

研究課題別事後評価結果

1. 研究課題名： 連想に基づく情報空間との対話技術

2. 研究代表者名及び主たる研究参加者名(研究機関名・職名は研究参加期間終了時点)

研究代表者

高野 明彦 (情報・システム研究機構国立情報学研究所連想情報学研究開発センター センター長・教授)

主たる共同研究者

丹羽 芳樹 ((株)日立製作所中央研究所 主管研究員 (平成13年12月～))

3. 研究内容及び成果:

あらゆる文書情報の電子化が進み、それらの有効活用なしに、充実した知的生活は考えられなくなっている。しかし、現在の電子化された情報空間は、高信頼な知識獲得手段としてはもちろん、創造性を刺激するメディアとしても極めて不十分であり、遠く書籍に及ばない。「情報爆発時代の情報技術」などという問題の捉え方に端的に示されているように、現在日本で盛んに研究されている情報技術は、扱う情報の量的な増加にいかに対処するかという観点ばかりが重視され、人が電子情報を活用する場面での質的な問題は見過ごされがちである。しかし、当然ながらこの質的な問題こそが情報技術の社会的な価値を決める。

人が自分の考えを深めたり、新しい着想を得たりするために電子情報を活用するとき、そこで扱う情報の総量や網羅性はそれほど重要ではない。むしろ、望んでいるような思考の飛躍や気づきが自分の頭の中で実際に起きるかどうかの問題なのである。これは、数時間を費やしてGoogleがオートマジカルにサジェストする「重要」なサイトを片っ端から眺めて歩くウェブサーフィンよりも、大した内容でもない1冊の新書を批判的に読む読書の方が自分の考えや発想を鍛えるのに役立つことから明らかである。

読書の有効性は、情報を理解しやすい構成・順序で摂取することにより、その情報についての系統的な文脈が頭の中に形成されるためだと考えられる。少なくとも良書はそうのように書かれている。その文脈の移り変わりに沿って、著者の示すアイデアと読者である自分の頭に浮かぶ問いかけをぶつけ合わせることで、新しい自分だけのオリジナルな発想の手がかりとなる。ここに読書の醍醐味がある。

本研究は、電子情報空間においても、このような読書に匹敵する効果的な情報体験の場を実現し、思考や議論を深めるための情報技術を確立することを目的とした。取り扱う情報の量ではなく、その情報利用環境で得られる経験の質を追及した。

研究のアプローチは「連想の情報学」という標語で要約できる。ここでは2つの「連想」を考えている。1つは利用者である人間の頭の中で起きている「連想」である。人は何らかの問いかけを外から受けると、無意識のうちにそれに対応する記憶が連想的に想起され、いくつかの文脈がウォームアップされる。1つ目はこの連想である。2つ目は、電子情報空間における「連想」を指す。統計的処理の安定する数万件以上の文書集合について、文書間や単語間の類似性を確率的計量で測ることにより、与えられた文書と類似する文書の収集や、文書集合の内容を要約する単語群の自動抽出が可能となる。これが2つ目の連想である。「連想の情報学」では、これらの2つの「連想」がどのように創造的に相互作用可能かを研究してきた。

本研究では、研究開始時には既に完成していた汎用連想計算エンジンGETA (Generic Engine for Transposable Association) を基本コンポーネントとして活用することで初めて実現可能となる高信頼な連想的情報アクセスシステムの構築を目指した。さらにそれらを実用上も意味のある本物のコンテンツに適用して、社会的に意義のある情報サービスの立ち上げに繋げることを最優先した。これにより提案技術の実用性を実証するとともに、上質の情報サービスを一種の公共財として社会へ提供することを最大の目標とした。

研究代表者はこの研究方針を、論文等の学術的成果を偏重する普通の基礎研究プロジェクトにならないようにとの自省を込めて設定した。論文では研究成果の実用性を謳っていないながら、実用性のまったくないおもちゃのようなデータへの適用実験しかクリアできない「基礎技術」や、研究プロジェクト終了と同時にデータ更新が止まってしまう「実証サイト」の構築には意味がないと考えた。本研究における連想情報学が従来の基礎研究とは一線を画する本物の実用的情報技術であることを実証することを目標とした。

5年間の研究では、この新分野に計算機科学の手法によりアプローチして、膨大な情報に基づく連想過程の計算的な基礎づけや、情報空間の内容的概観をフィードバックして人の連想を効果的に刺激する対話手法を開発した。この情報技術を実用上も意味のある文書情報に適用して、連想的対話環境を特徴とする12種類の情報サービスを実際に構築した。そのうちの8サイトは一般公開され、ユニークな情報サービスとして多くのユーザを集める人気サイトになっている。特に「Webcat Plus」、「文化遺産オンライン」、「新書マップ」、「Book Town じんぼう」、「闘病記ライブラリー」の5サイトは、それぞれの分野の情報サービスとして国内No.1の評価を得ている。

さらに、2006年7月には、これらの特徴ある情報源を自在に組み合わせて仮想的に1つの情報源として連想探索できる技術を開発して、「想-IMAGINE Book Search」を公開した。この「想・IMAGINE」は電子情報空間での新しい「読書環境」と呼ぶに相応しいシステムである。本研究プロジェクトで追及してきた実用的な情報技術としての「連想の情報学」研究の最大の成果となった。

4. 事後評価結果

4-1. 外部発表(論文、口頭発表等)、特許、研究を通じての新たな知見の取得等の研究成果の状況

「連想の情報学」という新しい研究分野の開拓に向けて連想的文書検索システム「Webcat Plus」をはじめ、「図書館情報検索サービス」、「文化遺産オンライン」、「新書マップ」など多くの連想技術を用いた情報サービスを開発して実用的なレベルで公開している。システムとしてユニークであり、実用的で非常に有効なシステムを開発して公開したことは高く評価できる。

情報学分野では理論的な新しい成果を得る方向の研究と、新しいシステム構造を提案し、その有効性(実用の立場からの評価)を実証する方向の研究があるが、本研究は後者に属し、実用という面に重点を置いている。「生活情報技術」という観点からすれば、研究成果が広く社会に受け入れられて実用に供されることに意味があり、そのために多くの実用システムを公開してその成否を一般社会に問うという新しい研究手法の提案ともいえる。ただ、これらのシステムの評価は背景にある技術の高さや新規性もさることながら利用者の支持をどれだけ得られるかにかかっている。このシステムは最近何カ所かで実用され、またネット上の利用者も増えてきているところから社会に受け入れられつつあると考えられるが、社会的な評価が定まるまでには今しばらく時間が必要であろう。

論文発表は国内7件、国際論文なし、口頭発表は国内25件、国際10件と極めて少ない。これは実用性に主眼を置くという研究代表者の考え方によるものではあるが、もう少し論文や学会での成果発表の努力が欲しかった。また特許等にはなじまない研究課題ではあるが、実用システムの開発というフェーズを考えると、特許出願が全くなかったことには不満が残る。

4-2. 成果の戦略目標・科学技術への貢献

ますます雑多な情報の溢れるインターネットの世界で、質の高い情報の効率的な探索と人間の思考への支援の技術の可能性を高い完成度で示したことは、今後に大きな一石を投じたものといえよう。

連想検索技術を様々な情報サービスシステムへ適用して実用化を図っており、今後更に多くのニーズへの対応、展開が期待できる。日本国内だけでなく、多言語展開を図ることにより国際的にも有用なシステムとして広く

社会に受け入れられ、利用される「生活情報技術」となることを期待したい。

4-3. その他の特記事項(受賞歴など)

特になし