

ツールの操作履歴の共有によるテキストマイニング時の発想支援

Creativity Support for Text Mining System by Sharing Tool Operation History

金谷 直哉 砂山 渡
Naoya Kanaya Wataru Sunayama

広島市立大学大学院 情報科学研究科
Graduate School of Information Sciences, Hiroshima City University

Abstract: The total environment for text data mining, includes a number of mining tools, which allows the text analysis in a variety of ways. When performing text mining, it is necessary to analyze the results of the analysis widely collected. However, it is limited in the idea of one person. If it is possible to see how to use the unknown tools, it may be connected to the idea of obtaining new results. Therefore, this study suggests function for TETDM user to share with each other the operation history of the tool.

1 はじめに

近年、非常に多くのアプリケーションソフトウェアが登場し競争が行われている。これによりアプリケーションソフトウェアの多くは非常に多機能になり、様々な目的に利用することができるようになった。それに伴い我々はコンピュータを日常的に利用し、様々な活動を効率的に行えるようになったが、便利になる反面、豊富な機能の中から有用な機能を選び出し使いこなすことは非常に難しくなっているという点が課題となっている。

豊富なテキストデータマイニングツールを提供するテキストデータマイニングツールのための統合環境として TETDM[1][2] がある。TETDM では多彩なテキストデータマイニングツールを提供しており、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。TETDM で利用できるツールの機能は多様で様々な目的に合わせて使うことができる。しかし、それぞれのツールの使い方は異なるため、目的に応じすべてのツールを使いこなすことは難しい。さらに TETDM では柔軟に新しいツールを開発しシステム上に実装することができ、新たなツールの提供も行われるため、最適なツールの選択を行うには、高い技術が求められる。

そこで本研究では、TETDM 上の機能の 1 つとして、操作履歴を保存する機能、他者の操作履歴を元に操作を再現する機能を実装し、様々な利用者が行ったツール操作を自由に閲覧できるようにすることで、TETDM のツールによるテキストマイニングを行う際の発想を支援し、利用者のテキストマイニングを支援することを目的とする。

2 関連研究

本章では TETDM によるマイニングの発想支援の関連研究を、操作履歴共有、操作の再現手法の観点からまとめる。

機能実行履歴の共有によって、多機能なソフトウェアの有用な機能の発見を促す研究がある [3]。この研究は、ソフトウェア使用目的がよく似た人同士で、機能使用履歴を参照可能にすることで、ユーザーがそこから有用な機能発見を可能にすることを目指した研究で、Word、PowerPoint の利用機能を共有する評価実験を行ったところ、平均 16.1 個の未知の機能を発見することができ、被験者はそのうち約 40 % は有用な機能であったと解答した、という結果になった。本研究では、この理論を用いて、TETDM のツール利用の発想支援を目指す。

また視線行動を記録し注視点を可視化することで、着眼スキルを発達させる支援を行う研究がある [4]。この研究では可視化された他ユーザーの注視点と自身の行動を比較することで、知識や思考の違い感じ取り着眼スキルを発達させることができるミラーエージェントシステムを開発した。視線行動の入出力はマウスカーソルによって行われた。評価実験として自動車運転時の危険予知トレーニングを行い、着眼スキルが有意に向上していたことが分かった。この研究から他者の行動履歴の再現から他者の感覚を感じとることができることを示唆していた。そこで本研究では他者のツール操作の再現から操作の発想を獲得する支援を目指す。

Web アプリケーションの動作検証を目的として、操作のリプレイを再現する手法を開発した研究がある [5]。

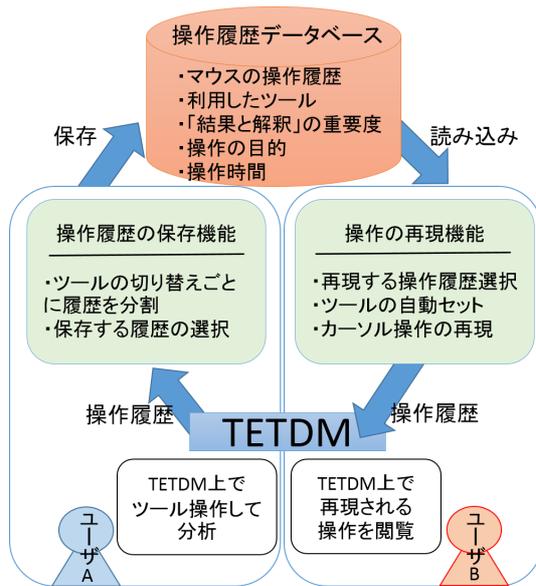


図 1: テキストマイニング時の発想支援の枠組み

このシステムは Capture 部と Replay 部で構成され、Capture 部では Web アプリケーション上で操作が行われると、ページを進む、戻る、ブックマークの操作や、クリック、スクロール、マウス移動などのカーソル操作のイベントを保存する。Replay 部ではこうして保存された操作履歴の内の Web アプリケーションのページ遷移の再現を行う。このシステムではマウスカーソルの操作の記録を行っていることから、より詳細な再現を可能にすることでアプリケーションの操作ガイドに応用できる可能性がある、と述べている。本研究では TETDM のツールを使ったマイニングの発想支援を目的として再現を利用する。

3 テキストマイニング時の発想支援

TETDM には数多くの機能が実装されており、TETDM 上でのテキストマイニングの方法も一人ひとり異なる。本研究では、TETDM のツールの操作履歴を再現できる機能を実装し、利用者がお互いに操作履歴を共有できるようにすることで、テキストマイニング時の発想を支援し知識の発散を促すことを目的とする。

TETDM は、ツールを利用して特徴的な結果を得た際に、それを記録するための「結果と解釈」の登録インターフェース (図 2) を備えている。ここで、さまざまな分析結果を得て、より多くの結果と解釈を登録し、知識創発における思考の材料をより多く集めることが、根拠があり質の高いアイデアの生成につながると考え、この結果と解釈の登録数が増えることを目指す。

発想支援の機能は、操作履歴のデータベース、操作

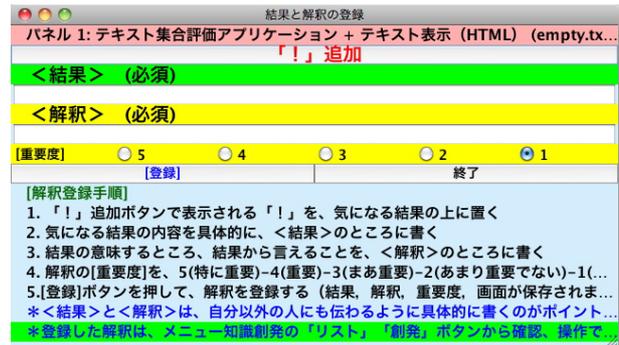


図 2: 結果と解釈の登録インターフェース

表 1: 記録されるツールの操作

TETDMの操作	記録される操作	再現の可否
カーソル移動	○	○
ボタン	○	○
チェックボックス	○	○
スクロールバー	○	○
コンボボックス	○	○
ドラッグ	○	○
ホイール	○	○
テキスト入力	×	×
メニューウィンドウのボタン操作	○	×
ツールの変更操作	○	×
「結果と解釈」の入力	○	×
TETDMのその他の機能	×	×

履歴の保存機能、操作履歴の再現機能で構成されている。テキストマイニング時の発想支援の枠組みを図 1 に示す。以下でこの枠組みの詳細について述べる。

3.1 操作履歴のデータベース

操作履歴保存機能と操作履歴の再現機能の間でデータベースを介して、操作履歴の受け渡しを行う。データベースにはツールを利用した際の以下の情報が保存される。

- 表 1 の操作が行われたディスプレイ上の XY 座標
- 操作履歴でセットされたツールの組み合わせ
- ツールを操作していた時間の長さ
- ツール操作から得られた出力結果を元に登録された「結果と解釈」の重要度の合計点
- PC の OS, ディスプレイサイズ, 処理性能
- 操作した人が保存時に入力した操作の目的

3.2 操作履歴保存機能

3.2.1 操作履歴の分割

TETDM によるテキストマイニングでは、目的とする結果を得るため様々なツールに切り替えながら処理

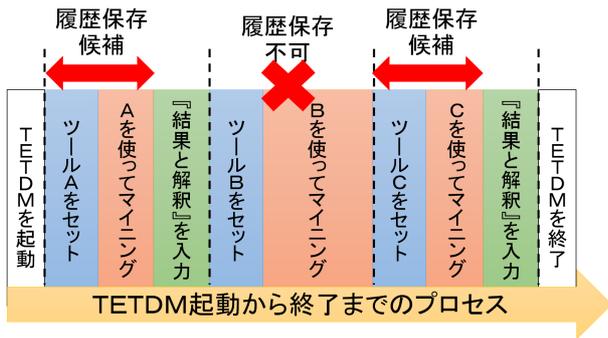


図 3: 操作履歴分割の仕組み

結果の分析を行う。データベースに保存する操作履歴を作成するため、ツールの組み合わせが変更されたタイミングでカーソル操作の操作履歴の分割を行う。そのため、1つの操作履歴には1つの組み合わせのツール操作のみの履歴が保存される。また本機能では、TETDMにあらかじめ用意されたツールの組み合わせを利用したツール変更のタイミングを分割するタイミングとして利用した。図3に操作履歴保存の仕組みを示す。

3.2.2 操作履歴の保存

TETDM 操作終了時に、利用者は分割した操作履歴の中から有用な操作を行ったツールを選択し保存する。ツールを操作した後、出力結果を元に「結果と解釈」を登録した場合、そのツール操作は有用な可能性が高いと考えられることから、保存操作を行う際は「結果と解釈」の入力操作を行ったツールのみ選択し保存することができる仕様にした。保存されるデータは図3の仕組みで分割された TETDM のツール操作とそのツール ID、そのツール操作時に登録された「結果と解釈」に与えた重要度の合計点、PC に関する情報があり、保存操作がされたとき 1つの再現用操作履歴としてデータベースに保存される。

3.2.3 操作履歴保存を行うインターフェース

TETDM の利用を終了する際、有用な結果が得られたツールの操作履歴を選択し保存する機能を TETDM の機能の1つとして実装した。TETDM の終了操作を行うと、TETDM の終了と入れ替わりに操作履歴保存ウィンドウがポップアップし、このウィンドウから操作した履歴の保存を行うことができる。操作履歴保存ウィンドウには、TETDM でセットしたツールの組み合わせ一覧が表示される。一覧にはツールを使って登録された「結果と解釈」の重要度合計点、ツールを操作した時間も合わせて表示される。

操作履歴一覧にはセットしたツールの組み合わせが

全て表示されるが、重要度合計点が 0 点の場合は選択することができない。保存可能な操作履歴を選択した後、『操作履歴保存』ボタンをクリックするとさらに操作履歴のタイトルと操作の目的を入力するウィンドウがポップアップする。各項目を入力することで、操作履歴の保存が完了する。

3.3 操作履歴の再現機能

3.3.1 再現する履歴の選択

TETDM 上で再現機能を起動すると、データベースから読み込んだ操作履歴の一覧が表示され見たいものをクリックすることで操作履歴の選択が行える。操作履歴一覧をクリックした段階でカーソル操作の履歴がデータベースから読み込まれ TETDM には自動的に操作履歴で利用されたツールがセットされるが、カーソル操作の再現は行われずさらに再現開始を確定する操作を行う必要がある。利用者は再現を閲覧するかどうか、実際にツールを見ながら選択することができる。

3.3.2 再現の実現方法

操作履歴の再現機能では、データベースに保存された操作履歴を参照し TETDM のツールの操作を再現する。操作履歴の再現は、ツール上でのマウスカーソルの動きの再現によって行い、本研究ではカーソルの動作の再現を行うため、Java の Robot クラスを利用した。

本機能ではツール上での操作のうち、クリック、ドラッグ、ホイール操作の再現を行い、TETDM を操作する。操作履歴に記録されたディスプレイ座標にマウスカーソルを移動させ、マウスの左ボタンの入力、ホイールの回転操作の命令を Robot クラスにより入力することで、操作履歴と同じ操作の再現を実現した。

再現中でもカーソルの操作が可能な時間として、一定回数操作が行われるたびに 5 秒間の再現中断間隔を設けた。再現を開始するとマウスカーソルが自動的に動き始め自由に操作することができなくなり、利用者は自分のしたい操作をするために再現がすべて終わるまで待たなければならなくなる。利用者は再現中断間隔を利用することで、再現が終わるまで待たず操作を中断したり調整をすることが可能である。再現中断間隔までの操作数は再現開始時に選択することができる。

3.3.3 再現を行うインターフェース

操作履歴をデータベースから選択し再現の閲覧を行う機能を TETDM の機能の1つとして実装した。TETDM

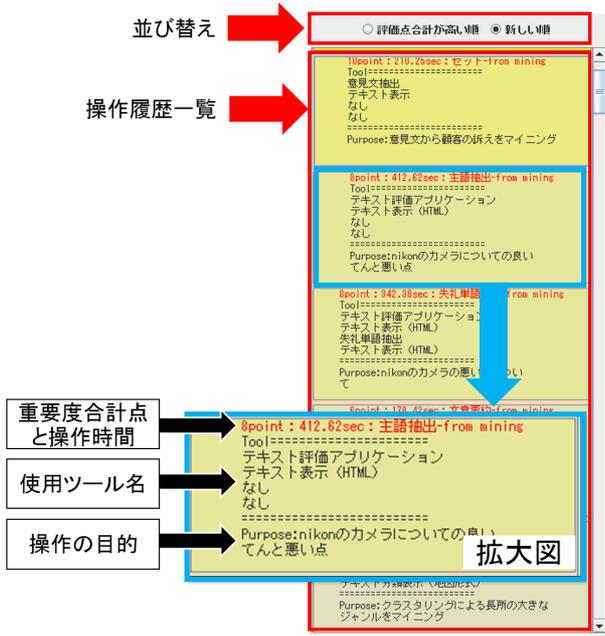


図 4: 再現ウィンドウの操作履歴一覧



図 5: 再現制御ウィンドウ

のメニューウィンドウから『再現』の項目を選択することで、再現ウィンドウがポップアップし再現を閲覧するための操作を行うことができる。このウィンドウでは再現速度、再現中断間隔の設定を変更することができる。

再現中の速度は、再現が始まった後でも再現中断間隔を利用して変更することができる。

さらに再現ウィンドウでは再現したい操作履歴の選択を行うことができる。再現ウィンドウの右側に操作履歴が一覧で表示されており、各操作履歴に関する以下の情報を確認することができる。

- 操作から得られた重要度の合計点
- 操作履歴の操作時間
- 利用されていたツールの種類
- 操作履歴保存時に入力された操作の目的

操作履歴一覧から閲覧する操作履歴を選択しクリックすると、TETDMに操作履歴のツールの組み合わせが自動的にセットされる。最後に『再現開始』ボタンを押すと、再現ウィンドウの表示が消え再現が開始される。

3.3.4 再現制御ウィンドウ

再現中断間隔を利用して再現速度の変更や停止操作を行うことができる図5の再現制御ウィンドウを開発した。このウィンドウは再現ウィンドウの『再現開始』ボタンを押すと同時にディスプレイ左上に表示される。

再現制御ウィンドウの上部の三角は現在の再現速度を表し、色のついた三角が多いほど速度が速くなっていることを表している。次に中央の水色のメーターは再現進行度を表し、水色のメーター量で再現がどれくらい進んでいるかいつでも確認することができる。下部の4つのボタンは再現の制御用のボタンで、再現の一時停止、終了、速度変更を行うことができる。

4 操作履歴の再現機能の評価実験

本章では、TETDMのツール操作履歴の再現機能を利用し他者のマイニングを閲覧することで、新たなテキストマイニングの発想を獲得し、そこから新たな知識を創発ができるのかを検証した実験について述べる。

4.1 実験方法

実験は、TETDMのチュートリアル『初心者』『初級』の課題をクリアしたことがある大学生と大学院生12名を対象として行った。実験では全被験者に同一スペックのPCを使用してもらい、2日間でTETDMによるテキストマイニングを行ってもらった。

マイニング課題においては、通信販売サイトで販売される nikon 製カメラのレビューを収集し使用した。レビューは全40件をまとめたもので、多様な知識を発見できるように評価値5, 4のものと1, 2のものからほぼ同数になるよう収集した。“（被験者名）カメラ”の社員であると仮定して、ライバル会社 nikon に勝つためのカメラのアイデアを考えてもらい、そのために「結果と解釈」機能に、nikon のカメラの特徴とそこから得られた解釈を登録を行ってもらった。

1日目の実験では、被験者にTETDM上でレビュー記事をテキストマイニングしてもらい「結果と解釈」を登録してもらった。「結果と解釈」は最低でも10件以上の登録を目指してもらい、行った操作の履歴を保存してもらった。

2日目の実験では、TETDMでレビュー記事をテキストマイニングし1日目のものに追加する形で「結果と解釈」を登録してもらった。「結果と解釈」は可能な限り登録することを目指してもらった。2日目では再現機能により1日目の他の被験者の操作を閲覧することができ、分析に取り入れることができるようにした。これによって個人では得られなかった分析結果の獲得が可能になるか検証した。

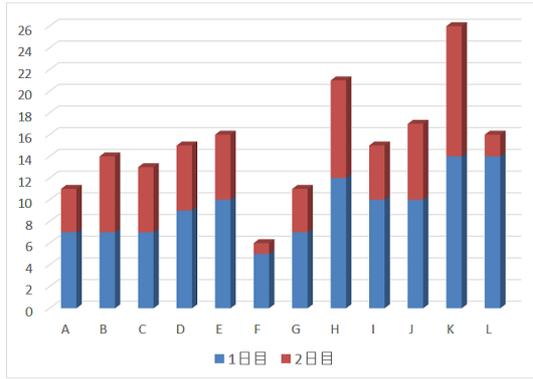


図 6: 登録された「結果と解釈」の数

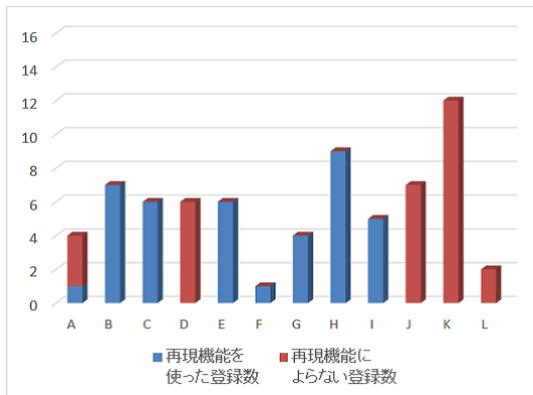


図 7: 操作履歴閲覧後の登録数

4.2 結果と考察

4.2.1 「結果と解釈」登録数の伸び

TETDMを利用して新しい登録できなくなるまで「結果と解釈」の登録を行った被験者でも、再現機能を利用することで、マイニングの新たな発想を獲得し、新たな「結果と解釈」を入力することができるようになるのかを検証するため1日目の時点の「結果と解釈」数と2日目の累計「結果と解釈」数を比較した。結果を図6に示す。1日目のマイニングでは「結果と解釈」を最低でも10個以上登録することを課題としたが、半数の被験者は10個以上「結果と解釈」として登録する結果を得ることができなかったが、図6を見ると、全被験者において再現機能を使って再度マイニングをすることで、「結果と解釈」数が増えており、マイニングの新発想を得ることができたことがわかる。

4.2.2 操作履歴の閲覧から得られた「結果と解釈」

「結果と解釈」の登録数の増加が再現機能によって発生したものあることを検証するため、再現機能を利用して登録された「結果と解釈」の集計を行った。

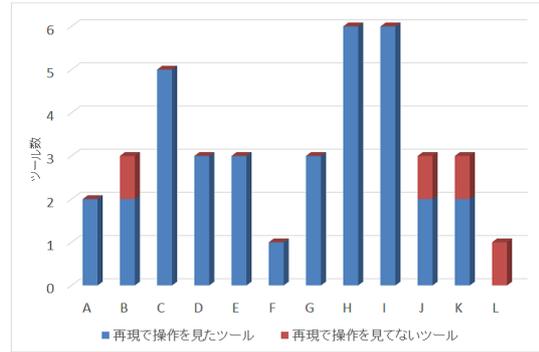


図 8: 2日目の「結果と解釈」登録に繋がったツール数

再現機能を利用し他者の操作を閲覧した直後、再現機能で見たツールを利用して登録された「結果と解釈」は、再現機能の閲覧によって得られたマイニングの発想によって、創発されたものであるといえる。そこで2日目の実験において登録された「結果と解釈」のうち、再現機能を使って操作を見た後そのツールから別のツールに変更せず登録された「結果と解釈」の数の集計を行った。結果を図7に示す。

図7から、再現機能を使い新たに登録された「結果と解釈」が多数あったことが分かる。この被験者の再現機能の使い方を詳細に検証すると、再現を見てその後操作せず「結果と解釈」を登録している場合と、再現を見た後ツールを操作して「結果と解釈」の登録を行った場合が見られた。これは他者の操作が出した出力結果から新たな発見をしたという場合と他者の操作方法から操作方法の発想を得た場合があると考えられる。さらに、再現途中で再現を停止させて、自身で操作を行った例も見られ、これはカーソル制御によって再現を行う利点として考えられる。

4.2.3 「結果と解釈」登録に繋がったツール数

図7では操作履歴閲覧直後そのツールを利用した登録のみを操作履歴の閲覧により得られた登録とした。しかし、操作履歴を閲覧してしばらくして、後に「結果と解釈」登録に繋がる可能性も考えられる。

被験者が「結果と解釈」の登録に利用できたツールは『そのツールでマイニングを行う発想がある』といえる。そこで2日目で各被験者が「結果と解釈」の登録に利用したツール数の内、「結果と解釈」を登録するより以前にそのツールの操作履歴を見ていた数がどのくらいあるかについて検証を行った。図8では、2日目の実験で各被験者が「結果と解釈」の登録に利用したツールの種類と、事前に操作履歴を見ていた割合を示す。

図8から、ほとんどの被験者は再現によって見たツールを使って「結果と解釈」の登録を行っていたことが

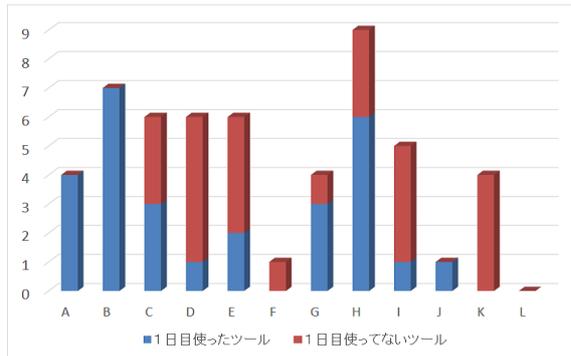


図 9: 2 日目に再現で操作を見たツールで登録した「結果と解釈」数

分かる。このことから操作履歴再現の閲覧が、ツールを使い方の新しい発想に繋がったことが分かる。また、アンケートのコメントに「使わなかったツールを使うきっかけになった」というものが複数あったことから、再現の閲覧がツールの使い方を考えるきっかけになり、それが新しい「結果と解釈」の登録に繋がったことが考えられる。

4.2.4 再現で操作を見たツールで登録した「結果と解釈」数

図 9 では、2 日目に再現機能で操作を見たツールから得られた各被験者の「結果と解釈」数と、その内の 1 日目に使っていないツールから得られた「結果と解釈」数の割合を示す。

この結果、1 日目に使っていないツールを利用して「結果と解釈」を登録する例がほとんどであることが分かった。このことから、再現機能で操作履歴を閲覧する機能は特に、操作方法を知らないツールの使い方を知る、有効に使えなかったツールを活用する発想を得る、使わなかったツールを使うきっかけになるという目的において効果的であると考えられる。または、1 日目に使っていたツールで新たな「結果と解釈」を得ていたという例も複数見られ、これは使い方を知っているツールの新しい使い方の発想を得る目的において効果を発揮したためであると考えられる。

4.2.5 「結果と解釈」で登録された内容の比較

1 日目と 2 日目で同一のツールを利用して登録された「結果と解釈」の比較を行った。失礼単語を抽出するツールを使った登録を見ると 1 日目では「ボタンの操作性はいまいち」と曖昧な文章が登録されていたが、2 日目では「AF ON ボタンが無い」という比較的具体的な文章になっていた。他にも詳細な文章になってい

る例は複数見られ、このことから他者の操作履歴を閲覧することでツールの使い方の発想を獲得することで、以前は行えなかった詳細な分析を行えたと考えられる。

5 結論

TETDM を利用したマイニングを行う際の発想を支援するため、TETDM の機能の 1 つとしてツールの操作履歴の保存、保存した操作履歴の再現機能を開発し、操作履歴の共有を可能にした。

操作履歴の再現機能の効果を検証するために評価実験を行った。レビュー記事についてテキストマイニングを行い商品の特徴を発見するという課題を、操作履歴の再現機能を導入して行ってもらったところ、再現機能で他者の操作履歴を閲覧することがより多くの特徴の発見に繋がること、多くのツールを利用したマイニングに繋がることが分かった。

今後の課題として、操作履歴の再現の精度の向上、再現に対応した操作の充実、操作意図の伝達効果の向上によって、マイニング時の発想をより効果的に支援することを目標としていきたい。

参考文献

- [1] Total Environment for Text Data Mining (テキストデータマイニングのための統合環境), (URL) <http://tetdm.jp>.
- [2] 砂山渡, 高間康史, 西原陽子, 徳永秀和, 串間宗夫, 阿部秀尚, 梶並知記, 松下光範, Bollegala Danushka, 佐賀亮介, 河原吉伸, 川本佳代: テキストデータマイニングのための統合環境 TETDM, 第 29 回人工知能学会全国大会, 2E3-NFC-01a-1, (2015).
- [3] 森崎修司, 白石裕美, 門田暁人, 大和正武, 松本健一, 鳥居宏次: 機能実行履歴を用いたソフトウェア機能の発見支援システム, 電子情報通信学会論文誌 D - I, Vol.J84 - D - I, No. 6, pp.755-767, (2001).
- [4] 藤本武司, 砂山渡, 山口智浩, 谷内田正彦: 視線行動の可視化による着眼スキル伝達支援, 人工知能学会論文誌, Vol.71, No.1, pp.517-518, (2009).
- [5] 柿元宏晃, 佐野博之, 大園忠親, 新谷虎松: Web 行動リプレイシステムに基づく Web アプリケーション動作検証システムとその応用, 第 71 回情報処理学会全国大会講演論文集, Vol. 19, No. 3, pp.174-183, (2004).