

NFC-01 「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3回目のサバイバルとなった。5年以内に実現という近未来チャレンジの制約に対して、3年間で終了したことになり、今後2年間で統合環境の完成を目指している。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>) をぜひともご覧ください。

今年の全国大会では、9件の発表があり、40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の基本的な環境づくりの話から、テキストマイニングツールの実装例、既存ツールとの連携、今後の拡張の可能性などについて行われた。以下で発表の概要を報告する。

「統合環境 TETDM を用いたテキストマイニング初心者のためのツール選択支援」(中垣内李菜：広島市立大学) では、テキストマイニングに不慣れなユーザが、TETDM の多様なツールの中から適切なツールを選んで用いられる環境の提案を行った。

「TETDM モジュール構成に基づくテキストマイニングメソッドの概念化に関する一考察」(阿部秀尚：文教大学) では、TETDM の多様なツールの入出力に着目し、ツールを整理したうえで、利用者と開発者の双方がツールを利用しやすくなる方法の検討を行った。

「ユーザの観点に基づく電子掲示板からのコミュニケーション抽出」(里中晴日：広島市立大学) では、電子掲示板で行われているコミュニケーションに対して、ユーザの興味を表す観点に沿ったやり取りが行われている書込みを抽出することで、要約を行う方法を提案した。

「トップダウン・ボトムアップな文章構造作成のための推敲支援システム」(山手砂都美：広島市立大学) では、段落間の関係を明示するとともに、段落構成をよりトップダウンに変化させるための具体的な方法を提示し、トップダウンな文章作成を支援する方法の提案を行った。

「TETDM を利用した情報系専門教育の実践例」(梶並知記：神奈川工科大学) では、マイニングのアルゴリズムを学習させる目的で、TETDM 上にマイニングの簡易的なプログラムを実装しその効果を実験的に確認、検証できる教育の実践例について示した。

「テキストデータマイニング統合環境を利用した看護記録からの専門用語辞書作成支援ツールの提案」(高間康史：首都大学東京) では、電子カルテデータに含まれ

る専門用語の辞書を構築するために、医者などの専門家に専門用語候補を提示し、その選別を行ってもらえるツールの提案を行った。

「Rによるテキストマイニング用 TETDM モジュール開発」(徳永秀和：香川高等専門学校) では、既存のソースの一つとしてあげられる R の機能を TETDM 上で利用することを目的として、両者の連携の可能性と、その具体的な方法の提示を行った。

「FACT-Graph と TETDM の融合の可能性」(佐賀亮介：大阪府立大学) では、時系列データの関係と変化を明示するグラフをベースに、TETDM との融合により、より効果的にデータの解釈を促すインタフェースづくりの可能性についての検討を行った。

「テキストデータマイニングのための統合環境 TETDM の活用と実践」(砂山 渡：広島市立大学) では、現在までに構築された TETDM の概要とその実践例を示した。今後は、モジュール収集に向けたモジュール開発者支援と、知識創発に向けた利用者支援を目指すことを示した。

便利で効果的なツールはどんな分野でも望まれるものだが、現在は、ビッグデータ、クラウド、リンクトデータなどにも見られるように、多くのデータがつながる時代となっている。ただしデータのみが繋がっても、その解釈を行う人間をアシストする機能が整わなければ、新しいものは生み出されない。

本チャレンジでは、データを分析するためのツールをつなぐことを目指し、またデータのパイプライン処理が可能というレベルではなく、ツール間の有機的な運動と、データの解釈を生み出すための機能を実装していくことで、これからの時代に不可欠なデータ分析ツールとしての TETDM を構築している。

本チャレンジは、チャレンジャーのみによって達成されるものではなく、現時点でチャレンジャーを含め10名以上のコアメンバとともに開発を進めている。また、本全国大会だけでなく、本学会の「インタラクティブ情報アクセスと可視化マイニング研究会」(主査：加藤恒昭先生) (<http://must.c.u-tokyo.ac.jp/sigam/>) と連動して、多くの方との意見交換により、汎用的な環境づくりを目指している。

少しでもご興味をもたれた方がおられましたら、メンバとしての参加、ツールの試用、全国大会、研究会へのご参加を、常に歓迎しておりますので、皆様のご高配を賜ることができましたら幸甚に存じます。

(砂山 渡 (広島市立大学))