

審査の結果の要旨

氏名 矢田 竣太郎

近年、急激な書店数の減少や電子書籍の普及により受動的に図書に遭遇する機会が減っている。それらの代替として展開しているデジタル環境では受動的接触環境を十分に提供できていない。その一方で、最近の研究で、積極的読書だけでなく受動的に図書に接触する環境がリテラシーに与える影響の重要性が指摘されている。

こうした状況を背景に、本論文は、デジタル環境とりわけ SNS (twitter) 上で受動的な図書遭遇環境を代替するメカニズムを構築するために必要な課題と概念を整理し、メカニズムの中核を担う基礎技術を開発するものである。対象言語としては日本語を扱う。論文は V 部に分かれ、全体で 10 章からなる。

第 I 部第 1 章ではこの問題背景を踏まえデジタル環境上に受動的図書遭遇環境を構築するという課題を明らかにしている。第 2 章で、関連研究としてデジタル環境における読書推進手法とシステム、受動的接触の要件に関する概念をレビューし研究の位置を定めている。

第 II 部では受動的図書遭遇環境の概念を整理する。第 3 章ではデジタル環境における図書遭遇に関与する属性として日常性・近接性・好評性・非脅迫性の 4 つを定義し、システム構成とその要件を示す。第 4 章では属性を実証的に評価しその妥当性を確認している。

第 III 部は図書に言及するツイート (TMB) からなるデータの整備を扱う。第 5 章で包括的な書名データベースを用いて約 2 万ツイートからなる TMB コーパスを構築している。このコーパスには人手で好評性と非脅迫性のタグが付与される。これは、ツイッター上での図書関連情報の実態把握と機械学習用教師付きコーパスとして利用される。第 6 章でコーパスが分析され、TMB の属性の分布とそれに関わる言語表現上の性質が整理される。

第 IV 部第 7 章では TMB 識別器の開発を行っている。ボットフィルターとノイズ除去の 2 段階からなる、SVM と最大エントロピー法を用いた手法を実装し、既往手法を上回り実用に耐えうる性能を実現している。第 8 章では好評性と非脅迫性の同定器の開発を行っている。好評性は正・負・中立の 3 値分類タスクとして機械学習を用いた手法を、非脅迫性は表現の強さを示すルールに基づく手法を開発し、優れた性能を達成している。

第 V 部第 9 章では研究成果を整理し、第 10 章で今後の展開の方向性を示す。

本研究は、図書および図書情報に受動的に接することを可能にする環境という、図書館や書店の存在が暗黙に担っていた基盤的機能が、デジタル環境でどうなるかという理論的な問題を明確にし、受動的図書遭遇機能を担う機構を SNS 環境の中で実現するために求められるシステムの概念を示したという点で、デジタル環境時代の図書館情報学に対し理論的に重要な貢献をなすものである。また、システムの中核的な機能を実装し高い性能を実現した点でシステムの実現可能性を示すと同時に言語処理研究にも貢献している。よって、本研究は博士 (教育学) の学位を授与するにふさわしい水準にあるものと判断された。