

たった一つの検索窓から学術情報を自由自在に活用 Web Scale Discovery Service ～ Summon

松下 茂
株式会社サンメディア
matsushita@sunmedia.co.jp

概要：図書館が所蔵する資料や契約している電子リソースを図書館利用者が教育や研究のために有効に活用するためのディスカバリーサービスと呼ばれる検索システムが注目されている。中でも、Web-Scale Discovery Service と呼ばれる「Summon」は、検索対象となるリソースをあらかじめインデックスし、高速で確実な検索結果を提供するサービスとして急速に普及しつつある。対象となるリソースを提供する出版者の協力も増えてきており、ディスカバリーサービスのデファクトの位置を占めつつある。

キーワード：ディスカバリーサービス、ウェブスケールディスカバリー、OPAC、検索、Summon

1. はじめに

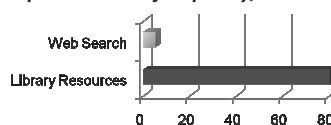
大学図書館や研究図書館では、膨大な学術資料を利用者が効率的に利用するためのサービスやシステムをいかに構築するかが大きな課題となっている。

この10年間で学術資料の多くは電子的に提供されるようになり、電子ジャーナルとして提供されている学術雑誌の数は3万タイトルを超えている。また、電子書籍も増大の一途をたどっている。一方では、物理的な蔵書も過去からの蓄積で膨大になっている。このような中で利用者が、探し求める学術情報を的確に見つけることは簡単ではなくなっている。

他方、インターネット上では無料で提供されている学術関連情報も多く存在しているために、利用者の多くは図書館が所蔵する資料（ローカルリソース）や図書館が契約する有料コンテンツ（電子リソース）を探すよりも、GoogleやYAHOOなどの検索エンジンを使って情報を探す傾向が強くなっている。

過去の米国での調査によれば、有用な学術情報の所在は図書館にあると認識している利用者が多いにもかかわらず、図書館利用者の圧倒的多数が最初に使うツールは、Googleであった。（図1、2）

The superior source for quality, credible content



Preferred for academic research and course assignments

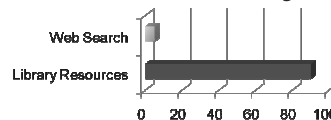
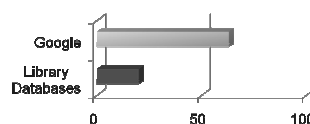


図1 一般の Web サーチと比較した
図書館資料の有用性

The easiest place to start the research process



出所：Information Online 2009.

BRINGING LIBRARY RESOURCES INTO USERS' CONTEXT: A MULTI-PART RESEARCH STUDY EXAMINING ALTERNATIVES FOR INTEGRATING LIBRARY ACCESS POINTS INTO USER ENVIRONMENTS

John Law, Serials Solutions, Ann Arbor, Michigan, USA

URL: http://www.information-online.com.au/sb_clients/og/bin/og_programme_sandbox_to_search_box_B1.cfm?vm_key=76E5F84C-1422-0982-EBEFF3E0A9FA463F

図2 学生が最初に使う検索ツール

このような状況を背景に、図書館が提供するリソースを有効に活用できるようなサービスやシステムが開発されてきた。それらは、複数データベースを串刺しに検索して結果を表示する統合検索システムや図書館のOPACを探しやすく見やすく表示できる次世代OPACなどのディスカバリーサービスとして提供されている。

前者の例としては360Searchがあり^[1]、様々なデータベースや電子ジャーナルを一度に検索することができる。日本語のデータベースでは、CiNiiや医中誌Web、JDreamII、日

経 BP 記事検索サービスなど、多くを検索対象としている。また、後者の例としては AquaBrowser などがある^[2]。AquaBrowser は、従来の OPAC にはなかった所蔵図書の書影（表紙写真のサムネイル画像）や目次、所蔵資料の種類別アイコンの表示（DVD や CD、書籍などの違いが一目で分かるアイコン）、貸し出し状況のリアルタイム表示などを実現することで、視認性に優れたインタフェースを実現した。

このようなディスクバリーサービスの延長線上にあるのが、ウェブスケールディスクバリー（Web-Scale Discovery）サービスとしての Summon（サモン）である^[3]。

以下、Summon の特徴を述べるものとする。

2. Summon とは

2.1 Summon とは

Summon は、Serials Solutions 社^[4] が開発した図書館向けの検索システムのことで、Google や Yahoo などの検索エンジンと同様に、データをあらかじめハーベストしインデックスすることで検索の速度や精度を上げる手法を使っている。

「summon」は英語で、なにかの目的のために人を召還、招集するという意味を持っているが、Summon は、まさに図書館の所蔵物や図書館が契約しているすべての電子情報資源を集めて検索し一つの検索結果として表示する。

Summon がクロールしてハーベストするデータは、契約する機関の OPAC データや図書館システム（ILS）のデータと出版者やデータベースベンダーのコンテンツである。

現在、Summon に協力している学術出版者や学術データベースベンダーは 6800 社余りとなっており、コンテンツは 94000 誌以上の雑誌やデータベースを含んでいる。またインターフェースは、15ヶ国語に対応している。

2009 年 7 月にリリースされて以来、全世界で 160 機関を超えて利用されており、国内でも 2 機関が利用している。

2.2 次世代 OPAC の限界

Summon が登場する前に注目を集めていた次世代型 OPAC は、既存の図書館システムや OPAC のデータを収集して前述のような視認性の良い表示を行うが、図書館が契約

する外部の電子リソースと一緒に検索することができない。したがって、リソース全体の統合的な検索のためには、次世代型 OPAC 以外に 360Search などに代表される統合検索機能を別途に契約して利用することになる。

ところが統合検索対象のデータベースの数が多い場合は、パフォーマンスは各データベースのシステムに依存するため、検索効率が落ちる可能性もあった。そこで、外部リソースも含めた情報資源を一気に検索して回答を示すツールが必要となり、Summon の登場となった。

3. Summon の機能

3.1 メタデータのハーベスティング

前章で記載したように Summon の最大の特徴の一つは、図書館が契約している電子リソースのメタデータを最初から搭載していることであるが、そのデータは個別の図書館がデータを提出したものではなく、Summon に協力する出版社やデータベース製作者のメタデータをクロールリングによってハーベストすることによって実現している。クロールは、24 時間に 1 回のタイミングで行われており、常に最新のコンテンツや情報を検索することができる。

ハーベストしたデータの中に、同じ書誌のデータが含まれている場合は、ノーマライズという一定の書式にもとづくデータとして 1 本化される。たとえば出版者サイトの電子ジャーナルからハーベストしたデータと、アグリゲーターと呼ばれるコンテンツプロバイダーからハーベストしたデータに、同一の書誌データがあった場合などである。検索した結果のタイトル部分をクリックすれば、リンクリゾルバー経由で機関が契約しているそれぞれの全文（フルテキスト）にアクセスできる。

しかし、Web of Science（WOS）などのように個々の書誌レコードに引用回数など有用な分析データが付与されている場合は、その分析データそのものに価値があるので、ノーマライズの対象とならず、その検索結果から WOS に遷移するようになっている。

3.2 Web-API の実装

Summon は、外部からの検索連携のために Web API を実装し、Z39.50 のほか SRU、

SRW, Open Search にも対応している。そのため利用する機関によって様々なインターフェースの構築が可能である。

検索の主な基本機能は下記のようになっている。

3.3 検索インターフェース

● シンプルなサーチボックス：

Google などの検索エンジンと同様に複雑な操作を必要とせず、たった一つの検索窓から探したい情報を簡単に深く探することができる。

もちろん詳細検索モードへの切り替えも自由である。(図3、4)

入力の際には、スペルミスを防ぐために“Did you mean ?” 機能で正しい候補語を示す機能がある。



図3 シンプルなサーチボックス

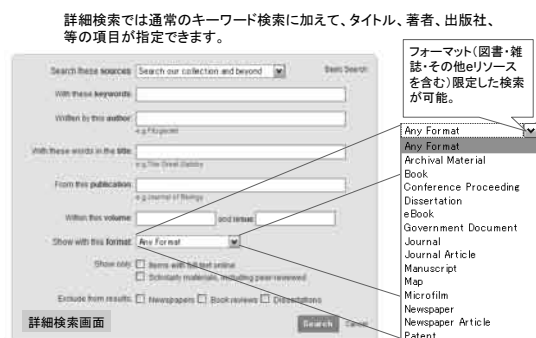


図4 詳細検索

● 優れた関連順検索結果：

関連性の高いものから順に検索結果を表示する。また、検索結果の詳細情報もワンクリックで確認可能。(図5、6)



図5 優れた関連順検索結果



図6 検索結果の詳細情報

● 検索結果のリファイン：

検索結果は、フィルタ、ファセット、ソート機能で簡単に絞り込みが可能。

ここから、フルテキストだけに検索結果を絞り込んだり、年代を指定しての検索が可能となる。(図7～図9)

その他、検索結果から著者名をクリックすると同一著者による執筆資料を検索する機能など、細かい絞り込み検索の機能が充実している。

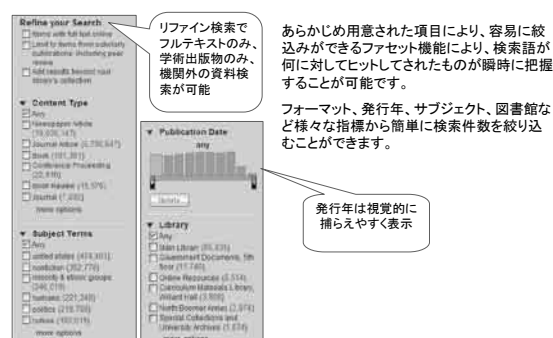


図7 リファイン検索

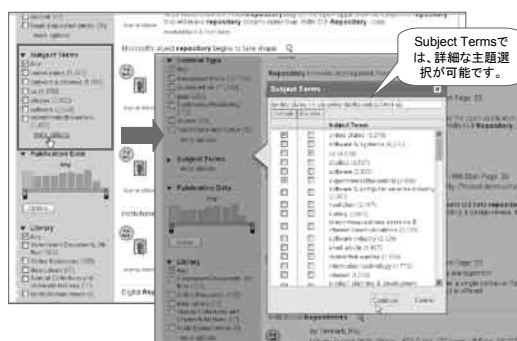


図8 Subject Terms

複数の場所(学部や図書館)にリソースがある場合は、どの場所からリソースを得るかの選択も可能です。



図9 検索結果のリファイン

- 資料の種類をアイコンで表示：
検索した結果が、どのような種類の資料なのかを一目で分かるアイコンで表示。(図10)

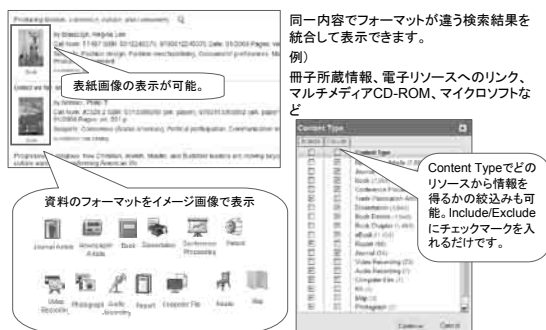


図10 検索結果の表示

- リソース管理ツールへのエクスポート：
RefWorks, EndNote, ProCite など個人リソース管理ツールへのエクスポート機能を搭載。(図11)



図11 リソース管理ツールへのエクスポート

- 図書館システムとの連携：
冊子体の貸し出し情報などを表示。(図12)

ILS連携も実現し、貸出状況をリアルタイムに表示できます。



図12 図書館システムとの連携

4. Summon の今後

Summon は、従来の統合検索システムや次世代 OPAC の先のディスカバリーサービスとして多くの大学図書館や研究図書館で採用されていこうとしている。一方、Summon の利用者が増えれば増えるほど、個々の利用者の要望や利用地域・国・言語の違いによる要望が増えてくる。また、収載コンテンツの要望も増えてくる。これらローカルデマンドを吸収し実現することが、これからの Summon の発展の鍵となる。

また、図書館システムを提供している ILS ベンダーとの協力も不可欠である。そのためにも多くの ILS ベンダーの声を聞き、Summon の仕様に反映させていくことも今後の拡大の二つめの鍵となるであろう。

参考サイト

- (1) <http://www.serialssolutions.com/360-search/>
(Accessed 2011.02.14)
- (2) <http://www.serialssolutions.com/aquabrowser/>
(Accessed 2011.02.14)
- (3) <http://www.serialssolutions.com/summon/>
(Accessed 2011.02.14)
- (4) <http://www.serialssolutions.com/> (Accessed 2011.02.14)