

「高楠日記」の自然言語処理による解析

研究員 時弘 哲治

佐々木 多希子

上山 大信



本稿では、1934年から1938年までの「高楠日記」の自然言語処理による解析と題する研究を紹介します。高楠順次郎（1866-1945）先生は、戦前を代表する仏教学者・教育者であり、武蔵野大学の創設者でもあります。武蔵野大学には、高楠先生が自筆で記した日記が13年分現存しています。本研究ではそのうち1934年から1938年までの日記を対象としました。本研究は、自然言語処理の手法を用いて高楠順次郎先生の日記の読解に数理工学の視点から裏付けを与え、時代や文化の違いを越えて、高楠先生の思想や足跡への理解を深めることを目指すものです。

自然言語処理とは、人間が用いる言葉をコンピュータで処理・分析する技術であり、感情分析、文脈解析、文章分類、機械翻訳など幅広い応用を持ちます。本研究では、日記の文章を言語データとして整理し、場所や人名の抽出、出現頻度の集計、年別頻度の比較、共起ネットワーク、共起行列、クラスタ分析などを行いました。これにより、記述の背後にある行動範囲や人的関係の構造を、定量的に可視化することを試みています。分析結果からは、場所として東京、沼津、上

野、ルンビニ、京都、三島などが高い頻度で現れ、人名としては須本、樺山、中橋、加藤、大久保侯、宇井、高楠順次郎、小野などが多く確認されました。また、共起ネットワークの解析により、特定の場所や人物が複数の記述を結び付ける中心的な役割を果たしていることが示されました。こうした結果は、高楠先生の移動、交流、関心の広がりや、資料全体の構造として捉え直すための有効な手掛かりとなります。さらに、年ごとの頻度の違いを追うことで、高楠先生の活動拠点や交友の変化を時間軸の中で捉えられる点も、本研究の重要な成果の一つです。また、こうした分析は、史料を通覧しただけでは見えにくい関係性や傾向を浮かび上がらせ、従来の読解を補完する新しい手法として期待されます。本研究の特色は、仏教研究や歴史資料の読解に、自然言語処理と数理工学の方法を導入した点にあります。従来は個別の記述を丁寧に読むことに重きが置かれてきた日記資料に対し、頻度、共起、ネットワーク構造といった視点を与えることで、高楠順次郎先生の思想と活動をより立体的に理解できる可能性が開かれます。今後は、個別記述の内容解釈と計量的分析とを往還させながら、人文学と数理工学の連携の新たな展開を目指したいと考えています。

なお、本研究の遂行にあたっては、石上和敬先生より多大な御支援、御協力を頂きました。ここに深く感謝申し上げます。