

## NFC-1 「Total Environment for Text Data Mining (TETDM)」

チャレンジャー：砂山 渡 (広島市立大学)

本チャレンジは、世の中に分散しているテキストマイニングツールを統一的に扱え、分析に没入して新たな発想を得られる環境の構築を目指している。全国大会では、チャレンジの採択後、今回で5回目のサバイバルとなり無事に「卒業」となった。本年度の全国大会に先駆けてTETDMバージョン1.0を完成させ、現在公開している。本チャレンジの詳細、ならびにツールのダウンロードについては、TETDMサイト (<http://tetdm.jp>) をぜひともご覧いただきたい。

本年度の全国大会では、11件の発表があり、延べ60名以上の聴講者とともに活発な議論を繰り広げた。発表は、「TETDMの達成内容、活用事例の報告」、「TETDMを用いたシステム構築」、「TETDMのシステム関連と今後に向けて」の三つのセッションに、オープニングとクロージングディスカッションを加えて行った。以下で発表の概要を報告する。

オープニングで、TETDMの特徴について述べた後、「テキストデータマイニングのための統合環境 TETDM」(砂山：広島市立大学)では、構築したTETDMの概要と本チャレンジの5年間の歩みと成果について述べた。

「教育機関におけるTETDMの活用事例報告」(梶並知記：神奈川工科大学)では、教育機関における、研究、講義、演習における実践的なTETDMの活用事例についての報告がなされ、今後の活用の可能性が示唆された。

「TETDMを用いた電子カルテ分析支援ツール」(串間宗夫：宮崎大学医学部附属病院)、および「TETDMを用いた電子カルテテキストデータ分析」(山崎友義：宮崎大学医学部附属病院)では、大学病院における電子カルテデータの活用に向けて、構築したシステムと現場からの意見について報告がなされ、今後の幅広い応用の可能性が示唆された。

続いてのセッションでは、TETDMに実装済み、あるいは今後実装して活用される可能性があるシステムについての提案がなされた。「議論の進捗を促した発言の抽出と議論の流れの可視化」(西原陽子：立命館大学)では、議論の会話データの中から、議論に貢献する有効な発話を抽出する方法の提案がなされた。

「TETDMを用いた協調的マルチビューによるインタラクティブ文書クラスタリングの提案」(高間康史：首都大学東京)では、対話的に文書クラスタリングを行う際に、複数の粒度のデータを閲覧可能にするシステムの提案がなされた。

「ARとテキストマイニングを用いた対話時の好感度推定によるコミュニケーション支援」(後藤賢悟：広島市

立大学)では、対話データをもとに、相手の好感度を上げるアシストをするシステムの提案がなされた。

「テキストから得られる複数特徴量を融合する絵本類似探索法」(服部正嗣：NTTコミュニケーション科学基礎研究所)では、複数の特徴量を利用して、利用者の絵本探索を支援するシステムの提案がなされた。



図1 TETDMセッションの様子

最後のセッションでは、TETDMの利用や開発のための環境についての発表がなされた。「TETDMの利用者向けチュートリアルシステムの開発」(川本佳代：広島市立大学)では、TETDMの利用者を支援する、目的指向あるいは目的非依存のチュートリアルの提案がなされた。

「テキストマイニング共通語彙基盤の構築とツール実装への適用に関する検討」(阿部秀尚：文教大学)では、TETDMで取り扱うオブジェクトとメソッドの概念化により、汎用的で扱いやすい開発環境の提案がなされた。

「探索過程の視覚化によるテキスト集合からの知識獲得支援に関する研究」(岡本香帆里：関西大学)では、知識獲得に向けて、分析を繰り返す試行錯誤のプロセスを支援する枠組みの提案がなされた。

最後にクロージングディスカッションとして、TETDMの「実用化が期待される場面」、「利用者拡大の方策」、「知識創発への期待」の三つのテーマについて議論を行った。(議論の詳細は前述のTETDMサイトを参照されたい)

本チャレンジは「卒業」という一つの区切りを迎えたが、統合環境TETDMが真価を発揮するのは、利用者層の拡大と開発者の増加が相乗的に繰り返される状況と考えている。そのため、今後も研究会活動、その他公開講座などにおいて、TETDMの普及を積極的に進めたいと考えている。

皆様にもぜひ一度、TETDMをお試しいただき、忌憚のないご意見を頂戴するとともに、企業、大学での積極的な利活用を視野に入れていただけると幸いです。

(砂山 渡 (広島市立大学))