

文部科学省 国立情報学研究所

図書検索サイト“Webcat Plus”のサービスを開始

- 連想検索機能でベストの本探し -

文部科学省 国立情報学研究所 [NII] (東京都千代田区一ツ橋, 所長 末松 安晴) は, 平成 14 年 10 月 8 日から, 図書を効率よく探すことができる“Webcat Plus”(ウェブキャット・プラス: NII 図書情報ナビゲータ) サービスを開始する。

国立情報学研究所では以前から, 全国の大学図書館が所蔵する図書や雑誌を検索できるサービス“Webcat”を公開しており, 国内外から非常に多く利用されているが, “Webcat Plus”では, 検索機能・性能を大幅に強化し, 使い勝手を向上させている。

“Webcat Plus”の機能面での最大の特徴は, 膨大な情報の中から知りたいテーマに関連するものを即座に探し出す「連想検索機能」である。

単語からはもちろんのこと, 長い文章を入力してもトピックスを自動的に解析し, 関連する図書を探してくれる。さらに, 検索結果の図書一覧から関心のあるものを選んで検索を重ねることで, より関連度の高い図書を見つけることができる。

また, 検索結果を基に, 関連するキーワードが自動抽出されるので, これらを用いてもう一度検索できる。このようにして, 自分が考えつかなかった言葉や概念を発見し, イマジネーションを広げていくことができる。

この連想検索機能は, 国立情報学研究所と(株)日立製作所中央研究所が共同で開発した, 大規模文書を高精度かつ高速に処理するソフトウェア《汎用連想計算エンジン“GETA”(ゲタ)》を使用して実現している。

“Webcat Plus”は, 毎週追加される最新の図書から, 明治期以前に発行された古いものや既に絶版となって図書館にしかないものまで同時に検索できるという点において, 国内随一の図書情報検索サイトである(平成 14 年 10 月現在, 日本の図書約 235 万件)。

これらについて, 目次や内容の情報を見ることができる(昭和 61 年以降発行分)だけでなく, 図書を所蔵している大学図書館の一覧や, 図書館の利用条件を知ることでもできる。

また, 毎年 6 万点を超えて新しく発行される図書について, 従来はその全貌をつかむ術がなかったが, “Webcat Plus”により, 一般の人が自分の興味に最も適した図書を探し出すこともできるようになる。

“Webcat Plus”は, サービス開始後も常に内容や機能の充実を図っていく予定である。

当面は日本の図書のみを収録するが, 海外の図書情報へも拡大する予定である。また, さらに便利な機能を追加し, 図書を探すための日本の代表的なサイトとして発展させていくことを目指している。

なお, 国立情報学研究所では, インターネット上の各種学術情報を連携させ, 研究に必要な情報を総合的に利用できる環境“GeNii”(ジーニイ: NII 学術コンテンツ・ポータル)の構築を進めている。“Webcat Plus”は, その中で図書情報提供の役割を担う。

“Webcat Plus”の URL: <http://webcatplus.nii.ac.jp/>

用語解説

(1) Webcat (ウェブキャット) <http://webcat.nii.ac.jp/>

国立情報学研究所が提供している目録作成システム (NACSIS-CAT) を通じて、全国の大学等の図書館 (約 1,000 の大学・研究機関等) が共同で図書・雑誌情報の総合目録データベースを作成している。

この総合目録データベースを基にして、インターネット上でだれでも無料で検索できるようにしたサービスが“Webcat”である。

全国の大学図書館等が所蔵している図書や雑誌について、タイトル、著者/編者名、出版社、出版年、標準番号 (ISBN/ISSN) や、所蔵図書館の情報を知ることができる。

平成 10 年にサービスを開始して以来、国内外から非常に多く利用されている人気サイトである (1 日平均 約 12 万ページビュー、検索回数約 5 万回)。

なお、当面は“Webcat”と“Webcat Plus”とを並行してサービスするが、“Webcat Plus”の内容や機能が充実した時点で一本化する予定である。

(2) 連想検索機能

現在広く使われているキーワード検索では、探したいテーマを表す言葉をいくつか指定して検索を行うが、抽象的なテーマや多義性のある言葉を使った場合、検索漏れが生じる、無関連な情報が多く紛れ込むなどの問題がある。あらかじめ知識のない分野について、ユーザが適切なキーワードを漏れなく指定することは大変難しい。

“Webcat Plus”の「連想検索機能」は、このようなキーワード検索の限界を補完して、ユーザが求める情報を素早く探索できるよう手助けする。

[文章からの検索] 新聞記事のような文章を丸ごと入力して検索できるため、ユーザは検索キーワードを考える必要がない。システムは入力文中の特徴的な言葉を自動抽出して検索に用いるため、検索漏れが少なく、関連性の高い検索結果が得られる。

[関連ワードからの検索] 検索結果を要約する関連ワード一覧が示されるので、その中から関心に合致する言葉を選んで再検索することで、よりシャープな結果が得られる。

[図書からの検索] 検索結果中の図書を複数指定して、それらと関連する図書を再検索できる。検索結果にいくつかの異なるトピックが含まれている場合、特に関心のある図書を指定して再検索することにより、関連図書を効率的に集めることができる。

なお、“Webcat Plus”では、タイトル・著者名・出版社名・出版年・ISBN からのキーワード検索が可能な「一致検索機能」も提供している。調べたい図書が明確で書誌情報がある程度わかっている場合に便利である。連想検索機能とは容易に切替可能であり、一致検索結果の中から図書を選択して連想検索を行うことも可能である。

(3) GETA (ゲタ) <http://geta.ex.nii.ac.jp/>

汎用連想計算エンジン“GETA”は、1 千万件規模の大規模文書 DB を対象に、文書間や単語間の関連性を高精度かつ高速に計算するソフトウェアである。文書に現れる単語の頻度を索引化して、PC クラスタで並列分散処理することにより高速化を実現している。

“GETA”は、情報処理振興事業協会 (IPA) が実施した、平成 13 年度「独創的情報技術育成事業」の一環として、国立情報学研究所と(株)日立製作所中央研究所が共同で開発した。オープンソース形態でプログラム及びドキュメントを無償公開している。

Webcat Plusで何ができるか?

- ◆より関連度の高い本に絞り込める
- ◆いろいろな分野から関連する本を探し出せる
- ◆思いがけない本を発見できる



本を探したい

あるテーマについて調べたい

連想検索



知りたいテーマのキーワードや文章で探す

一致検索



書名, 著者名, 出版年, ISBNなどで探す

書名か著者名がわかっている



どんな本?

図書詳細情報

- ・目次
- ・内容
- ・書名, 著者名, 出版社, 出版年, ページ, サイズ,

所蔵図書館一覧

図書館情報

- ・住所, URL
- ・利用条件

どこの図書館が持っている?



GeNiiとは?



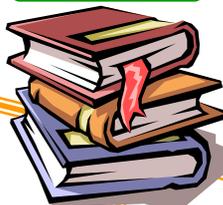
NII 学術コンテンツ・ポータル

インターネット上に点在する各種の学術情報資源を連携させ、研究に必要な情報を総合的に利用できる環境を構築しています。

論文情報



図書情報



研究者情報



電子ジャーナル



研究機関情報



研究課題



研究成果



特許情報

