

■神門典子 情報社会相関研究系教授

検索質問の背後にある情報ニーズを満たす

有用な情報を収集することは、人がよりよく生きていくために不可欠です。しかし、探したい事柄が明確でない場合、キーワードを入れて検索を行う現在の方法は有効ではありません。答えが予想できないものを探するときや、どこから手をつけたらいいのかわからない場合でも、適切に対応してくれる検索システムの研究を進めています。目指す方向は、検索質問の背後にある情報ニーズを満たすことです。

探索をしながら関心を深めていける検索

そのためには、行間を読むなど、人が自然に行っているテキスト理解技術が必要です。つまり、これからの情報検索には、言葉の意味だけでなく、それがどのように使われているかなど、語用論や高次な言語処理技術が必要になってきます。また、利用者がどのように検索機能を使うのかなど、利用者の行動も融合することが求められます。

ある明確な事柄について調べるのではなく、探索をしながら関心を深めていける「探索型検索」の1つの例が、CEAX（デジタル・アーカイビングにおけるコンテンツ統合・利活用技術に関する研究）プロジェクトです。博物館が所蔵する文化財の画像や情報を利用した学習支援ソフトを開発し、それを西東京市の小学6年生の授業で実際に使っています。パソコンの画面上に現れた埴輪などの画像を簡単な操作で拡大したり、違う角度から眺めたり、時代、場所、遺物の種類などに応じて画像を並べ替えたりすることもできます。生徒たちは自分の興味に沿って探索を進め、さまざまな発見をしています。

また、「探索型検索」のもう1つの例は、MEW（探索・学習を支援する情報アクセス技術）の研究です。これは、Web上の情報資源を対話的・探索的に検索するシステムで、テーマ、時代、地理、資源のタイプなどの多面的なディレクトリーをたどっていったり、質問を入力して検索した結果がディレクトリーに沿って分類表示されたり、さらに、これから検索する方法を提案してくれるシステムです。

情報検索システムの向上を目指して

このように新たな検索システムの研究に取り組んでいますが、システムを向上させるには、その機能を評価することが必要です。実験に繰り返し使えるデータセットを作りましたが、それを多くの人に使ってもらえたらいいのではないかと思い、1997年末、情報検索システムを評価するプロジェクトNICIR（エンティサイル）を始めました。一人で研究すると一度に試せるアルゴリズムの数は限られていますが、多くの人で試せば、研究が短時間で発展することが期待されます。NTCIRはNIIの重点プロジェクトになっています。15カ国の100以上の研究チームとともに、検索技術の研究を高める研究基盤として、情報検索、要約、質問応答、情報抽出、テキストマイニング、意見分析、機械翻訳などの有効性テストに繰り返し使用できる実験用データセットを構築し、また比較評価、テスト手法の研究を行っています。1年半ごとに成果報告会を開催していますが、毎回参加チームが増えています。それだけ情報検索の研究が盛んになっている証拠だといえます。皆で切磋琢磨しながら今後も検索システムの向上に努めていきたいと考えています。

（取材・構成 村上朝子）