

Title	Googleになる?Googleにする?
Author(s)	阿部, 潤也
Journal	大学図書館研究, (79): 62-67
URL	http://hdl.handle.net/10130/168
Right	

Googleになる？ Googleにする？

阿部 潤也

抄録： 学術情報検索においてOPACが利用されることは少ない。OPACとサーチエンジンは同じウェブ上でサービスを提供しているが、なぜOPACが利用されないのか、なぜサーチエンジンがこれほどまで利用されるのかを考察した。OPACにはユーザに対する、検索結果の積極的な提案がないことが利用されない原因であるとし、代表的なサーチエンジンであるGoogleをモデルとしたOPACの理想形を提案した。また、Googleの図書サービスへの展開にも着目し、GoogleをOPACに利用することの可能性を論じた。

キーワード： OPAC, Google, サーチエンジン

はじめに

学術情報の検索ツールとして最初に使う情報源は果たして何であろうか？ EDNER Projectの調査¹⁾によると、英国の大学生が研究プロジェクトやレポート作成に着手する際に用いる情報検索ツールの64%が各種サーチエンジン（Googleが筆頭で45%）であり、大学提供OPAC（Online Public Access Catalog）は10%であった。

情報検索と図書の所蔵検索が異なることは言うまでもないが、情報検索手段の視点において、OCLCサイトに掲載されている「Google is disintermediating the library（Googleが図書館を無用化しつつある）」²⁾という言葉は決して大袈裟なことではなく、インターネット上にGoogleやAmazonが存在している今、OPACを検索する人は近い将来いなくなるかもしれない。

そこで、本稿ではサーチエンジンとOPACの発展の過程から、なぜOPACが利用されないのかを考察し、Googleの技術革新から現在のOPACに求められる機能について述べることにする。

サーチエンジンの発展

増え続けるウェブページを効率よく検索するための手段として提供されているサーチエンジンであるが、今までのところ第一から第三世代までに区分できる（表1）³⁾。

第一世代はYahoo!に代表されるディレクトリ型である。ディレクトリ型の場合、手作業によるウェブサイトの収集や分類を基本としているので、選りすぐりのページを収集できる反面、網羅的な収集、鮮度の維持が難しい。人海戦術型とも言える。

第二世代の特徴はクローラ（Crawler）の利用である。クローラと呼ばれる、ウェブページに張られたリンクをたどってすべてのページを網羅的に自動的に収集するロボットの利用により、第一世代の弱

点であった大規模収集が可能となった。また、大量のウェブページを対象とした全文検索も、高速なレスポンスを提供するために並列全文検索が開発された。検索対象のページ規模と、ユーザからのアクセス頻度に対するスケーラビリティを確保するための技術を実現している。この二つの技術を利用して提供されたサーチエンジンの代表的なものはLycos, gooである。

また、第二世代の構成を踏襲しつつリンク解析に基づく検索結果ランキング方式を導入したのが、第三世代の技術である。メジャーなスコアリング方法はPageRankと呼ばれ、多数引用されるページは信頼でき、信頼できるページに引用されるページも信頼できる、という考えに基づいている。これがまさにGoogleで採用されている方式である。

表1 これまでの代表的なサーチエンジン

第一世代	Yahoo!	ディレクトリ型
第二世代	Lycos, goo	クローラ, 全文検索
第三世代	Google	第二世代 + PageRank

検索手段としてのOPACの進化

一方「検索」という視点から図書館を見ると歴史はかなり長い。米国議会図書館ではジェファーソン元大統領の蔵書（約6,500冊）の分類目録を冊子目録として1815年に発行している⁴⁾。また、米国で最初のカード目録は1854年にボストン公共図書館で使われており、そのカード目録をコンピュータで処理できる形態に変換する方法についての研究が、米国議会図書館を中心にMARC（MACHINE READABLE CATALOG）プロジェクトとして1966年に発足している⁵⁾。インターネット時代においてはTELNET等のOPACからWebOPACと進化し現在に至っている。

OPACの現状

冊子目録からカード目録，オンライン目録の最新形のWebOPACへの進化は図書館に多大な影響をもたらした。しかしながら現在のWebOPACがユーザに真の満足を与えているか，と言えそうではないであろう。OPACが利用されないのは，サーチエンジンで言えば，第一世代に相当する機能で停滞しているからではないだろうか？

また，それまでローカルの環境でしか検索できなかった図書館所蔵の目録をインターネット上に公開したことは画期的なことであった。しかし，それはこれまでの長い労力の積み重ねを公開したに過ぎず，システムそのものについては何ひとつ目新しいことは無かった。極論すれば，冊子目録からWebOPACへの進化は所蔵情報を公開するためのインターフェースが変化しただけで，そこで公開されるデータの内容は今も相変わらず伝統的な目録法に則った書誌事項を主体としている。

では，OPACはGoogle時代の今，どうすればユーザの満足を得ることが出来るのであろうか？ OPACに，サーチエンジンにおける第三世代の機能を実装することは出来ないものであろうか？ 書誌事項以上の何を公開すれば良いものであろうか？

Googleになる？

東京歯科大学図書館におけるOPAC（富士通：iLiswave）の検索結果の表示順は「タイトル」「著者名」「出版者」「出版年」である。その表示順から導かれる検索結果はあまりに機械的でユーザへの積

極的な提案が無い。ピンポイントで資料を探す場合は，現状の機能で問題はないと思われるが，ある主題や特定のキーワードに関連する図書を求めるユーザの満足を得させる検索結果を提供することは不可能であろう。

国外のOPACはどうであろうか？ 米国医学図書館のOPACであるNLM Locator Plus⁶⁾のOPAC（Endeavor：Voyager）では検索結果の表示順のデフォルトが関連順の「Relevance」であり，関連度をアイコンの数で表現している（図1）。積極的な検索結果の提供という点で，ある程度評価できる。

では，Google（サーチエンジンの第三世代）的に検索結果を積極的に提供するためには，OPACの表示順にどのようなアルゴリズムを実装すれば，ユーザの満足を得ることが出来るのであろうかを提案したい。Googleがウェブページに索引を付与して，そのページの重要度を基準に結果を表示するのと同様に，図書資料の重要度を基準に検索結果を提供するのである。

表2 図書資料評価指標

- | | |
|---|--------------------|
| 1 | 貸出回数 |
| 2 | 読者の評価 |
| 3 | 図書館員の評価 |
| 4 | Google的なPageRank評価 |

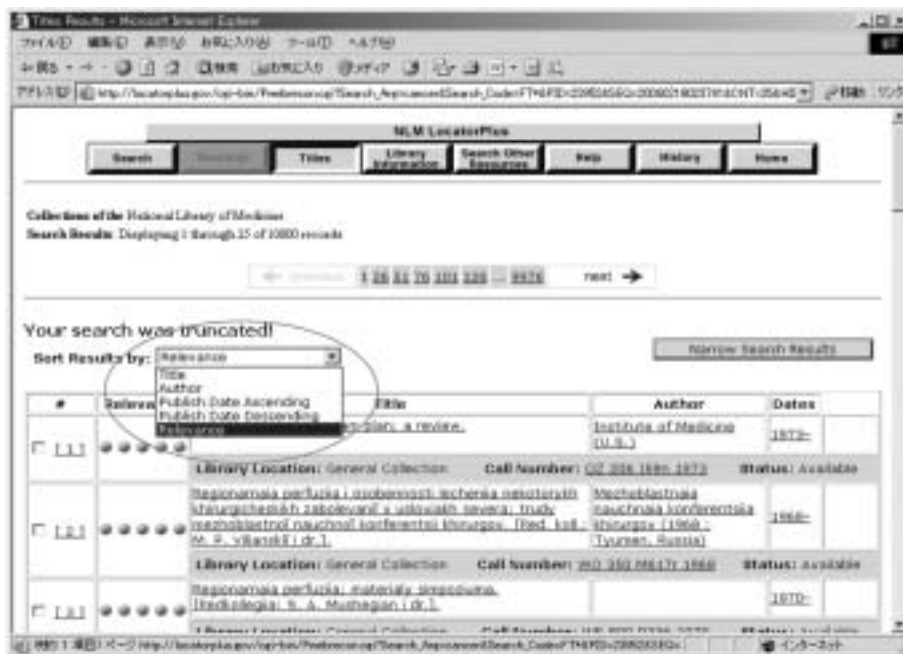


図1 NLM Locator Plus

検索語に対する、タイトルや著者名、出版者名の一致度の他に図書としての新鮮度を示す最近の出版年のものを上位に表示させる。これらは現在のOPACにおいて実現している静的なデータによる検索結果の提供である。そこに動的なデータとして、表2に挙げた項目を図書の重要度としてデータ付与することで、より積極的な検索結果の提供が可能となるであろう。

1の貸出回数は書店で言えば売りに相当する評価指標であろう。現在のOPACにおいても貸出履歴から頻繁に借りられている図書を集計するベストリーダーの機能を実装している。それらのデータを検索結果に反映させるのであるから、十分に実現可能であろう。2の読者の評価は、実際に図書を借りたユーザーに返却後に評価してもらい、それをデータとして蓄積していくのである。その図書に対するコメントなどがあれば、より有効なデータベースとして育っていくであろう。3の図書館員の評価は選書以上にシビアで、専門力が必要な作業になるはずである。その分だけ、読者の評価よりも重要度のポイントは高くすべきである。4のPageRank評価を用いるには、書誌が静的なウェブページである必要があると同時に、特定の書誌が他のウェブページにどれだけリンク(引用)されているかを解析する必要がある。現時点では実現は難しいかもしれないが、書誌の閲覧回数で代用できるかもしれない。

以上にあげたようなアルゴリズムを実装することで、積極的な検索結果を提案をするOPACにすることが可能であろう。つまりは、『OPACがGoogleになる』のである(図2)。

表3 プロフィールによる判断基準

1	身分
2	専門分野
3	シラバス
4	検索動向

また、それぞれのプロフィールを検索結果に反映させてはどうだろうか(表3)。1の身分では提供する資料の専門度を教員か学生か、あるいは学年によって判断する。2の専門分野の情報は、ユーザーのニーズの基本的な情報源となるであろう。3のシラバス情報はそのユーザーが学生ならば、これから受ける授業に関する教科書等の関連資料を提供することになる。これら1~3に共通して言えるが、同じ身分、専門分野、履修科目の他ユーザーの検索や貸出の動向も蓄積することである。そうすることで同じよ

うなニーズを持つ他のユーザーの情報をも検索結果に反映させることが可能となる。また、4の検索動向は、Googleでは「Googleパーソナライズド検索」⁷⁾として同様のサービスを始めている。例えば、「Matrix」を検索した場合、過去の検索履歴からアカデミックな情報を求めているユーザーなのか、あるいは娯楽的な情報を求めているユーザーなのかを判断して、最適な検索結果を提供するのである。

これらのプロフィールは1, 2をベースに3は時期毎に変化し、4では検索を行う度にOPACが学習し、ニーズに合わせた検索結果の提供を行うことが可能となる。以上にあげた項目にどれだけの重要度を付与するかは、検索結果に直接影響が出るので慎重に考慮する必要があるであろう。

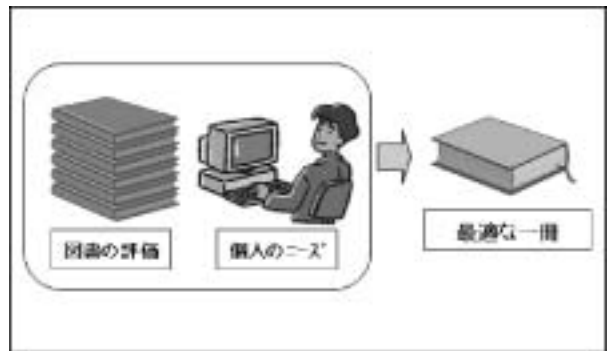


図2 GoogleになったOPAC

近年「My Library」というポータルサービスが普及しつつあるが、提供される内容は貸出情報や予約が主であり、「My Library」と呼べるほどパーソナライズされたページとは言えない。しかしながら、こうした既存のインターフェースを利用してそれぞれのプロフィールを登録することにより、そのユーザーに特化した検索結果を積極的に提供することが可能になるであろう。

Googleにする？

2004年10月にGoogleから、図書全文検索サービス「Google Print(現在はGoogle Book Search)」⁸⁾が発表された(図3)。Google Book Searchは著作権の切れている図書の全文と、著作権の切れていない図書の前後数ページ(目次や索引)を検索・閲覧対象としている。また、出版社の許可した図書だけではなくハーバード大学、スタンフォード大学、ミシガン大学、オックスフォード大学、ニューヨーク公共図書館が参加するプロジェクト(Google Books Library Project)⁹⁾では、これらの図書館蔵書の一部をGoogle Book Searchで電子化して公開してい

る。

従来の書誌情報に加え目次情報等を付加した図書検索システムは「ブックコンテンツ・データベース」^{10) 11)}として東京大学が提供している。その有用性は従来のOPACに比べて網羅的な主題検索が可能であることが既に証明されている。目次情報等に加え、全文情報をも検索、閲覧できるGoogle Book Searchとタイトルや著者名情報だけのOPACのどちらをユーザが選択するかは言うまでもない。とはいえ、図書資料の全文電子化は法的な課題もあり、なかなか普及していないのが現状である。

Google Book Searchの所蔵データについては「Open WorldCat」¹²⁾にリンクしているため、Worldcat参加館はシームレスに自館の所蔵確認が可能である。あるいは、ユーザは自分の利用する図書館に所蔵が無いことが分かれば、すぐ隣にリンクしているAmazon等のオンライン書店にそのまま注文をするかもしれない。Googleから公式な発表はされていないが、Amazonに対抗するオンライン書店を検討中¹³⁾という話もある。

また、Google Book Searchを自館OPACの代替検索インターフェースとして利用するには、資料から所蔵データへのリンクが張られていれば良いわけだが、学术论文検索用のサーチエンジンであるGoogle Scholarでは既に「Library Links」¹⁴⁾として実現している（図4）。

Library Linksは図書館側で用意しているリンク

リゾルバとリンクすることで、図書館で利用可能な電子リソースへリンクすることが可能である。Google Book SearchとGoogle Scholar、リンクリゾルバとOPACが連携することでユーザは図書、雑誌の情報を本文情報や図書館の所蔵状況と共にシームレスに入手することが出来るようになるであろう（図5）。つまり、OPACは自館の所蔵や貸出情報等の動的な所蔵データを提供するだけで、書誌検索はGoogleに任せるのである。この意味で、『OPACをGoogleにする』のである。Google Book Searchを検索インターフェースとして、積極的に利用したOPACも近い将来現れるかもしれない。

中国国家図書館ではすでにGoogleとの提携により「電子図書館」サービスを展開しており¹⁵⁾、詳細については公開されていないが、国家的な取り組みであるこのプロジェクトには今後注目すべきであろう。

おわりに

サービスは時代に合わせて常に進歩していく必要がある。電子図書館構想も図書館にとっては、その先進的なサービスの一つであったのであろう。しかし、今や実際に真の電子図書館を実現しそうなのはGoogleやMicrosoft¹⁶⁾である。とは言え、OAI-PMHの実装やWebcat Plusによる連想検索等OPACも着実に進化を続けている。第三世代であるサーチエンジンの次なるキーワードは「目的特化」「状況適応」¹⁷⁾



図3 Google Book Search



図4 Google Scholar Library Links

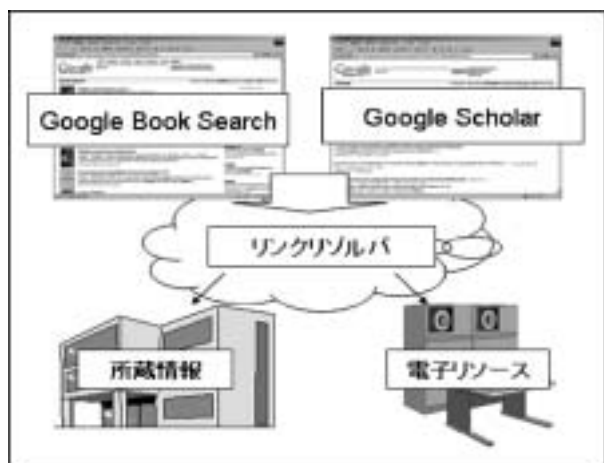


図5 Googleを検索の入り口に

である。冒頭で述べた「Google is disintermediating the library」を現実のものにしないために、次世代OPACは、その機能を実装すべきであろう。できないのであれば、これからのGoogleの展開に期待すれば良いだけなのだが。

注記・引用文献

- 1) EDNER Project. "How students search : information seeking and electronic resource user. Issue Paper 8, 2004". available from <http://www.cerlim.ac.uk/edner/ip/ip08.rtf>, (accessed 2007-1-6).
- 2) OCLC. "Google is disintermediating the library. OCLC 2003 Environmental Scan Introduction." available from <http://www.oclc.org/reports/

escan/introduction/default.htm>, (accessed 2007-1-6).

- 3) 福島俊一. 検索エンジンの仕組みと技術の発展. 情報の科学と技術. Vol.54, 2004, p.66-71.
- 4) 丸山昭二郎編. 目録法の歴史, 現状, 今後. 洋書目録法入門 マニュアル編. 東京, 日本図書館協会, 1988, p.1-16. (4-8204-8802-3)
- 5) 川崎良孝, 薬師院はるみ. テクノロジーと図書館. 図書館ハンドブック. 6版. 東京, 日本図書館協会, 2005, p.25-31. (4-8204-0503-9)
- 6) National Library of Medicine. "Locator Plus." available from <http://locatorplus.gov/>, (accessed 2007-1-6).
- 7) Google. "Google パーソナライズド検索." available from <http://www.google.co.jp/psearch>, (accessed 2007-1-6).
- 8) Google. "Google Book Search." available from <http://books.google.com/>, (accessed 2007-1-6).
- 9) Google. "Google Books Library Project." available from <http://books.google.com/intl/en/google-books/library.html>, (accessed 2007-1-6).
- 10) 茂出木理子, 杉田いづみ, 前田朗. 東京大学ブックコンテンツ・データベースサービスの紹介. 薬学図書館. Vol.44, 1999, p342-344.
- 11) 東京大学. "ブックコンテンツ・データベース." available from <http://contents.lib.u-tokyo.ac.jp/contents/top.html>, (accessed 2007-1-6).
- 12) OCLC. "Open WorldCat program." available from

- <<http://www.oclc.org/worldcat/open/default.htm>>, (accessed 2007-1-6).
- 13) Nikkei BPnet. “Google, Amazon 対抗のオンライン書店か.” available from <<http://www.nikkeibp.co.jp/archives/419/419647.html>>, (accessed 2007-1-6).
- 14) Google. “Library Links.” available from <<http://scholar.google.com/scholar/librarylinks.html>>, (accessed 2007-1-6).
- 15) 人民網日本語版. “国家図書館が「電子図書館」開始-グーグルとの提携で.” available from <http://j.peopledaily.com.cn/2005/12/21/jp20051221_56112.html>, (accessed 2007-1-6).
- 16) ITmedia ニュース. “Microsoft, 大英図書館の蔵書を

デジタル化.” available from <<http://www.itmedia.co.jp/news/articles/0511/05/news003.html>>, (accessed 2007-1-6).

17) 前掲3

なお、本稿は『デジタル・ライブラリアン研究会』第5期大学図書館コースの修了レポートとして執筆したものに加筆修正をほどこしたものである。また、一橋大・豊田裕昭氏には原稿を読んでもらい、貴重なコメントをいただいた。この場を借りて感謝申し上げる。

<2006.11.9 受理 あべ じゅんや 東京歯科大学図書館司書>

ABE, Junya

Become Google? Make it Google?

Abstract: People rarely use an OPAC when searching for scholarly information. Search engines and OPACs are both accessed over the Internet, but OPACs are rarely used and search engines are used frequently. The author believes that it is because the OPAC does not offer the user any suggestions based on search results, and proposes an OPAC model based on a search engine like Google. Furthermore, he posits that the Google book service has developed to the point that it is possible to use Google as an OPAC.

Keywords: OPAC / Google / search engines