

新任教員よりの挨拶

連想検索「想-IMAGINE」とデータベース

丸川 雄 三

「想-IMAGINE」について

「想-IMAGINE」は、国立情報学研究所（NII）連想情報学研究開発センターが開発した連想検索システムである。「想-IMAGINE」の主な特徴は、複数のデータベースを同時に検索できる番組表型インタフェースと、検索結果そのものを問い合わせに利用して、それらに関連する情報を引き出すことができる連想検索機能の二つである。

連想検索とは、多種多様なデータベースの中から、あるキーワードや新聞記事に関連する項目を集めてまとめて表示する技術である。同じくNIIが開発した連想検索エンジン「GETASSOC」により実現されている。システムは種となるキーワードや記事が与えられると、それをいわば「横串」として複数のデータベースに横断的な問い合わせを行い、その結果を関連度の高い順に返し表示する。これらの計算を一瞬で行っているのが、高速演算ライブラリ「GETASSOC」である。

二〇〇六年六月、「想-IMAGINE」は、書籍を中心とした情報探索サイト「想-IMAGINE

(111) の公開によって初めて一般に披
露された。「想-IMAGINE Book Search」では、
二〇一二年九月現在で一八のデータベースを横
断して検索することができる。主な収録データ
ベースは、大学図書館の蔵書を中心に一千万冊
の和書洋書を検索できる「WebcatPlus」⁽¹¹²⁾、新書
と新書にひもづく一〇〇〇のテーマを検索でき
る「新書マップ」⁽¹¹³⁾、日本国内にある美術館や博
物館の収蔵品を写真とともに楽しむことができ
る「文化遺産オンライン」⁽¹¹⁴⁾等である。

早稲田大学演劇博物館／国立美術館との連携

筆者は二〇一二年三月までNIIでの研究開
発に従事していた。そこで「想-IMAGINE」シ
ステム制作に関わるとともに、その活用の一環
として、外部組織向けの「想-IMAGINE」を立
ち上げて公開する取り組みをこれまで実施して
きた。二〇〇八年度には早稲田大学演劇博物
館および国立美術館と連携し、一般公開サービ
スを提供した。図は二〇〇九年三月に一般公開し



た「想-IMAGINE 早稲田大学演劇博物館」の画面である。

早稲田大学演劇博物館版の「想-IMAGINE」が検索対象とするデータベースは、「浮世絵」、「現代演劇上演記録」、「近代演劇上演記録」、「AV資料」である。これらのデータベースと、「文化遺産オンライン」や「新書マップ」、その他、国立国会図書館の「貴重書画像データベース」などのデータベースが相互に検索可能である。

「想-IMAGINE」の使い方を、この早稲田演劇博物館版を例に示す。まずサイトを開き、演劇博物館の四つのデータベースの説明文が並んでいることを確認する。次に検索窓に「平賀源内」と入力して「IMAGINE」ボタンを押す。すると「近代演劇上演記録」の軸には平賀源内が脚本を執筆したという「神霊矢口渡」が上位に表示される。次にこの「神霊矢口渡」左上のボックスをチェックして再度「IMAGINE」を押す。その結果、画面に表示されている検索結果がガラリと変わり、「浮世絵」の軸にはこの作品の上演場面を描いた浮世絵が数多く表示される。これらの浮世絵の写真を選択してリンクを開けば、演劇博物館の「浮世絵データベース」で作品の詳細を見ることが出来る。さらに、左側のデータベース一覧から「文化遺産オンライン」を選び検索対象に追加すると、今度は平賀源内が開窯した「源内焼」や、源内の筆による油絵「西洋婦人図」を見つかることもできる。複数のデータベースと連想検索とを自在に組み合わせることによって、江戸の昔に活躍した平賀源内の多才ぶりを一望することができるのである。

このように「想-IMAGINE」を活用することで、単独のデータベースでは難しかった「横串」を入れるような検索（連想検索）を比較的容易に実現することができる。特に文化遺産など文化財を扱うデータベースを対象とした際に、連想検索は意味のある（つながりのわかる）形で

結果を出してくれることが多い。それは文化財にまつわるデータベースの質の高さと、文化財が本来持ち合わせている価値や来歴における多義性や多様性によるものであらうと考えている。

今後について

連想検索は、データベースの潜在的な可能性をこれまでにない形で引き出し得る技術であり考え方である。その意味で、連想検索とデータベースとを具体的につなげることができる「想-IMAGINE」は、データベースの活用に役立つシステムであると言える。一方で「想-IMAGINE」は連想検索のひとつの実現例であり、いわばコンセプトモデルの域を出ていないという側面もある。このような横串型の検索に対するニーズは、まだ世の中にそれほど認識されてはおらず、システムに対する具体的な要望などのフィードバックが十分には得られていないという問題もある。

筆者は二〇一二年四月に国際日本文化研究センターに着任した。日文研には、長年にわたって構築され維持管理されている文化資料の貴重なデータベースがある。これらのデータベースは各分野の専門家や関係者に活用され鍛えられ続けてきたものであり、そこには活用例など数多くの知見が合わせて蓄積されているものと思われる。今後は「想-IMAGINE」などこれまでの成果や経験に加え、日文研での新しい環境を活かして、文化資料にまつわるデータベースの価値や可能性をさらに引き出すための研究開発に取り組んでいきたいと考えている。

参考

- (一) 小池勇治、西岡真吾、森本武資、丸川雄三、高野明彦『分散連想計算サーバー群を統合する

連想検索システム「想-IMAGINE」』情報処理学会研究報告、NL 自然言語処理研究会報告、No. 186, pp. 31-36 二〇〇八年

- (二) 想-IMAGINE Book Search <http://imagine.bookmap.info/>
- (三) 連想検索エンジン GETAssoc <http://getassoc.cs.nii.ac.jp/>
- (四) WebcatPlus <http://webcatplus.nii.ac.jp/>
- (五) 新書マップ <http://shinshommap.info/>
- (六) 文化遺産オンライン <http://bunka.nii.ac.jp/>
- (七) 想-IMAGINE 早稲田大学演劇博物館 <http://imagine.enpaku.waseda.ac.jp/>

(国際日本文化研究センター准教授)