人工知能とバイオ関連産業の未来

東京大学 大学総合教育研究センター 副センター長・准教授
美馬秀樹

近年の科学の著しい発展や、工学の拡大、専門分野の深化を背景に、知識の量も爆発的に増加している. また、その結果として、分野が過度に細分化されているのが現状であるが、全体的な知識の俯瞰がなされているとは言えず、知識間の相互の関連が非常に分かりにくくなっている。つまり、1)知識の膨大さによる量的な問題、2)相互の関連の不明瞭さによる知識源としての質的な問題、により知識の活用がより困難な状況になっていると言える。

一方で、環境、エネルギーのような、地球規模での複雑で多様な問題が顕現化し、学際的、分野横断的知識の活用を促す仕組みの構築がより重要性を増しており、ものづくりと環境問題、さらには、安全性やデザイン性、倫理、法令遵守等は切っても切れない関係にある。つまり、社会の課題に対応するためには、分野によらず、より「広く」「深く」知識を習得する必要があるにもかかわらず、必ずしもこのような知識の獲得を効率的に支援する仕組みが提供されているとは言えないのが実情である。例えば、より具体的には、ナノマテリアル分野では、異なる材料分野を横断する知識を活用し、複合されたナノ材料を設計するのが一般的であるが、分野間の用語の統一性が乏しいため、従来の検索手法により必要な知識を探すのは非常に困難な状況になっている。

他方、 このような問題に対し、我々は先より、人工知能技術を応用することで、異なる分野における知識を効率的に抽出し、かつ得られた知識間の関連性から浮かび上がる新たな知の発見を支援する「知の構造化」に関するシステムの研究開発を行ってきた。本システムの特徴は、膨大なテキストデータを対象に、自然言語処理技術や機械学習によるオントロジー構築を自動で行うことで、知識の分析を高速かつ高精度に行うことにある。さらに可視化技術を統合することで分析した結果をグラフィカルにマップ表示し、関連する知識全体の横断的俯瞰を可能とする。

実際に本システムは、MIMA サーチとして、東京大学を始めとした主要大学における授業カタログ(http://catalog.u-tokyo.ac.jp/)、国立国会図書館での書誌情報検索・可視化システム等として実用化され、講義や文献の構造化マップ作成や関連性の俯瞰等の機能を実現している。また、近年では、バイオ関連テキストからの知識抽出や、臨床医療への応用研究も進められている。

本講演では、上記「MIMA サーチ」システムの概要を述べると共に、人工知能技術のバイオ関連産業への応用やその未来について概観する。