

## 新書本を用いた学問発見支援システムの構築

清水 花菜子

現在の科学コミュニケーションは、科学に興味がある人々が主な対象となっており、科学に興味がない人々に対するアプローチ不足が指摘されている。そのため、特に学問への興味が少ない人々に対し、学問に興味を持てるようにすることが重要である。また、大学入学希望者の多くは、入学以前から希望する学部・学科等を選択しなければならない。学部・学科等を選択する上で、学問に興味を持ち、興味のある学問分野を発見する必要がある。

以上から、本研究では、学問に興味を持ち、興味のある学問分野を発見するための学問発見支援システムを構築し、その有用性を確認する。提案システムでは、多様な学問分野の基礎的な内容を扱っており、手にされやすいという特性をもつ新書本を用いる。提案システムでは、利用者が選択した新書本に類似する学問分野を、興味のある学問分野として提示する。

具体的なデータセットとして、新書マップの「主なノンフィクション新書シリーズ」に該当するものを新書本であると定め、新書本に関する情報として、「BOOK」データベースのタイトル、内容説明、目次を用いた。また、学校基本調査の学科系統分類表を基に学問分野を定め、学問分野に関する情報として、逆引き大学辞典の学科紹介文を用いた。対象データは新書本 793 冊、学問分野 54 分野である。類似度算出手法は、BM25 による単語の重みを利用し、コサイン類似度で新書本に対する学問分野の類似度を算出する手法 1 と、BM25 で新書本と学問分野の類似度を算出する手法 2 を検討した。人手による判定結果から正解データセットを構築し、評価指標 nDCG で結果を評価した。手法 1 の nDCG は 0.78、手法 2 の nDCG は 0.59 であり、手法 2 より手法 1 のほうが有効であると言える。

続いて、提案システムの有用性を確認するために、提案システムとキーワード選択システムを比較するユーザ評価実験を、大学生 10 名を対象に行った。事前及び事後アンケート、比較アンケートの分析結果として、提案システムが興味のある学問分野を発見する、学問分野への興味の理由を明らかにする、学問分野への興味を深くするために有用なことが確認された。

今後の課題として、手法による結果の向上及びシステムにおいて表示する情報の検討が挙げられる。手法による結果を向上させる上で、新書本と学問分野間で単語が一致しないことが問題であった。そのため、新書本の本文を追加した。しかし、本文追加による結果の大幅な改善は見られなかった。今後は、語彙拡張及び対象データの追加や変更によって、結果の向上を目指したい。また、システムにおいて表示する情報の検討として、学問分野に適合しない新書本は、システム利用者視点からの有用性を考慮し、除外することが検討される。そのため、学問分野に適合しないと思われる新書本の判別手法も検討したい。さらに、学問分野に関する情報として、学問分野の紹介文を加える等、学問分野提示時に表示する情報も検討される。

(指導教員 高久 雅生)