宝庫」を意味するギリシャ語です。英語では、「倉庫」および「辞書」という意味で用いられていましたが、1852年にイギリスの医師ロジェ(Peter Mark Roget, 1779-1869)が、「Thesaurus of English Words and Phrases」を刊行してから、「シソーラス」には語句の意味によって分類配列し、各語句の同義語や類義語、上位・下位語などを記述した辞典という意味が加わりました。

「シソーラス」には、以下のような役割があります。

(1) キーワードの表す概念の階層関係を明示。

複数のキーワード間の関連は、その分野に精通していないとわかりにくいものです。

専門家でない人でも、例えば、

「コンピュータ(上位)」>「パソコン(中位)」>「ノートパソコン(下位)」

といった階層関係が分かれば、検索の際のヒントになります。

また、概念の縦の関係だけでなく、関連語といった横の関係も示されています。

(2) 利用者を非統制語から統制語に導く。

統制語として採用されていないキーワードを「非統制語」と言います。

利用者が思いついた言葉が非統制語の場合、シソーラスを利用することによって対応する統制語を探し出すことができます。

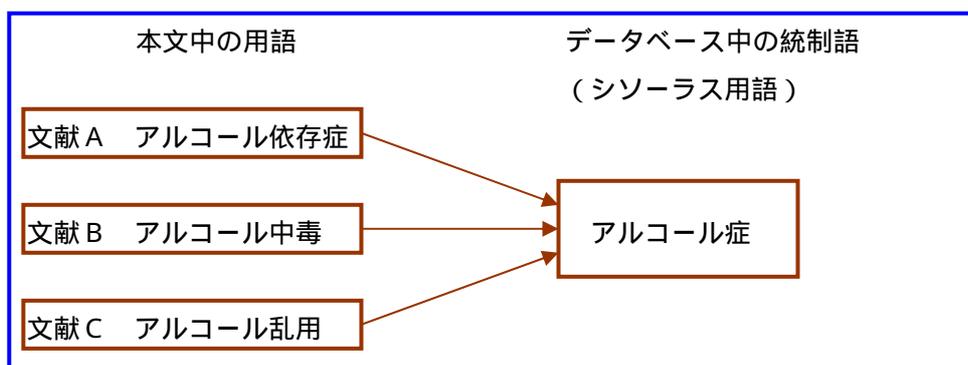
統制語で検索するファイルの場合には、キーワードが統制語か非統制語かをチェックすることは、非常に重要です。

(3) 索引者(インデクサー)がキーワードとして付与。

索引者が論文を読みシソーラスの中から適切な用語を選択するため適合率は高いといえます。

コンピュータが自動付与する場合(自然語)

データの中にある単語の意味を考慮せず、文字通り機械的に認識するため、例えば、「パソコン」と入力すると「パソコンショップに強盗が入った」などという、目的外のデータが検索される可能性が比較的大きくなります。



8-2-1 医学用語シソーラス

医中誌 Web では医学用語シソーラスを参照できる機能(候補語辞書検索)が搭載されています。

統制語リスト(シソーラス用語, フリーキーワード) ▶ 閉じる

入力した語... 産業事故

採用

同義語	見出し語	種別
産業事故	労働災害	シソーラス

同義関係

採用

シソーラス

▶ 閉じる

労働災害

「労働災害」はシソーラス内で1個所の階層構造に位置しています。

下位語も検索する

採用

環境と公衆衛生(g3+)

衛生学(g3-10+)

公衆衛生(g3-10-30+)

事故(g3-10-30-130+)

労働災害(g3-10-30-130-90+)

放射線事故(g3-10-30-130-90-10)

上位語

下位語

8-2-2 MeSH(MEDLINE のシソーラス)

MeSH

ページ: 前へ | 次へ | ブラウズ: AIDS 検索

同義関係

分類を選択し、次の条件で検索に追加する: or 追加 拡張

(用語をクリックすると、詳細が表示されます)

AIDS 使用: [Acquired Immunodeficiency Syndrome](#)

Home Health A シソーラス

ツリー表示

[Virus Diseases +](#)

[RNA Virus Infections +](#)

[Retroviridae Infections +](#)

[Lentivirus Infections +](#)

[HIV Infections +](#)

Acquired Immunodeficiency Syndrome

[AIDS-Associated Nephropathy](#)

[AIDS Arteritis, Central Nervous System](#)

[AIDS Dementia Complex](#)

上位語

下位語

検索:

ボジャー コンセプト

サブヘディングに限定?

Blood

Cerebrospinal Fluid

Chemically Induced

Classification

Congenital

Complications

Diet Therapy

8-3 キーワード選びのコツ

いかに検索テーマに合致したキーワードを選択するか、これが検索精度を高める重要なポイントといえます。

(1) データベース専用のシソーラスを活用すること

キーワードは、標題あるいは抄録に出現しているフリーワードや同義語等をコントロールしているシソーラスの統制語(ディスクリプタ)が対象となります。

そこで、著者の標題標記によっては、検索の適合率を下げる結果があります。

検索漏れやノイズを防ぐために、データベース専用のシソーラスを活用することが肝要です。

(2) フリーキーワードの利用に注意すること

専用のシソーラスを持っていないデータベースは、著者が表現している語がキーワードになり、同義語等をコントロールしていませんので、次のことに注意してください。

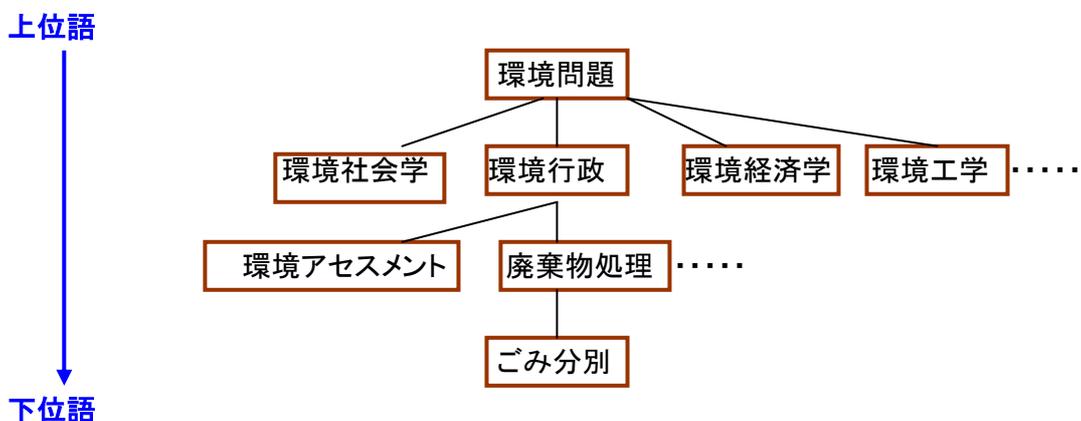
- ① 同義語・関連語を網羅すること。
- ② 単数形と複数形の語尾変化や英国式と米国式の綴りの違いを考えること(英語 DB の場合)

(3) シソーラスとフリーキーワードとを組み合わせること

フリーキーワードと統制語または、分類コードが併用できるシステムでは、抽象的な概念を統制語や分類コードで、固有名詞など具体的な概念はフリーキーワードで、と組み合わせると検索するのが効果的です。

(4) 上位語と下位語を検討すること

用語には、より概念の広いものと狭いものがあります。



「環境問題」は「環境社会学」や「環境行政」、「環境経済学」などを含んだより包括的な用語です。このとき「環境問題」と「環境行政」は、上位と下位の関係になります。

さらに「環境行政」の下位には「環境アセスメント」や「廃棄物処理」のような用語が属します。

広いテーマで検索したいときはより上位の用語を選び、特定のテーマに絞りたいときはより下位のキーワードを選ぶと適切な結果を得ることができます。

(5) 同義語・関連語を検討すること

より網羅的に検索するには、同義語・関連語を使って検索をする必要があります。

「環境問題」の関連語としては「エコロジー」や「公害」などが考えられます。

どのような同義語・関連語があるかは、一度得た検索結果の書誌情報をよく読み、そこに含まれる用語を調べることで分かります。また、辞書・辞典などにより調べることも有効です。

(6) 自動マッピング機能の取り扱いに注意すること。

自動マッピング機能とは、考えついた用語をキーワードとして入力すると、システム側で適当なシソーラス用語に置き換えて検索してくれる機能のことです。

仕組みは、システム側に搭載している「同義語辞書ファイル」と入力した用語をマッチングさせ、同義語・関連語を自動で検索します。

なお、マッピングされた用語が不適切な場合は、解除することも必要です。

例】医中誌 Web PubMed

(6) 複合語の扱いに注意すること

「環境行政」は「環境」と「行政」という単語からなる複合語です。

データベースによっては、「環境行政」1語での複合語検索が可能であったり、「環境」と「行政」と2語に分けて検索しなければならない場合があります。

また、ダブルクォーテーション(“ ”)でキーワードを囲むとフレーズ検索を行うデータベースもあります。

(7) EXPAND(エクスパンド)機能を利用すること (参照⇒p. 21)

キーワードをアルファベット順または五十音順に表示します。

例】『Yahoo』『Google』⇒ “評判のいい病院”

8-4 ストップワード

出現頻度の極めて高い、したがって検索キーワードとしては意味をもたない文字列を事前に登録しておき、検索対象から除外される用語を一般にストップワード(不要語)と呼びます。英文データの場合には、冠詞、接続詞、前置詞などが典型的なストップワードとされます。

例】「の」「は」「to」「of」「the」

9 検索機能

9-1-1 論理演算機能とは

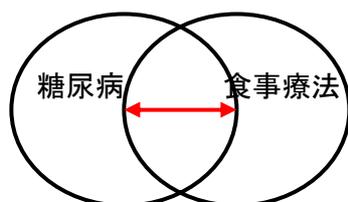
検索に用いるキーワードの選定が終了すると、次に、複数のキーワードを組み合わせることでキーワード間の関係を定義します。複数の概念が互いにどのような関係にあるか、概念と概念の関係を定義することを「**論理演算**」といいます。

例えば、『糖尿病についての文献はないか』というテーマでは、「**糖尿病**」だけがキーワードになるので、論理演算の必要はありません。

しかし、「糖尿病が高脂血症のどちらかを持った人で肥満について扱った文献」というテーマでは、それぞれのテーマが、複数の概念（**肥満**、**糖尿病**、**高脂血症**）から成り立っているため検索に際してはキーワード間（**肥満**、**糖尿病**、**高脂血症**）の論理演算が必要になります。

論理演算には、AND演算、OR演算、NOT演算があります。

①AND 検索

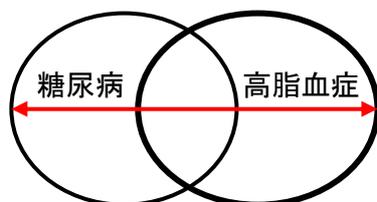


それぞれのキーワードが**両方とも付与**されているデータを検索します。

「糖尿病の食事療法」→「糖尿病 **AND** 食事療法」

「パソコンのソフト」 →「パソコン **AND** ソフト」

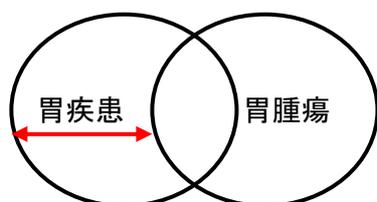
②OR 検索



それぞれのキーワードの**いずれか、あるいは両方**が付与されているデータを検索します。

「糖尿病または高脂血症」→「糖尿病 **OR** 高脂血症」

③NOT 検索



あるキーワードが**付与されていない**データを検索します。

「胃腫瘍以外の胃疾患」→「胃疾患 **NOT** 胃腫瘍」

注意 キーワードの位置に注意してください。

「胃疾患 NOT 胃腫瘍」と「胃腫瘍 NOT 胃疾患」では、まったく結果が異なります。

NOT の後にくるキーワードが「除かれる」キーワードになります。

注意

システムによっては、AND、OR、NOT の代わりに「*」「+」「#」等の記号が使える場合がありますので、必ずマニュアル等で確認してからご利用ください。

9-1-2 論理演算の優先順位

AND、OR、NOT の演算を一度に組み合わせる場合、これらの演算の優先順位に注意する必要があります。

一般にカッコ「()」を用いた演算、次ぎに NOT 演算、そして AND 演算、最後に OR 演算の順に優先されます。

順位を間違えると、期待した結果が得られないことになります。

(例1)「肥満と糖尿病か高脂血症をみたく文献」について検索するつもりが…

×

とすると、「高脂血症」と「肥満」の間の AND 検索が行われた後、その結果と「糖尿病」が OR 演算されます。

従って、高脂血症はある程度肥満に限定されますが、糖尿病については肥満以外に関するデータも検索する結果になります。

正解

(例2)「日本以外のアジアまたは環太平洋国家」について検索するつもりが…

×

「環太平洋」と「日本」の間の NOT 検索後、その結果と「アジア」が OR 演算されます。

従って、環太平洋諸国の中から日本は除かれますが、アジア諸国の中には日本は含まれたままとなります。

正解

注意

システムによってカッコが使えなかったり、優先順位が異なったり、カッコとキーワードの間に空白が必要なものや、逆に空白があってはならないものなど様々です。

また、「キーワード入力欄」が複数用意されているデータベースの場合は、論理演算子をプルダウンメニューから選択します。この場合、優先順位は縦のラインで実行されます。

The screenshot shows a search interface with a search condition input field containing 'stomach or intesine and perforation'. The search condition is enclosed in a green circle. To the right of the input field is a 'MEDLINE' button. A green arrow points from the search condition to the 'MEDLINE' button. The interface also shows a search button and a search results section.

検索の際には必ず、各システムのマニュアルを参照してください。

9-2 トランケーション機能(ワイルドカード)

9-2-1 トランケーション機能とは

トランケーション(truncation)とは、円錐などの先端を切断するという意味です。

トランケーション機能は、英語等の語尾変化や単数形の綴りの一部分が異なるものを検索したい場合に発揮する機能です。

任意文字とする部分に使用する入力文字を、《**マスク文字**》あるいは《**ワイルドカード**》といいます。マスク(mask)とは、覆い隠すという意味です。

システムによってマスク文字は、「?」、「!」、「\$」、「*」、「@」マークなどが使用され、それぞれ意味が異なったりするので、システムのマスク文字の意味を確認してから使用してください。

例えば、「Develop?」と入力すると「Developer」「Developments」という「Develop. . .」で始まる単語を検索してきます。

この機能を利用した検索に、①前方一致検索②中間一致検索③両端一致検索④後方一致検索があります。

9-2-2 前方一致検索

前方一致検索——語の頭の部分が共通で語尾を特定しない場合に使います。

例】「心臓?」と入力すると「心臓」の他に「心臓外科」「心臓移植」も検索。

9-2-3 中間一致検索

中間一致検索——中間だけが共通で語の頭と語尾を特定しない場合に使います。

例】「?電子?」と入力すると「電子」の他に「超電子」「高速電子解析」も検索。

9-2-4 両端一致検索

両端一致検索——語の両端が共通で中間の語を特定しない場合に使います。

例】「情報?システム」と入力すると「情報システム」のほかに「情報検索システム」「情報管理システム」も検索。

9-2-5 後方一致検索

後方一致検索——前方一致検索と逆に語尾が共通で語の頭の部分を特定しない場合に使います。

例】「?癌」と入力する「癌」の他に「胃癌」「腎臓癌」「大腸癌」も検索。

9-2-6 トランケーションとワイルドカード

意味が異なったトランケーション機能を持っているシステムの場合は、使いかたによってノイズ（不要情報）が検索されますので注意が必要です。

例】 MEDLINE/CINAHL/Biomedical Reference Collection の場合

タイプ	記号	意味	例
トランケーション	*	0文字以上の文字の代わりになります。	“William, H*” ★外国人の氏名の場合は、セカンドまたはサードネームを漏れなく検索するために使用してください “cat*” ⇒ “cats” + “catalog” + “category”を検索
ワイルドカード	?	1文字または0文字の代わりになります。	“cat?” ⇒ “cats”を検索 “an?esthesia” ⇒ “anesthesia” + “anaesthesia”を検索 “wom?n” ⇒ “woman” + “women”を検索

猫(cat)を検索したいのに、“cat*”と入力すると、“cat”や“cats”も検索されますが、“catalog”や“category”なども検索されます。

このように短い語幹に対して文字数無制限のトランケーションを行うと、様々なものが検索されてノイズの原因になるので注意しなければなりません。

例】 insideWeb の場合

タイプ	記号	意味	例
トランケーション	\$	0文字以上の文字の代わりになります。	“antibod\$” ⇒ “antibodies” + “antobody”を検索

9-2-7 トランケーション機能使用の注意点

システムによっては、すべてのタイプのトランケーションが使えるもの、前方一致検索のみ使えるもの、中間一致検索は使えないものなど様々です。

キーワードの選び方で、利用される事の多い言葉でトランケーションを使うと、検索結果が表示されるまでにかなり時間を必要とする場合があります。

どのシステムでどのトランケーションが使えるか、あるいは時間がかかるか、といったことは各システムのマニュアルを参照してください。

9-3 横断検索

複数のデータベースを同時に同じインターフェイスで検索することができます。

また、検索システムによって重複レコードを除去する機能備えられています。

同じインターフェイスで一度に検索が可能で便利ですが、データベースの組み合わせにより検索項目や検索機能が制限されますので、注意が必要です。



9-4 EXPAND (エクスパンド) とは

キーワードを入力するとき、どのようなキーワードを入力すればよいか迷うことがあります。

このようなとき、「EXPAND」という機能を利用すると便利です。

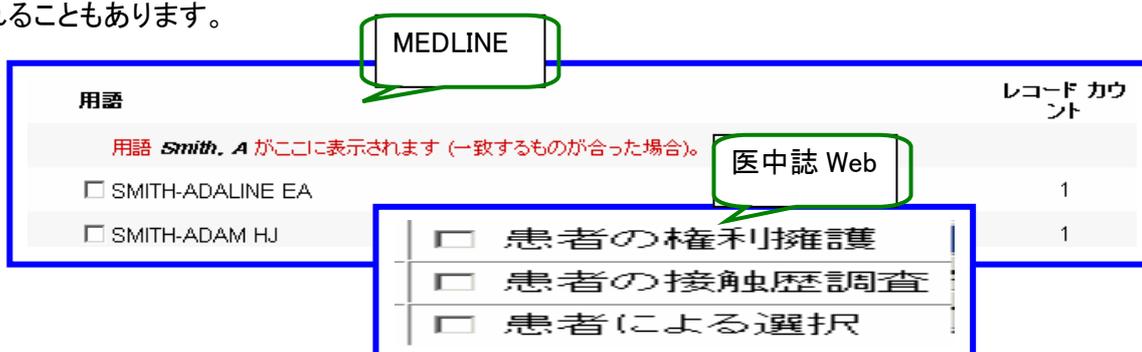
EXPAND は、アルファベット順または五十音順にキーワードを表示する機能です。

また、キーワード表示だけではなく現在収録されている件数が表示されるものもあります。

この機能を利用すれば、最初は「タバコ」という言葉しか思い付かなかった場合でも、「タバコ依存」「タバコ煙公害」「タバコ産業」など、語順が近い言葉は次々と表示され、的を得たキーワードが見つかる場合があります。

また、英文データベースで米語・英語で綴りが違ったり、社名や人名等が短縮されている可能性がある場合は、EXPAND を使うとスペルが近い単語を列挙するため、スペルミスのチェック、或いは人名等の省略形の確認に便利です。

なお、EXPAND 機能の名称は、システムにより、「ブラウジング」「検索語一覧」などの名称が用いられることもあります。



9-5 近接演算機能

近接演算とは、複数のキーワードが同じ検索項目の中で一定の位置関係にあるときに限定して検索する方法です。

「Personal Computer」について検索する場合に、「Personal and Computer」と入力したのでは、「**Personal** Role for Operation of Super **Computer**」なども検索されてしまいます。

そこで、キーワードの位置関係を指定することにより、このようなノイズを除去することができます。

近接演算を指定する記号を《**近接演算子**》と言います。

なお、近接演算は、英文データベースで利用できますが、日本語データベースでは少ないようです。これは、日本語が英語のように単語間の区切りが見た目に明確でなく、単語単位の近接演算が難しいことも原因の一つでしょう。

日本語データベースでは、複合語(フレーズ)検索でノイズを除去することができます。

inside Web	gastric adj cancer	10,308件
	gastric and cancer	12,611件

ポイント 近接演算機能を上手に利用することにより、適合率の向上が期待できます。

9-6 完全一致検索機能

入力されている文字列が検索の対象となるシステムでは、不要な情報を検索してきます。

そのノイズを除去する機能として、入力したキーワードの文字列に完全に一致する文字列を含むもののみを検索する完全一致検索があります。

システムによって、入力するパターンが異なりますので、マニュアルで確認してください。

医中誌 Web	[藤本淳]	78件
	藤本淳	146件 (藤本淳也+藤本淳治なども検索します。)

9-7 範囲検索機能

記事検索における発行日や企業検索における売上高など、数値情報を検索する時は、〇〇から××までといったように、データの範囲を指定して検索することができます。これを「範囲指定」といい、データを「:(コロン)」や「-(ハイフン)」などでつないで範囲を指定します。

また、「<」や「>=」といった比較記号を用いて、「〇〇未満」「××以上」といった指定を行うこともできます。

範囲指定は、キーワードなどで分野を限定せずに絞り込む際の有効な方法の一つですが、データベースのファイルによって使用する記号が異なっていたり、範囲を広く取りすぎると検索に時間がかかってしまうこともあるので、事前にマニュアルなどで確認することが大切です。

9-8 検索式一覧機能(検索履歴の表示)

これまでに入力した検索式を表示するための機能です。

検索の途中でこれまでにどのような検索式を入力したかを確認したり、それらの集合番号を利用するために以前に入力した検索式を見ることができます。

9-9 検索式の保存

検索システムを終了する前に、作成した検索式を保存しておいて、再度検索システムを起動させたときにその検索式を呼び出して利用できる機能です。

同じテーマで何度も検索を行う場合や検索の途中で終了せざるを得ない時などに便利です。

また、検索を続けながらデータベースだけ変更するときに、それまでのデータベースに対して作成した検索式をそのまま使いたい場合などにも利用できます。

10 検索結果の絞り方

検索した文献が多い時、どのような絞り方があるかを紹介します。

おすすめ

(1)サブヘディング(副標目)を使用

サブヘディングは、その主題がどのような視点から扱われているかを限定します。

シソーラスと組み合わせてより検索効率の良い検索ができます。

The screenshot shows a search interface with a section titled "副標目" (Sub-headings) and a list of sub-headings to limit results. The sub-headings are categorized into four groups:

- 1群: 薬物・化学物質に関する副標目 (Sub-headings related to drugs/chemical substances)
- 2群: 診断に関する副標目 (Sub-headings related to diagnosis)
- 3群: 治療に関する副標目 (Sub-headings related to treatment)
- 4群: 疾患の原因などに関する副標目 (Sub-headings related to causes of diseases, etc.)

Each group has a dropdown menu to select a sub-heading. The "副標目" section also includes a note: "※Ctrlキーを押しながらクリックすると複数選択できます" (※You can click multiple items while holding the Ctrl key).

The "サブヘディングに限定する:" (Limit by sub-heading:) section shows a list of sub-headings with checkboxes to select them. The list includes:

- Blood
- Cerebrospinal Fluid
- Chemically Induced
- Classification
- Congenital
- Complications
- Diet Therapy
- Diagnosis
- Drug Therapy
- Economics
- Microbiology
- Mortality
- Nursing
- Pathology
- Prevention & Control
- Physiopathology
- Parasitology
- Psychology
- Radiography
- Rehabilitation

A "選択をクリア" (Clear selection) button is located at the top right of the sub-heading list.

Annotations in the image include:

- A red box around the "おすすめ" (Recommended) text.
- A green box around the "MEDLINE" text.
- A green box around the "医中誌 Web" (Ishinshi Web) text.
- A red box around the text: "シソーラス用語と同様にインデクサー(索引者)が付与していますので、適合率は高いといえます。" (Since indexers are assigned to MeSH terms, the fit rate is high.)

- (2) 雑誌(本)の発行年を限定
- (3) 収録雑誌を限定
- (4) 使用言語を限定
- (5) チェックタグで限定

標題および抄録中に極端に出現する一般的概念をもつ用語がチェックタグとして用意されています。

- ① 研究対象を限定(人・動物等)
- ② 性別を限定(雄性・雌性)
- ③ 研究の種類を限定
- ④ ヒトの年齢区分を限定
- ⑤ 標題に出現する用語のみを特定

おすすめ

11. アラート機能

関心のあるテーマを継続的に追跡したい場合があります。

……こんな時あなたならどうしますか？

日を決めて定期的に同じ検索を行うことが考えられます。

……ついつい忘れてしまうことがあるかもしれません。

……毎月、毎週あるいは毎日同じ検索をするのは面倒です。

あらかじめテーマに基づいて、それにマッチする検索式を登録しておく、データベースに新しい情報が追加されたときに、自動的に検索され文献情報をメールで受け取ることができるサービスです。

→とりこぼしなく情報収集を行うことができます。



12. ドキュメント・デリバリー・サービス

文献の複写を郵送の他に EDD、FAX 等で取り寄せるサービスです。

文献をすぐ入手したい場合は、大変便利です。

但し、文献の手配料金に著作権料金が加算されますので相互貸借より割高になります。

情報検索は“習う”より“慣れる”といわれています。

検索の上達は、試行錯誤を繰り返し経験を積むことです。

ただし、各データベースのマニュアルをチェックすることをお忘れなく！！