

## 参考資料に着目したレファレンス事例に対する NDC の自動付与

石野 夏海

レファレンスサービスとは、何らかの情報あるいは資料を求めている図書館利用者に対して、図書館員が求められている情報あるいは資料を提供するサービスであり、利用者とのコミュニケーションスキル・資料に対する習熟・探索する能力など図書館員に高度な能力を要求するものである。そのため、レファレンス事例を記録・共有することが図書館員のレファレンスサービスに対する知識や経験の共有につながり、有効であるとされている。国立国会図書館が主導し、全国の図書館が参加するレファレンス協同データベースは、レファレンス記録の蓄積を実現したものである。レファレンス記録を有効活用するためには、利用者が必要とするレファレンス事例を効率よく検索できるようにする必要がある。それにはレファレンス記録への NDC の付与が有効な手段の一つである。しかし、レファレンス協同データベースでは約 4 分の 1 のレファレンス記録には NDC が付与されていない。原因としては、NDC の人手による付与には多大な労力を必要とすることがあげられる。

本研究では、レファレンス事例で用いられている参考資料に着目して、レファレンス記録に対して NDC を自動付与する可能性を検証した。レファレンス記録に対する NDC の自動付与の研究として、原田らのレファレンス事例中の質問文・回答文と NDC の組を学習用データとする機械学習による手法があるが、付与精度は高いとは言い難く、研究の余地が大きい。レファレンスサービスとは、資料に基づいて行われるものであり、レファレンス事例の特徴を端的に表しているものが参考資料であると考えられる。参考資料を用いることで、これまでの質問文・回答文を用いる手法より高い精度で分類を行える可能性があると考えた。手順として、(1)レファレンス事例データの参考資料の項目から国立国会図書館サーチ API を用いて検索を行い、参考資料の NDC を特定し、(2) 特定した参考資料の NDC とそのレファレンス事例の NDC の組を学習用データとして機械学習を行った。機械学習にはデータマイニングソフト Weka を用い、機械学習アルゴリズムは SVM、ナイーブベイズ、決定木の 3 手法を用いた。9 割を学習用データとし、残り 1 割のデータを対象に分類を行い、人間が付与した NDC と比較し、再現率・精度を求めた。レファレンス事例の NDC 1 桁目を自動分類する実験と、NDC2 桁目までを自動分類する実験を行った。また、参考資料に着目することの有効性を検証するために、原田らの先行研究手法を再現し、比較を行った。

実験の結果、すべてのアルゴリズムで先行研究手法を上回り、最も結果が良かったナイーブベイズで再現率 74.3%、精度 74.3% という高い値が得られ、参考資料に着目してレファレンス事例に NDC を自動付与することが有効であることが明らかとなった。また、NDC1 桁の分類だけでなく、NDC2 桁の分類においてもナイーブベイズで再現率 61.3%、精度 61.8% が得られ、実用に近づいたと言える。

(指導教員 辻慶太)