

第4章 情報検索・文献検索

この章について

高等学校までと大学が決定的に異なる点のひとつとして、「知」の蓄積量があげられます。「はじめに」にも書いたように本書でいう「知」は書籍・学術雑誌・オンラインジャーナル・データベースなどの「文献」を意味することとします。早稲田大学の場合、中央図書館、高田記念図書館、戸山図書館、理工学図書館、所沢図書館といった各キャンパスの中核となる図書館が設置されています。また独自に学生読書室を設置している学部もあります。この学生読書室は小規模ではありますが、それぞれ約5万冊と、普通の高等学校の図書室よりも蔵書数が多く、また各学部の専門性に沿った選書が行われています¹。系統的な学習体系をとる学問分野では、この選書というのは重要です。手当たり次第に本を読むのではなく、学習の基礎段階で「評価の定まった文献」からきちんと学ぶことが重要です。大学全体としての蔵書数は約478万冊、最新の学問発表の場である学術雑誌は5万種類、ネットワーク経由で見ることのできるオンライン・ジャーナルは6000タイトルを超え、貴重書、研究図書など一部の例外を除いて、大学の構成員であればこれらをほぼ自由に使うことができます。多少の規模の違いはあれ他大学でも同様で、初等・中等教育とは段違いの文献が蓄積されているはず²。このような充実した環境を構築するために、大学は毎年多額の予算を費やしています。それはこれら文献が学術そのものを支える基盤となっているからです。

この膨大な「知」たる文献を活用しない手はありません。こと英文を含めた学術書に関しては大学の図書館におよぶところはありません³。しかしあまりに膨大なため、「あるテーマに沿った図書を探し出す」という目的の時は、書棚を眺めたり、カード目録⁴をめくり本を探すといった方法では、もはや目的を達成できません。いまや、目録は電子化され、コンピューターで検索をするようになっています。

印刷物の本を扱っているという点で誤解を受けがちなのですが、図書館は大学の中で最も情報化が進んでいる部署の一つです。ですので、皆さんも図書館を漫然と利用するのではなく、情報化に対応した使い方を身につけないと、効率よく、また有効に使うことができません。例えば自宅からインターネット経由で書籍を検索し、貸し出し予約をかけたたり、所蔵していない図書を、提携している他大学から取り寄せる依頼をしたり、購入依頼をすることもできます。それだけではありません。大学が契約をしている電子書籍、データベース、オンライン・ジャーナルはPC上でその全文を見ることができます⁵。

では、図書館の情報化に対応するためにはどうすればよいのでしょうか。大量に蓄積された文献の中から必要なものを探すということは昔は大変な労力を必要としましたが、情報化された図書館では多少の知識があると比較的簡単にそれを抽出できます。GoogleやYahoo!などのWeb検索エンジンなどでもそうですが、検索方法（検索式）を工夫することによって飛躍的に検索結果を絞り込むことができます。本章ではそれら情報化された図書館環境の使い方について学んでいくことにしましょう。

¹その他に教員が主に使う教員図書室・研究所図書室が設置されています。利用条件など詳しくは <http://www.wul.waseda.ac.jp/etclibs.html> を参照してください。

²大学を新たに設置するときの審査において、図書館の蔵書数も判断基準となっているほどです。

³国立国会図書館は、国内で発行されたほぼすべての図書・雑誌が収められているという点では別格です。

⁴既に紙のカード目録が存在しない図書もあります。

⁵ただし契約上、学内のネットワークに接続した端末からしか参照できないことが多いことに注意してください。

4.1 情報検索の方法

4.1.1 データベースの選択

ここで「検索」とは、あらかじめ蓄積されているデータ（データベース）⁶から自分の求める情報を見つけ出すことを言います。

例えば図書館の検索であれば、以前ならば図書カードをめくったり本棚を巡ったりといった方法がとられていましたが、昨今では様々なデータが電子化されており、キーワードを指定することで目的とする情報の収められている書籍を探し出すことができるようになってきました。

この際、蓄積されているデータが多くなければ、キーワードを与えて検索した結果もさほど多くないので、検索結果を吟味するのに大きな労力は必要ありません。しかし、膨大なデータベースに対して一般的なキーワードを与えると膨大な量の検索結果が得られてしまいます。例えば、早稲田大学蔵書検索（WINE）に「早稲田」というキーワードを与えて検索すると11,709件の該当があり⁷、とてもすべての検索結果について検討することはできません。

一方で、どのようなデータベースが存在しており、それぞれのデータベースをどのように利用することができるのか、というのは場合によって異なります。

無償の、公開されたデータベースもあれば有償のデータベースもあります。有償のデータベースでも、早稲田大学の学生や教職員であれば無償で利用できるものもあります。また、オンラインのみで提供される情報源（Web ページや後述する Wikipedia など）、オンラインでも印刷物でも提供されている情報源（新聞記事や政府刊行物など）、基本的には印刷物のみで提供されている情報源（書籍や学術雑誌など）といった分類をすることもできるでしょう。

従来、印刷物のみで提供されている情報源を探すのは大きな労力を必要としましたが、WINE や各種論文検索データベースなどは、印刷物を手軽に検索できる手段を提供してくれるシステムであるといえます。

ここでは、無償のデータベースとして Web の検索エンジン（Google）を利用した情報検索を紹介するとともに、検索の一般的な考え方について学習します。また、早稲田大学の学生が利用することのできるデータベースの一部を紹介します。データベースによっては、教員図書室等でのみ利用することのできるものもありますので、自分の利用したい種類のデータベースが利用できるかどうか、図書館のレファレンスサービス⁸を利用して照会してみるのも良いでしょう。

初期の検索エンジンは Web 文書のみしか扱えないことが多かったようですが、現在では Web、PDF、Word、Excel、画像、動画など、キーワードを与えるだけで様々な情報を横断的に検索することができるようになりました。

特に最近ではブログ（blog）に代表される、Consumer Generated Media（CGM）と呼ばれるマイクロコンテンツ⁹が検索において比較的重視される傾向が高いことにも注意が必要です。CGM の代表的なサイトが Wikipedia¹⁰です。Wikipedia は非常に便利なサイトです。学術論文やレポートを作成する際に、調べごとの導入に利用するといった程度で参考にすることもできなくはありませんが、一般的に Wikipedia における記述を根拠に論を組み立ててはいけません。

これは Wikipedia に限らないことで、たとえ「ブリタニカ国際大百科事典」であっても不適切です。論文やレポートの引用に耐えるのは一次資料と、その一次資料に基づいて作成された論文のみ

⁶「データベース」とは、直訳すれば情報の基地ということになりますが、様々な情報を集めて管理し、検索や抽出、更新などの再利用をしやすいものを言います。コンピューター上に構成されたものだけでなく、図書館などもデータベースであると言えます。

⁷2011年2月時点の結果です。2007年5月には7,037件でしたので、順調に増加していることが分かります。

⁸<http://www.wul.waseda.ac.jp/CLIB/ref-j.html> 参照のこと。

⁹コンテンツ1つあたりの情報量が少ないが数が多く、また増える速度が速いようなコンテンツを総称して「マイクロコンテンツ」と呼びます。

¹⁰<http://ja.wikipedia.org/>

です。性質上、書籍に最先端を求めるのは難しいところがありますが、専門の論文誌であれば最新かつ信頼性の高い情報を入手することができます¹¹。

もともと、Google Scholar¹² や Google Books¹³ のように Web の検索エンジンでも信頼性の高い情報源からの検索サービスを提供しているケースもありますので、データベースといっても様々な質のものがあるということです。その質についての判断は、最終的に書き手と読者の手に委ねられています。

このように、情報検索では（１）データベースの選択、（２）効率的な検索、（３）検索結果としての情報源の信頼性の吟味の３つが必要であることに注意してください。

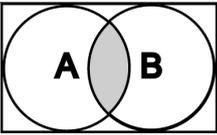
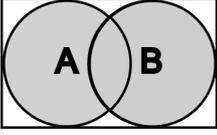
4.1.2 ブール演算子

どのようなデータベースから検索する場合であっても、条件を与えてデータの中から情報を抽出して絞り込み、その中から更に自分の欲しい情報と近いものをより分けていくというのが通常です。そこで、あるデータの中からどのようにして情報を絞込むのかという方法について説明します。

ここでは、あるデータの集まりを１つの集合と見なして、これに集合演算を行うことで新たな集合を得る、ブール演算という方法について解説します。このように書くと難しそうに聞こえますが、実際にはキーワードをいくつか指定する際にちょっとした工夫をすると、より便利に検索できますということです。

表 4.1.2 を参照してください。AND はキーワードを追加していくことで、絞り込むことができるということを意味します。逆に OR は、キーワードを追加することで範囲が広がります。NOT は検索結果から一部を除外することができます。

表 4.1: ブール演算子（論理演算子）

演算子	説明	ベン図
AND	論理積 両方の検索語を含む	
OR	論理和 いずれか一方の検索語または両方の検索語を含む	
NOT	否定 一方の検索語を含むが、他方の検索語は含まない	

この他にも NAND や XOR、NOR などもありますので、興味があれば調べてみてください。

ここでは、検索エンジンとして Google を例にとりてこの検索方法を試してみましょう。Web ブラウザで <http://www.google.co.jp/> にアクセスしてください。ここでは例として、温泉に関する情報を調べてみることにします。

最初に、AND について確認しましょう。Google では検索結果の件数が表示されます。次のようにキーワードを増やしながら、検索結果の件数が何件になるか確認してみてください。キーワード

¹¹もちろんその場合であっても議論の妥当性を批判的に検討しなければならないのは同じ事です。

¹²<http://scholar.google.com/intl/ja/>

¹³<http://books.google.com/>

第4章 情報検索・文献検索

が複数の場合、それぞれをスペースで区切りながら入力してください。

- 温泉
- 温泉 日帰り
- 温泉 日帰り 露天
- 温泉 日帰り 露天 電車

2012年1月現在の検索結果では、上記4つのキーワードによる検索結果件数は、それぞれおよそ「391,000,000」「63,500,000」「27,000,000」「7,090,000」でした。キーワードを追加すると、検索結果が減少していくのが分かるはずです。

次に OR について見てみましょう。次のようにキーワードを指定します。

- 箱根
- 箱根 OR 草津

ここで、Google の場合、「OR」は大文字である必要があります。これらのキーワードで検索した結果では、それぞれ「46,100,000」「71,800,000」となり、検索に当てはまったページ数が増加していることが分かります。箱根ないし温泉のどちらかのキーワードが含まれる Web ページが検索結果として表示されるためです。

最後に NOT ですが、次のように指定してください。

- 温泉
- 温泉 -箱根

含めたくないキーワードの前に、「-」（マイナス記号）を付けます。これで、箱根というキーワードを含む検索結果が除外されます。件数が大きく減ることを確認してください。

最後に注意したいことですが、前述の「OR」について2008年2月に行った検索では、逆に検索結果が減少するという現象が見られました。Google は検索結果を「賢く」操作することもあります。また、検索する人によっても結果を変更する可能性もあります。検索する時期によっても表示される件数は異なるでしょう。Google に限らないことですが、検索エンジンを利用した検索結果にのみ引きずり回されるのは、決して良い事ではないことは覚えておいて下さい。

4.2 学術情報リソースと実際の検索

ここでは、学術情報リソースとそれを対象とした検索を見ていくことにします。これらは学術情報リソースは組織外に公開されているとは限りません。ここでは早稲田大学図書館のシステム（WINE）の検索例と、主に無償でつかえるデータベースについて紹介します。

4.2.1 WINE-OPAC

OPAC（Online Public Access Catalog）とは、電子化された目録である MARC（MACHINE READABLE Cataloging）をネットワーク接続されたコンピューターから検索可能なシステムを意味します。Web ブラウザからアクセス可能なものを特に Web OPAC と呼ぶこともあります。早稲田大学図書館には総合的な学術情報検索システムである WINE（Waseda university Information Network System）が設置されており、その一部である WINE-OPAC は Web OPAC で、学内外問わず広く公開されています。だれで

4.2. 学術情報リソースと実際の検索

も Web ブラウザを用いてインターネット経由でアクセスすることができます。ここでは WINE-OPAC を例にして検索を見ていきます。WINE-OPAC トップページ (<http://wine.wul.waseda.ac.jp/>) にアクセスしてください。



図 4.1: 早稲田大学 学術情報検索システム

ここでは、タイトル検索、キーワード検索を取り上げます。

検索画面

WINE - Waseda University Library Catalog

タイトル検索

[\[最初に戻る\]](#)

*入力域内をクリックした後、タイトルを入力し<Enter>キーを押して下さい。

図 4.2: タイトル検索

書籍のタイトル名での検索 タイトル検索 (図 4.2 参照) では、書籍名を前方一致で検索します。例えば、温泉というキーワードを与えて検索すると「温泉」「温泉：歴史と未来：」などといった書籍が検索結果として返ってきます。

検索結果が多すぎる場合は、「限定」というボタンを押してキーワードを追加して絞り込んだり、その他の書誌情報を与えて検索範囲を限定することもできます。

WINE: キーワード検索

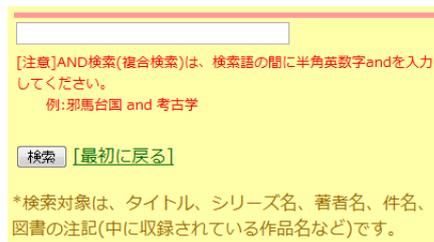


図 4.3: キーワード検索

キーワード検索 キーワード検索(図 4.3 参照)では、自由にキーワードを与えて検索することができます。ここで「温泉」というキーワードを与えると、タイトル検索とは異なる検索結果が出てくることに注目してください。

キーワード検索では、先ほどのブール演算を利用した検索が可能です。例えば、温泉かつ、箱根または草津というキーワードで検索する場合、「温泉 AND (箱根 OR 草津)」と入力してみてください。この際、検索キーワード以外はいわゆる半角文字で入力するのが一般的であることに注意して下さい。

その他にも様々な検索方法が用意されていますので、目的に合わせて使い分けると良いでしょう。

4.2.2 その他の学術情報リソース

蔵書検索 早稲田大学は膨大な蔵書数を誇りますが、早稲田大学だけですべての書籍を蒐集することができるわけではありませんので、場合によっては別の図書館の蔵書を検索して、必要に応じてそこから借り出したり一部の複写をお願いしたりといった作業が必要かもしれません。このような場合、WINE 以外の OPAC や Webcat と呼ばれるシステムを利用することになりますので、いくつか紹介しておきます。

- 国立国会図書館 NDL-OPAC
国会図書館の OPAC。
<http://opac.ndl.go.jp/>
- Z39.50 による OPAC 横断検索
「Z39.50」という米国規格協会が定めた標準的な通信規約を使う横断検索。
<http://wine.wul.waseda.ac.jp:211/z39m/>
- NACSIS Webcat
1200 を越える大学等の機関が参加する目録検索システム。
<http://webcat.nii.ac.jp/>
- NACSIS WebcatPlus
自然言語に近い言葉で検索できるという連想検索に対応しており、キーワードが曖昧な場合に有効。
<http://webcatplus.nii.ac.jp/>
- Google Scholar
特にアカデミックな分野に限定した検索エンジン。
<http://scholar.google.com/>

- Scirus (サイラス)
エルゼビア社によるアカデミックな分野に特化した検索エンジン。
<http://www.scirus.com/>
- NII 論文情報ナビゲータ
日本の学術論文を中心に情報を提供するサービス。
<http://ci.nii.ac.jp/>
- 早稲田大学リポジトリ (Space@Waseda University)
早稲田大学の研究者等が作成した学術論文、学位論文等の学術情報を保存・公開。
<http://dspace.wul.waseda.ac.jp/dspace/index.jsp>
- Books.or.jp
「国内で発行され、現在入手可能な書籍を収録する書籍検索サイトです。」
¹⁴<http://www.books.or.jp/>

この他にも有料ですが、JDreamII、Web of Science、SCOPUS など様々なサービスがあります。

早稲田大学図書館が契約しており、早稲田大学関係者であれば無償で利用することのできるサービスは膨大といいほどあります。早稲田大学図書館の「学術情報検索」のページ (<http://www.wul.waseda.ac.jp/imas/index.html>) にアクセスして、どのようなデータベースが利用可能か、一度確認しておくとい良いでしょう。

4.3 演習問題

- Google (<http://www.google.co.jp/>)、Yahoo! (<http://www.yahoo.co.jp/>)、Goo (<http://www.goo.ne.jp/>) のそれぞれで、教科書中の例と同じ検索を行い、それぞれ (1) 検索結果件数 (2) 検索結果として表示される Web ページのトップ 10 のリストを比較しなさい。
- 早稲田大学図書館 (<http://www.wul.waseda.ac.jp/>) にアクセスし、大隈重信侯が著者として登録されている本の冊数を調査しなさい。
- 2 の検索結果をもとに、出版されたのが 1920 年以降に出版されたものに絞り込んで冊数を調査しなさい。
- 早稲田大学図書館のサイトから、日本経済新聞社の各種雑誌をオンラインで閲覧する方法を探しなさい。
- Google ブックス (<http://books.google.co.jp/>) にアクセスしなさい。多くの書籍を「立ち読み」することができることを確認しなさい。書籍の全てを読むことができない場合が多いことも確認しなさい。書籍の全てを読むことができる場合があることを確認し、その雑誌のタイトルと発行年を示しなさい。

¹⁴<http://www.books.or.jp/> のトップページより引用。

