

特 集

自然言語処理の高度化による 知的生産性の向上



1. 知の共創のための自然言語処理技術—情報マネジメント技術を俯瞰する—
2. 自然言語処理技術による情報マネジメントの実際
3. 自然言語処理技術の高度化はいかにして組織の競争力を生み出すのか



編集にあたって

東京大学大学院

辻井 潤一

tsujii@is.s.u-tokyo.ac.jp

産業技術総合研究所

橋田 浩一

hasida.k@aist.go.jp

ポスト工業化や知価社会の到来が論じられて久しい。生産力を越える大きな市場があって、いいモノを作ればいくらかでも売れる、という「規模の経済」は終焉しつつある。範囲の限られた市場の中で自分の商品をアピールして顧客を獲得することが求められる「範囲の経済」の時代が始まっている。見込みに基づく画一的大量生産と大量消費に代わって、受注に応じた多品種少量生産と顧客のニーズに合わせた個別サービスに基づくビジネスモデルが台頭し、物質ではなく情報、あるいは製品のライフサイクル全体にわたるキメ細かいサービスが主たる商品になろうとしている。さらに、ITの普及による情報の氾濫、国境を越えた人々の往来、専門領域への知の細分化などもまた、社会の複雑性と多様性を増す要因である。経済のグローバル化は単なる市場の拡大というよりもむしろ、市場の複雑化と競争の激化をもたらす。

このように多様性の扱いが社会のさまざまな場面での主要な課題となるゆえに、物質に代わって情報が生産財の中心を占め、モノの管理に代わって情報のマネジメントが重要性を増している。優れた情報分析能力と戦略立案、およびその効率的な実践の基盤である情報マネジメント力が、組織の競争力をほとんど決するようになった。

もちろん情報のマネジメントはもとより必要だったが、その重要性がさらに高まるとともに、マネジメントの対象とすべき情報の種類が変わりつつある。大量生産・大量消費の時代には、全般的な傾向や一般性にかかわる統計情報の処理で十分だった。それは今もなお有効ではあるが、たとえば各顧客に細やかに対応するにはそれだけでは明らかに不十分であり、もっと具体的な情報、つまり、だれがいつなぜ何をどこでどうしたとかがどうしたかといった、個別的な命題内容の扱いが要請される。

このような細かい意味内容を表現するために多くの人

が使えるほぼ唯一の手段が自然言語である。実際、自然言語は具体的な意味内容を伝達するために最も広く使われており、本誌でもコンテンツの記述方式として日本語を用いているように見受けられる。社会的な情報マネジメントの中核を担うホワイトカラーの主たる業務は人間の間のコミュニケーションのネットワークを管理運営することであり、それは一種の自然言語処理であるといってもよい。

インターネットの普及と情報検索技術の進歩は情報の利用の仕方を大きく変えたが、それもまた自然言語の表現力に帰するところが大きい。WWWに蓄積された莫大な言語情報が、文書のみならず画像や音声にも及ぶ情報検索の主要な手掛かりである。このように大量の言語データの存在が、自然言語処理技術の発展を促すとともに、そうした情報の複雑性に対処し情報を効率よく利用する技術への需要と課題を生み出している。

新しい時代の情報マネジメントには、組織内に眠る断片的な知識の有効活用やネットワーク上の膨大な情報の検索と分析、複雑化する市場や社会システムに対応した情報システムの構築、学際的研究領域での情報交流や知識連携、言語の壁を越えた情報の流通など、いくつかの課題がある。これらの課題を解決するためには、知識管理技術と自然言語処理技術を融合した情報マネジメントのツールやプラットフォームの構築を進める必要がある。

本特集では、自然言語処理技術を中心に、情報マネジメントの基盤技術の意義について一般読者にも分かりやすく興味深く読んでいただけるように解説する。また、情報マネジメント技術が、実際に組織内の情報分析・情報活用の場面で活用されている事例を紹介し、情報マネジメント分野の現状と将来の課題について述べる。1章では情報マネジメント技術の過去と現在を俯瞰し、2章でいくつかの先端的な事例に即してその現状を紹介し、3章では重要な技術的課題とそれらへのアプローチについて論ずる。

現在、情報マネジメントの中心は、定型的なデータベースからXML等に基づく非定型なデータへと移行し、マネジメントの対象も、単なるデータからそれに細かい意味内容を与えるメタデータへと拡大しつつある。その先には、非定型データの典型としてのテキスト、また、究極の知識表現、情報表現としてのテキストを管理し、利用する技術がある。以下では、情報の「意味」に関与する情報マネジメントの将来像を自然言語処理の立場から展望する。

(平成15年9月11日)