

データ分析スキルの育成に向けたゲームモードと キャラクターアシストチュートリアル

砂山 渡

滋賀県立大学工学部

sunayama.w@e.usp.ac.jp

概要 コンピュータによる知能を生かすには人間の知恵が必要であり、人間がいかにコンピュータを使いこなすかが問われる時代となって来ている。また、世の中に溢れるデータを活用する必要性が高まる中、データ分析を行える人材の育成が求められている。我々のグループでは、これまでにテキストマイニングのための統合環境 TETDM を構築している。TETDM は、汎用的な分析を行うための多くのツールを備えるとともに、データ分析に不慣れな人においても、データ分析のスキルを身につけてもらうための機能として、ゲームモードとキャラクターアシストチュートリアルを提供している。本稿においては、これらの機能の説明、テキスト分析スキルとの関連、これまでの利用実績について述べるとともに、本環境の整備と普及につなげるための方策について議論したい。

キーワード テキストマイニング, データサイエンス, ゲーミフィケーション, マイニングスキル, TETDM

1 はじめに

昨今の AI ブームにおいては、AI が全てを解決してくれて、人間が何もしなくてもいい便利な社会をイメージする人も多い。しかし、コンピュータによる知能を生かすには人間の知恵が必要であり、人間がいかにコンピュータを使いこなすかが問われる時代となって来ている。また、世の中に溢れるデータを活用する必要性が高まる中、データ分析を行える人材の育成が求められている。大学等でデータ分析を専門的に学ぶ人の総数は少なく、データを幅広く活用するためには、データが集められる部署で働く多くの人や、情報系以外の大学生が、データ分析のスキルを身につけられる環境の整備が必要と考える。

我々のグループでは、これまでにテキストマイニングのための統合環境 TETDM[1]¹ を構築し、オープンソースのフリーソフト²として公開している。TETDM は、汎用的な分析を行うための多くのツールを備えるとともに、データ分析に不慣れな人においても、データ分析のスキルを身につけてもらうための機能として、ゲームモード [2] とキャラクターアシストチュートリアル [3] を提供している。

データ分析の初心者がデータマイニングのスキルを身につけるためには、1) 基本的なデータ分析の手順の理解、2) ソフトウェアの操作方法の理解、3) 継続して学ぶための意欲、が必要と考えられる。TETDM においては、この 1) と 2) を学ぶためのチュートリアルを実装している。また 3) のために、チュートリアルをキャラクターとの会話形式で進められるようにするとともに、

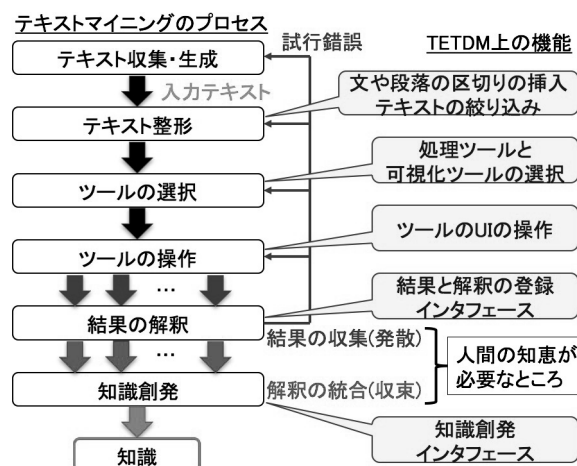


図1 テキストマイニングのプロセス

ソーシャルゲームの見た目で、さまざまなユーザの操作に対応して、経験値やコインを獲得できるゲームモードを実装している。

本稿では、これらの機能について説明した上で、多くの人がデータ分析のスキルを身につけるための、必要な環境や機能、ならびに TETDM の普及の方法について議論する。

テキストマイニングを行うことにより、実社会で使える知識を得るためには、図1に示すテキストマイニングのプロセスを経由する。一般的なテキストマイニングツールでは、単一のツールによる分析を主として、その結果を出力するにとどまっている。しかし TETDM は複数のマイニングツールを備えるとともに、結果の出力で終わるのではなく、試行錯誤によって多くの結果を集め、またそれらを統合する知識創発のプロセスを支援する機能を含んでいる。

Copyright is held by the author(s).

The article has been published without reviewing.

¹TETDM サイト : <http://tetdm.jp>

²本稿作成時には、2017/10/6 公開のバージョン 3.01 が最新版。

2 TETDM

統合環境 TETDM は、あるテキストを入力したときに、その分析処理を行う処理ツールと、処理結果を可視化して出力する可視化ツールを複数備えている。これらのツール群は、単独の開発者によって提供されるものではなく、任意の開発者により自由にツールを追加することができる。

TETDM (図 6) は、独立した複数のパネル内に、処理ツールと可視化ツールを 1 つずつペアとしてセットすることで動作する。各パネルにセットされる異なる開発者によって作成されたツールが、それぞれ独立に動作するだけではなく、それらを協調的に連動させることができる点が TETDM の特徴となっている。

2.1 起動モード

TETDM では、幅広い利用者を対象として、テキストマイニングの専門的知識をもたない人にも手軽に利用してもらえよう、次の 4 つの起動モードを用意している。

- スーパーライトモード：簡単に試したい人向け
- ライトモード：少しだけツールを操作したい人向け
- 通常モード：いろいろなツールを操作したい人向け
- 拡張モード：より詳細な分析をしたい人向け

スーパーライトモードとライトモードは、機能を制限してボタンを減らし、簡単に利用できるモードとなっており、主に趣味での利用を想定している。一方、通常モードと拡張モードは、仕事での利用を想定して、「結果の解釈」と「知識創発」を含む全てのマイニングプロセスを実行できるモードとなっている。

2.2 テキストの入力

テキストの入力は、TETDM 上部に表示されるメニューウインドウ内の「ファイル」ボタンからファイルを読み込む、または、図 6 右端のパネルにセットされている処理ツール「テキストエディタ」に、テキストをカット&ペーストすることで行う。

2.3 ツールの選択と操作

TETDM の各パネルには、処理ツールと可視化ツールを 1 つずつペアとしてセットする。利用者は、利用目的に応じたツールをセットして利用する。

処理ツールは、統合環境内のテキストデータをもとに、テキストの理解や分析に役立つ情報をテキストから抽出する。処理ツールの処理結果は、可視化ツールに渡されて出力される。可視化ツールは、処理ツールによる出力結果を可視化する。この処理ツールと可視化ツールのペアを、各パネルにセットして利用する。

ツールをセットした後は、ツールの操作によってさまざまな処理結果を得る。可視化ツールによってはマウス

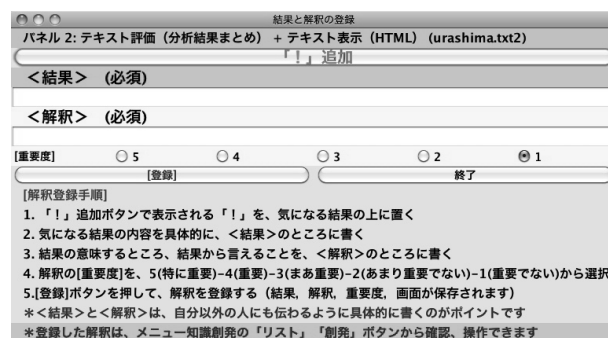


図 2 結果と解釈の登録インタフェース

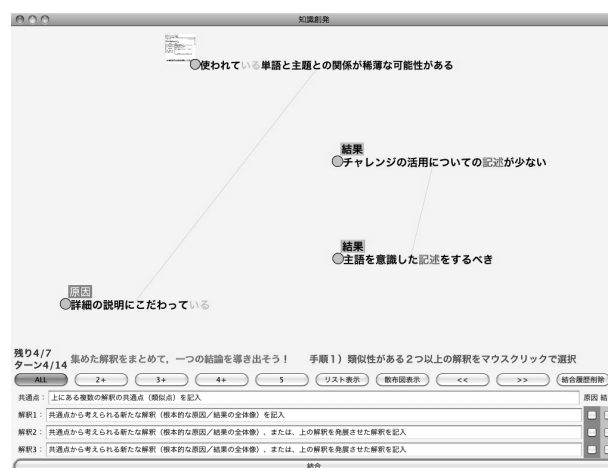


図 3 知識創発インタフェース

クリックやドラッグによって、表示形式を変更、調整した上で、処理ツールの再実行を行うことができる。

2.4 結果の解釈

データの裏に潜む知識を発見するためには、より多くの特徴的な結果を収集し、また記録しておく必要がある。結果を記録するための「結果と解釈の登録」インタフェースを図 2 に示す。利用者はインタフェース内にあるテキストフォームに、出力の中で気になった「結果」と、その結果の意味を「解釈」として記入して登録できる。試行錯誤を重ねる中で、さまざまな結果をこのインタフェースを通じて集めることができる。

2.5 知識創発

登録された「結果と解釈の集合」からわかることを、1 つにまとめた上で、次の行動につながるアイデアを得るための知識創発を支援するインタフェースを図 3 に示す。利用者は、解釈の集合を、原因と結果に分離しながら 1 つにまとめる作業を行うすなわち、類似性をもつ複数の解釈を選択し、それらに共通の解釈を記入し、原因または結果のチェックを入れて解釈の統合を繰り返す。これにより、集めた結果全体が表す因果関係を知識として得ることができる。

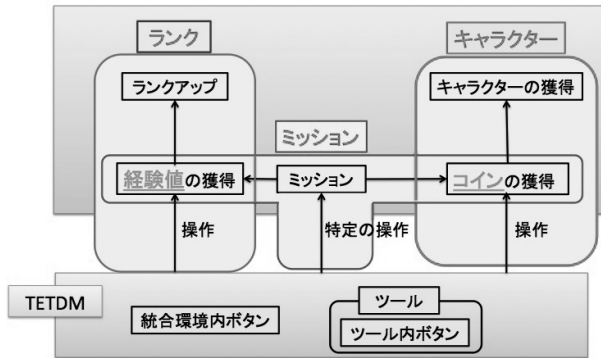


図 4 TETDM ゲームモードの枠組み



図 5 ゲームモード用キャラクター

3 TETDM のゲームモード

本章では、利用意欲向上のために TETDM に実装したゲームモードについて述べる。図 4 に示す枠組みに基づいて、ゲームの要素を TETDM に追加することで、利用意欲の向上を目指す。

ゲームモード内でユーザが獲得を目指す要素として、TETDM を利用することによって得られる「経験値」と「コイン」を用意している。また、この経験値とコインに関わる「ランク」「キャラクター」「ミッション」の 3 つの機能を実装している。すなわち、経験値をより多く獲得することによるランクアップ、コインをより多く獲得することによるキャラクターの獲得、また経験値とコインを入手するための手段としてのミッション、を機能として実装している。以下の各節で、これらの要素と機能の詳細について述べる。

3.1 経験値

経験値は、TETDM 上でテキストマイニングに関わる操作を行った場合に得ることができる。

まず、TETDM 内のボタンを押すなどの操作を行うたびに、1 から 10 の経験値を獲得できる。

次に、TETDM には、利用者がツールの処理結果を見る際に、その結果と解釈を登録できる機能がある。データ分析時に、気になった結果とその意味付けを考えられるようになることを意図して、この機能を用いて結果と解釈を 1 回登録するたびに、5 の経験値が得られるように設定している。

また TETDM には、利用者が登録した解釈を、1 つにまとめてアイデアの生成を促す知識創発の機能がある。データ分析の最終目的は、次の意思決定につながるアイデアを生成することと考えられるため、この機能を用いて解釈を 1 つにまとめた際には、まとめる前の解釈の数の 10 倍の経験値が得られるように設定している。

すなわち、積極的にツールを操作して、出力結果を見て有効な結果を集め、集めた結果を収束させてアイデアの創発につなげてもらうことを意図して設定している。

表 1 ミッションと獲得できるコインと経験値の例

No.	ミッション (獲得コイン, 獲得経験値)
1	ランクを 100 にしよう (5000 コイン)
5	知識創発を 5 回行ってみよう (5000 コイン, 50 経験値)
11	結果と解釈を 50 回登録しよう (5000 コイン, 50 経験値)

3.2 コイン

コインは、TETDM 上でテキストマイニングに関わる操作に限ることなく、より幅広い操作の結果として得ることができる。すなわち、前節で述べた経験値が得られる操作を行った場合には、経験値と同数のコインが得られる。また、3.5 で後述するミッションの達成時には、ミッションに応じて定められたコインが得られる。

3.3 ランク

ランク (またはレベル) は、ゲームユーザのやり込みの度合いを表す数値として用いられている。そこで、ランクを設定することにより、ランクの上昇を目指して積極的に TETDM を利用してもらうことを意図した。ランクは、3.1 で述べた経験値が 100 増加するごとに 1 増えるように設定した。継続的にランクの上昇による満足感を与え、複数回の利用にわたっての利用意欲の向上を目指すために、一定の経験値でランクが上昇するようにしている。この設定には、ランクが線形に増加することで、複数のユーザ間での利用状況が比較しやすくなる利点もある。

3.4 キャラクター

ユーザには、継続的な利用を進める要素として、キャラクターとそのコスチュームを用意した。キャラクターとコスチュームは、獲得したコインで購入することができる。「キャラクターなんとか機」[8]を用いて、全部で 6 種類のキャラクターを用意した (図 5)。

3.5 ミッション

ミッションは、ユーザが TETDM を利用する際の、具体的な利用目標の 1 つになるものとして設定している。ユーザが各ミッションを達成すると、ミッションごとに定められた経験値やコインを獲得することができる。用意したミッションの一例を表 1 に示す。

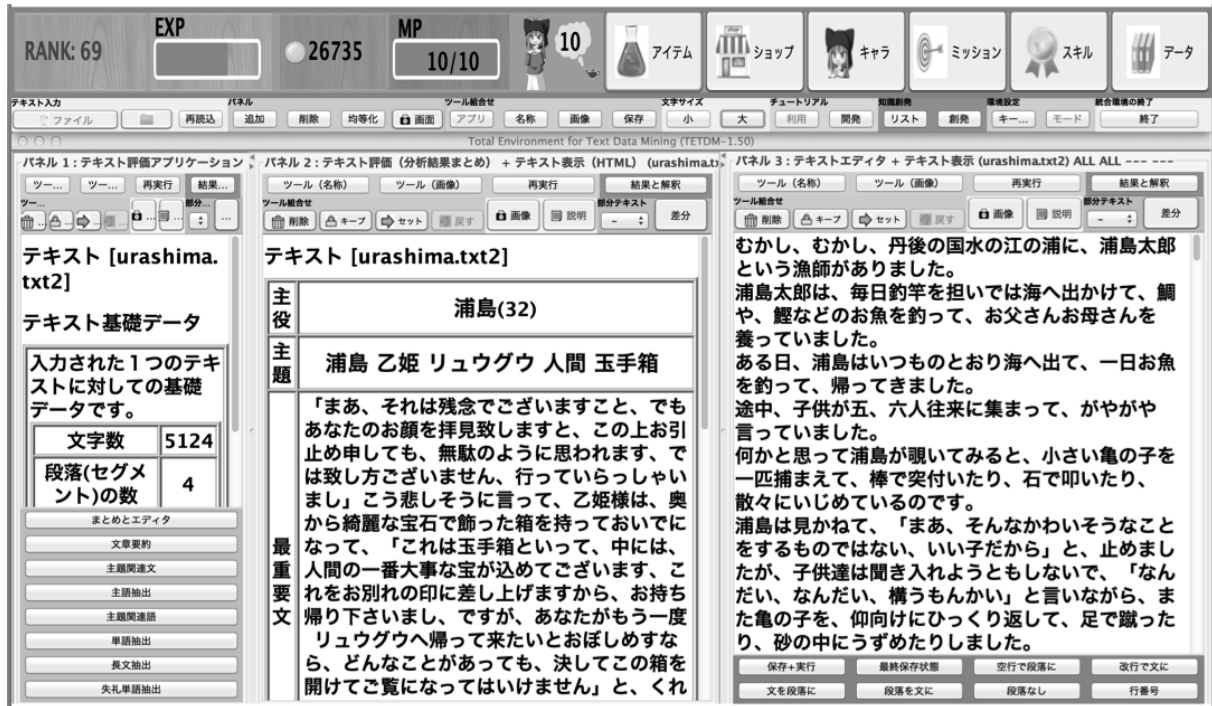


図 6 TETDM の画面とゲームモード用のウィンドウ（画面上部の赤枠部分）

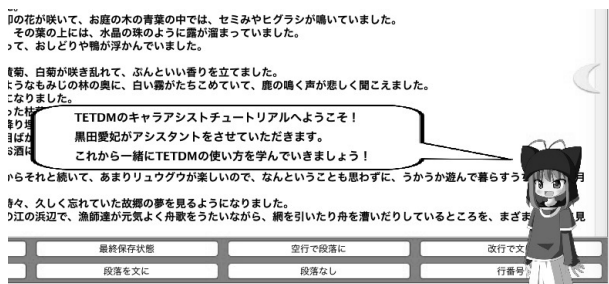


図 7 チュートリアル実行中の画面

4 TETDM のキャラクターアシストチュートリアル

本章では、ユーザのテキストマイニングスキルの獲得を支援する、キャラクターアシストチュートリアルについて述べる。既存の多くのソフトウェアのチュートリアルは、いわゆるヘルプ機能により説明を参照するものであったり、文章による説明を読み続ける形式となっていることが多い。しかし説明は、少しでもユーザの記憶に残りやすく、わかりやすいことが望まれるとともに、説明を聞き続けるための意欲が維持されることが望ましい。そこで TETDM においては、図 7 のようにキャラクターがセリフによって説明を行う形式にするとともに、直感的に操作を体験できる内容のチュートリアルを導入した。

表 2 チュートリアルの説明数（操作方法は、さらに操作の説明（左数字）とツールの説明（右数字）に分かれる）

起動モード	操作手法	マイニングの流れ
スーパーライトモード	11+15	12
ライトモード	9+6	6
通常モード	8+7	9
拡張モード	6+11	9
合計	34+39	36

4.1 チュートリアルによる説明の構成

本節では、本チュートリアルによって説明される内容の構成について述べる。本チュートリアルは、「TETDM の操作方法」と「テキストマイニングの流れ」の 2 つの説明群から構成されている。また「TETDM の操作方法」は、さらに「操作の説明」と「ツールの説明」に分かれている。TETDM では、各パネルに処理ツールと可視化ツールをペアで割り当てて使用するが、それら各ツールの使い方の説明は「ツールの説明」に含まれ、それ以外のツールに依存しない操作については「操作の説明」で説明する。

「テキストマイニングの流れ」では、前者の操作方法をある程度理解している前提で、テキストマイニングのスキル獲得に繋がる説明を行う。また両者には、4 つの起動モードに対応した内容が用意されており、上記の種類と起動モードを選択して、説明を開始する。

チュートリアルの説明数を表 2 に、またスーパーライトモードの「操作の説明」の内容を表 3 に、「マイニング

表 3 スーパーライトモード「操作の説明」の内容

1. はじめに
2.TETDM の概要
3. スーパーライトモード
4.TETDM のウインドウ
5.TETDM のボタン
6.TETDM の用語
7. ゲームモード
8. テキストマイニングの流れ
9. テキストの入力
10. テキストの段落分け
11. ツールの選択

表 4 スーパーライトモード「マイニングの流れ」の内容

1. はじめに
2. テキストマイニングの流れ
3. テキストを入力してみよう
4. テキストを整形してみよう
5. ツールの選択して結果を読み取ろう
6. 「長文抽出」の結果を読み取ろう
7. 「あいまい単語抽出」の結果を読み取ろう
8. 「主題関連文」の結果を読み取ろう
9. 「主題関連語」の結果を読み取ろう
10. 「主語なし文抽出」の結果を読み取ろう
11. 集めた解釈をまとめてみよう
12. 処理結果を画像で保存しよう

の流れ」の内容を表 4 に示す。また、「ツールの説明」の内容としては、各モードで利用できるツール名が並ぶ。

「ツールの説明」に関して、TETDM 内の各ツールは、TETDM のコアメンバー以外によっても開発されることを想定しており、各ツールのチュートリアルはツールの開発者が用意できるようにしている。すなわち、キャラクターに喋らせるセリフを並べたテキストファイルを用意することで、自動的にチュートリアルの説明に追加される仕様となっている。

4.2 チュートリアルの実施方法

チュートリアルによる説明は、メニューウインドウの「キャラ」ボタンを押して、開始することができる。また「ツールの説明」のチュートリアルについては、各パネル内の「キャラ」ボタンから、そのパネルにセットされているツールのチュートリアルを開始することができる。チュートリアルの説明は、先述のゲームモードで用意されたキャラクターのセリフによって行われる。以下で、キャラクターによる説明の詳細について述べる。

4.2.1 セリフの表示

説明文はキャラクタの横の吹き出し内に表示され、1 つの吹き出しにはおよそ 1 行 30 文字のテキストを 3 行以内で表示する³。この表示される吹き出しをクリックすることで、次のセリフに進むことができ、右クリック

によって 1 つ前のセリフに戻ることができる。また、説明をむやみに読み飛ばされないように、一つのセリフを表示してから 0.5 秒以内は、次のセリフに進めないようにしている。加えて、キャラクターとセリフはマウスでドラッグできるとともに、キャラクターの右上に表示される月をマウスで触ることで、ディスプレイ内の左右対称な位置にキャラクターを移動させることができる。これは、後述の課題をこなす際に、キャラクターの表示が邪魔になることを避けるための機能としている。

4.2.2 問題と操作課題

セリフの中に、4 択の問題や、特定の操作を行わないと次のセリフに進めない課題を用意することができる。4 択の問題は、説明内容を理解したかを確認する目的で行い、正解すれば次の説明に進むことができるが、不正解の場合は再度同じ説明を冒頭から聞く必要がある。また特定の操作を要求する課題は、操作内容の確認、ならびに手順を逐次実際に進めてもらうために行う。特定の操作には、特定のボタンを押す、特定のファイルを入力する、パネルに特定のツールをセットする、などがある。

4.2.3 キャラクターの表情変化

先の 4 択の問題に正解したり、特定の操作を実際に行うことができた場合には、キャラクターの顔が笑顔になる。また問題に不正解の場合には、悲しい顔に変化する。操作結果に対するフィードバックをユーザに行うことで、ユーザとキャラクターのインタラクションを促進し、意欲の維持を図る。

4.2.4 指アイコンと枠囲み

キャラクタによる説明時に、特定のボタンやパネルを指し示す際に、ボタンを指アイコンで指し示したり、対象パネルを赤い枠で囲む演出を行う。これは、具体的に見るべきポイントを直感的に理解できるようにするとともに、実際にユーザ自身が操作する際の助けとなることを意図している。

4.2.5 画像と音の出力

現在は、ごく一部でしか用いられていないが、チュートリアルの途中に用意した画像を表示させる機能がある。これによって、言葉では説明しづらい内容を、直感的に理解してもらう助けとすることができる。また次のバージョンにおいては、用意した音声ファイルの音を出力する機能を導入する予定となっている。

4.2.6 チュートリアルの中断

チュートリアルの実行時に、キャラクターをクリックすると、ポーズウインドウが開かれ、説明一覧の画面に戻ったり、チュートリアル自体を終了することができる。

³Mac においては、say コマンドによってセリフを音声で発話する機能の実装を予定している。

表5 各チュートリアルの説明1つあたりの平均所要時間
(分) (操作方法の左は操作の説明, 右はツールの説明)

起動モード	操作方法	マイニング	平均
スーパーライトモード	1.1, 1.6	0.7	1.1
ライトモード	1.5, 0.7	1.1	1.1
通常モード	2.3, 1.1	3.1	2.2
拡張モード	1.5, 1.1	4.5	2.3
平均	1.5, 1.2	2.3	1.7
合計時間 (分)	52, 48	83	183

5 TETDM の活用と普及に向けて

データ分析の専門家でなくても、ツールを用いて分析できる環境として TETDM の構築を進めてきた。これまでも、医師や看護師に TETDM を使ってもらい、電子カルテの特徴を捉えてもらう研究も行っている [4]。しかし、より多くの人に TETDM を活用してもらうためには、ツールの使い方と活用方法を学ぶための環境、ならびに学習意欲を維持する方策が欠かせない。

TETDM には、キャラクターを用いないチュートリアルも古くから実装されている [5]。TETDM は統合環境として、さまざまなツールを追加できることから、研究の場面において TETDM への実装を見越してツールを開発することが多く、その実験評価の際に、TETDM の操作方法を学んでもらう必要がある場面において、まずチュートリアルを実施するという活用を行ってきた。加えて、テキスト分析に関する講義、市民公開講座、AI ツールセミナーなどにおいても、このチュートリアルをベースに実施してきた。今回のキャラクターベースのチュートリアルは、それらの開催時に寄せられた意見と要望を踏まえ、より直感的に理解できるチュートリアルの実現に向けて実装したものとなっている。

テキスト分析のツールを全く触ったことがない工学系の学部3年生9名に対して、テキスト分析の実践というテーマの学生実験を3週間で実施した。内容は、キャラクターアシストチュートリアルを用いて、テキスト分析のツール TETDM の操作方法とデータ分析の手順を学び、自作の作文とサンプルデータを用いて、知識創発までを行う内容とした。その際に、合わせてゲームモードの紹介も行い、興味のある人は使っても良いこととした。

表5に、説明1つあたりの所要時間の平均を示す。全109の説明に対して、その多くは2分以内で終えており、全チュートリアルを約3時間で終えることができていた。実際にはすべてのチュートリアルを一度に終える必要はなく、モードごとや種類ごとに進める場合、それぞれにそれほど大きな時間がかからないことが確認できた。

またチュートリアルの後に、自作の作文、レビュー記事集合、学会論文の概要集合、を分析する知識創発を行ってもらい、全員がツールを使いこなした上で結果を

まとめることができていた。他にも、チュートリアルのみで学ぶことに比べ、ゲーム的な要素が入っていたことから、飽きずに続けることができた、という意見が多く得られた。中には、効率よく経験値を獲得するための裏技を発見するに至った学生もあり、ゲームモードが一定の興味を引くことができたと考えられる。ゲームの要素が教育に効果的であることはこれまでも確認されており、著者らも教育にゲーミフィケーションを応用したシステムを開発している [6]。

データ分析スキルの獲得に向けては、何かを学ぶ時には、まず「真似をすること」が大事と言われていることや、我々も要領がわからない作業をする時には、それをうまくやっている人を見本に、動作を学ぶことを行っている。データ分析においても、分析を上手に行っている人を見本にできることが望ましいと考え、現在 TETDM の操作履歴をサーバ上に集約して、お互いに共有できる環境の構築を進めている [7]。

6 おわりに

テキストマイニングのための統合環境 TETDM に実装した、ゲームモードとキャラクターアシストチュートリアルについて述べた。多くの人がデータ分析のスキルを身につけるための環境としての活用可能性を、引き続き検討していきたいと考えている。

参考文献

- [1] 砂山渡, 高間康史, 徳永秀和, 串間宗夫, 西村和則, 松下光範, 北村侑也: 統合環境 TETDM を用いた社会実践, 人工知能学会論文誌, Vol.32, No.1, NFC-A, pp.1-12, 2017.
- [2] 砂山渡, 竹岡駿, 西村和則: テキストマイニングのための統合環境 TETDM の利用意欲向上のためのゲームモードの開発, 日本知能情報ファジィ学会誌, Vol.29, No.2, pp.558-566, 2017.
- [3] 後藤賢悟, 砂山渡, 川本佳代: キャラクターを用いたテキストマイニングスキルの獲得支援, 第15回人工知能学会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資料, pp.40-47, 2017.
- [4] 高間康史, 串間宗夫, 砂山渡: TETDM を用いた電子カルテ分析支援ツールの開発と実カルテ分析での検証, 人工知能学会論文誌, Vol.30, No.1, pp.372-382, 2015.
- [5] 西原陽子, 中垣内李葉, 川本佳代, 砂山渡: TETDM を用いたテキストマイニングのスキル獲得を支援するためのチュートリアルシステムの開発, 日本知能情報ファジィ学会誌, Vol.27, No.5, pp.771-783, 2015.
- [6] 砂山渡, 高橋麻祐, 川本佳代: ストーリー提供機能とライフ機能を用いた学習意欲の向上と維持の仕組み, 日本知能情報ファジィ学会誌, Vol.29, No.3, pp.501-509, 2017.
- [7] 中江剛士, 砂山渡, 畑中裕司, 小郷原一智: 有効な操作履歴の評価によるテキストマイニングスキルの伝達支援, 第31回人工知能学会全国大会, 2M2-OS-34a-5, 2017.
- [8] キャラクターなんとか機:
(URL) <http://khmix.sakura.ne.jp/download.shtml>