

連想検索のためのコンテキスト情報を提供する web サービスの実証研究

研究代表者 青木隆平 有限会社セカンドブレーン 代表取締役
共同研究者 高野明彦 国立情報学研究所 教授

Abstract

世界的にも有名な本の街である神田神保町を舞台に、古書および古書店を対象に連想検索を可能にした BOOKTOWN じんぼう (<http://jimbou.info/>) は、2005年10月の開設以降、着実に古書

データが充実してきている。

本研究では、その BOOKTOWN じんぼうを維持・運営しながら、その貴重な古書データをさまざまな形態で連想検索と結び付けていく。

研究成果概要

■本研究のねらい

本研究では、BOOKTOWN じんぼうのような、比較的小規模かつ内容が偏ったデータを検索対象

とする環境においても、GETA による効果的な連想検索が実現できることを検証する。

表1 専門分野別参加店数

専門分野	参加店数
文学	8
古典籍	8
歴史	7
思想・宗教	4
外国書	2
社会科学	4
自然科学	4
美術・版画	4
趣味・芸術	10
サブカルチャー	2
古書全般	7
合計	60

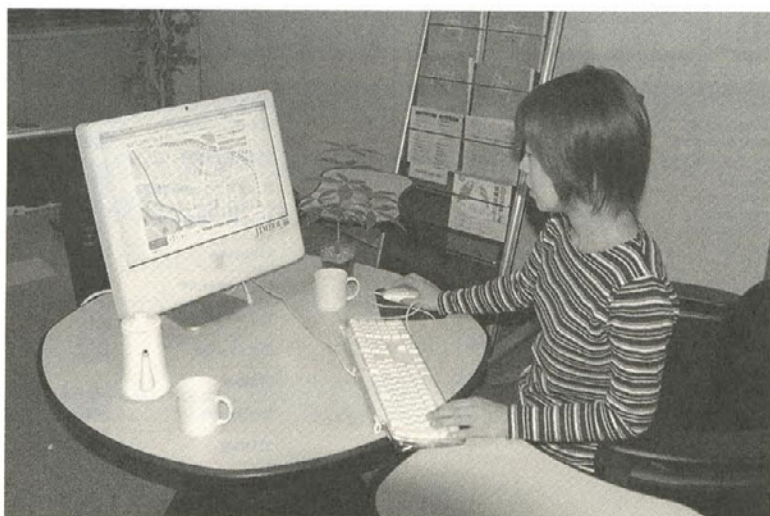


図1：じんぼう Kiosk 端末

■今年度の活動実績

今年度は、研究の土台となる BOOKTOWN じんぼうの維持・拡大に力を入れた。その結果、2006年2月現在では参加古書店が45店（参加途中古書店も含む）だったが、その後の活動により2007年2月現在では60店（同）となった。

また、BOOKTOWN じんぼうと神田神保町地域との連携を深めるため、NPO 連想出版にて「じんぼう Kiosk 端末プロジェクト」が立ち上がった。

じんぼう Kiosk 端末プロジェクトとは、神保町に実際に足を運ばれた方々が BOOKTOWN じんぼうを神保町めぐりの参考にしていただけるように、神保町の書店や喫茶店に BOOKTOWN じんぼうが気軽に閲覧できる Kiosk 端末の設置を進めるもので、今期は2台（三省堂書店、Folio）が設置に至った。

■BOOKTOWN じんぼうと想－IMAGINE Book Search

2006年7月28日に運営が開始された、想－IMAGINE Book Search (<http://imagine.bookmap.info/>) を利用することにより、BOOKTOWN じんぼうの古書データへの連想検索は、他の連想検索データベースとの相乗効果を得て、より膨らみを増したといえる。

想－IMAGINE Book Search に登録されているデータベースは、BOOKTOWN じんぼうも含めて、2007年2月現在で全8件。

想－IMAGINE Book Search では、Webcat Plus や Wikipedia などの大規模かつ、ある程度質の確保されたデータベースと BOOKTOWN じんぼうとを連動させて連想検索をすることができるので、BOOKTOWN じんぼう単体では得られない効果的な検索結果が期待できる。

以下、具体的な事例で説明する。（結果は2007年2月現在のもの）

例えば、「ザメンホフ」というキーワードで想－IMAGINE Book Search と BOOKTOWN じんぼうを連想検索してみる。

すると、この時点では、想－IMAGINE Book Search 内の BOOKTOWN じんぼうの検索結果と、BOOKTOWN じんぼう単体での検索結果は同じである。また検索結果数も5件と少なく、まだ「ザメンホフ」という単語そのものが含まれる古書しか検索されない。

この場合、BOOKTOWN じんぼう単体では、右側にある関連ワードを参考に連想をふくらませることができる。しかし、検索結果数5件の古書データから得られる関連ワードでは情報に乏しく、



図1：キーワード「ザメンホフ」での想-IMAGINE Book Search の連想検索結果

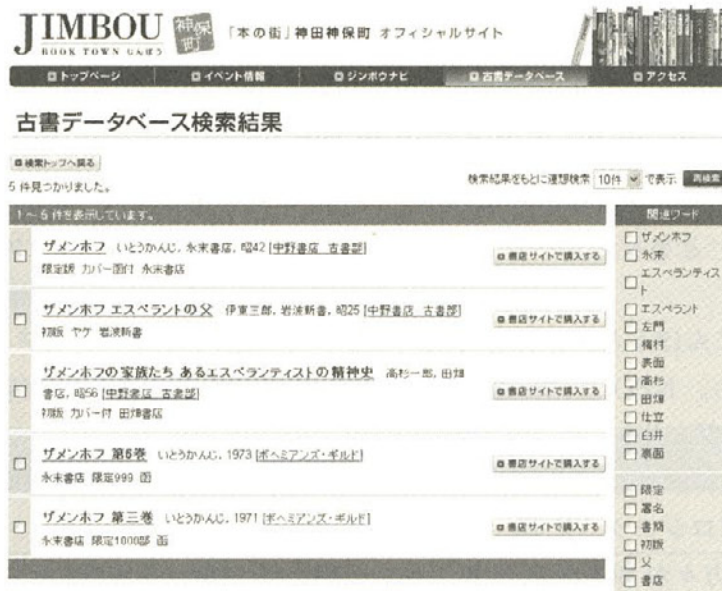


図2：キーワード「ザメンホフ」でのBOOKTOWN じんぼうの連想検索結果

効果的な連想を膨らませることは難しい。

そこで、想-IMAGINE Book Search 内の他のデータベースから得られるコンテキストを借用する。

例えば、Wikipedia データベースでは「ルドヴィゴ・ザメンホフ」という項目で、「ザメンホフ」を説明する文章がある。また、松岡正剛の千夜千冊データベースでは「ザメンホフ」という項目で、新書「ザメンホフ」の書評が得られる。

これらの項目を想-IMAGINE Book Search で選択し再度連想検索を行うことで、それらの項目にあるコンテキスト情報をもとにBOOK TOWN じんぼう内での連想が膨らんでいくことが確認できる。

例えば、Wikipedia データベースによる「ルドヴィゴ・ザメンホフ」のコンテキストからは英文の言語学関係の古書データへの連想が膨らみ、検索結果数も24,246件に増えた。また、松岡正剛の千夜千冊データベースによる「ザメンホフ」のコンテキストからは、主に日本の言語学に関する古

書データへ連想が膨らんでいき、検索結果数は10,850件となった。

このように、あるキーワードから BOOK TOWN じんぼうのような比較的小規模で内容が

偏ったデータベースを連想検索する際には、あらかじめ想-IMAGINE Book Search を用いて、他のデータベースから得られるコンテキスト情報を活用することで、より一層連想の幅が膨らむことが確認できた。