

## TETDMの利用者向けチュートリアルシステムの開発

Development of a Tutorial System for TETDM Users

川本 佳代\*<sup>1</sup>      中垣内 李菜\*<sup>1</sup>      西原 陽子\*<sup>2</sup>      砂山 渡\*<sup>1</sup>  
 Kayo Kawamoto      Rina Nakagochi      Yoko Nishihara      Wataru Sunayama

\*<sup>1</sup>広島市立大学大学院情報科学研究科

Graduate School of Information Sciences, Hiroshima City University

\*<sup>2</sup>立命館大学情報理工学部

College of Information Science and Engineering, Ritsumeikan University

Since TETDM, Total Environment for Text Data Mining, has various text mining tools, it is expected that the environment is utilized in many purposes. On the other hand, if a user can add original tools to the TETDM, the user can construct an environment for his/her smooth analysis cooperating with existing tools. However, it may be difficult for a user to learn the usage or to learn the method to add a tool by his/herself. Therefore, tutorial systems on TETDM is required for many purposes.

In this paper, two tutorial systems for users and a tutorial system for developers that have already implemented on TETDM are described. Tutorial systems that can be implemented on TETDM in future works are also described.

## 1. はじめに

テキストマイニングのための統合環境 TETDM[砂山 14] においては、さまざまなテキストマイニングのツールが用意されており、多様な目的に対しての有効な活用が期待されている。また、TETDM 上に独自のツールを追加することができれば、既存のツールとの連携により、自らの望む分析を行いやすい環境を構築できる。しかし、多くのツールを含む統合環境の使い方や、ツールの作成方法を一から独学で学ぶことは困難であり、それらを支援するチュートリアルシステムの作成が望まれる。

そこで本稿では、現在までに TETDM に実装されている以下の利用者向けのチュートリアル 2 種類と、開発者向けのチュートリアル 1 種類について述べるとともに、今後 TETDM 上に実装されるチュートリアルシステムの展望について述べる。

1. 利用者向け：TETDM の使い方とデータ分析（初級）
2. 利用者向け：文章推敲支援
3. 開発者向け：ツール開発支援（初級）

TETDM に実装されるチュートリアルシステムは大きく分けて、ツールの利用方法を学ぶための利用者向けチュートリアルと、ツールを自ら作成するための開発者向けチュートリアルに分かれる。また利用者向けチュートリアルは、特定の目的に特化しない、TETDM そのものの使い方や、各ツールの利用方法を学ぶためのものと、特定の目的を達成するためのツールの使い方を学ぶものとに分けられる。

1 は、利用目的に依存しない利用者向けのチュートリアルで、特に TETDM とテキストマイニングの初心者を対象として、TETDM の使い方の学習と TETDM を使ったデータ分析を行うためのものとなっている。すなわち、本チュートリアルを利用することによって、TETDM の基本的な操作の習得、ならびに TETDM に含まれる処理ツールと可視化ツールを使った

基本的なデータ分析のための操作方法を学ぶことを目的としている。特に TETDM を初めて利用する初心者にとっては、最初に何をして良いのかわからず、具体的な利用のイメージを持つことができないと考えたことから、本チュートリアルを用意した。現在実装されているものは、TETDM の基本機能の利用方法と各ツールの基本的な利用方法に限られているため、今後より高度な利用方法についてのチュートリアルの追加が検討される。

2 は、利用目的を文章推敲に限定した利用者向けのチュートリアルで、文章を作成しその推敲を行いたい利用者が、TETDM を使った文章推敲の方法を学ぶことができる。すなわち、利用者が文章の作成に際して、さまざまな視点があることを理解し、全体の構造から細部の表現に至るまでの一連の推敲についての考え方と、TETDM を利用した推敲の方法の習得を目的としている。TETDM は文章からさまざまな情報を抽出して分析できる環境として構築しているが、多くのツールが存在するため、それらをどのように区別して利用して行くべきかを利用者が自ら学ぶことは難しいと考えた。そのため、1 つの利用目的に対して、より具体的な項目に分類して系統的にツールを利用できるように本チュートリアルを用意した。今後は、他の利用目的に対してのチュートリアルの追加が検討される。

3 は、TETDM 上で動作する独自のツールを作成するためのチュートリアルで、基本的なツールの作り方を学ぶことができる。すなわち、ツール開発者が TETDM 上にツールを追加する方法、ならびにツール作成に際して必要な仕様を理解してもらうことを目的としている。ツールを作成したいと考えた開発者が、何を調べてどこから手をつけて良いのかわからない可能性が高いと考えた。そのため、具体的な開発のイメージをもたせ、まずは簡単なものでも TETDM 上で動作するツールを作成できた、という達成感をもってもらえることを期待して本チュートリアルを用意した。現在実装されているものは、ごく単純なツールの作り方に限られているため、今後他のツールとの有効な連携を図るためのチュートリアルの追加が検討される。

以下で、現在実装されている各チュートリアルの概要について述べ、これまでの利用実績について述べる。その後、今後実

装が検討されるチュートリアルの内容について述べる。

## 2. TETDM のチュートリアルシステム

チュートリアルシステムは TETDM の機能の一つとして実装される。チュートリアルシステムは、TETDM 上のチュートリアルボタンを押すことで開始することができる。最初に、図 1 のようなメインウィンドウが表示され、画面上部のプルダウンメニューによって、行いたいチュートリアルを選択する。

チュートリアルの選択後、画面下部に並べられた課題を表す宝箱を選択することで、課題に挑戦する。課題は類似するものでまとめられて、各グループには MISSION の番号が付与される。各課題をクリアすることによって、宝箱の中に宝物が表示され、経験値を得ることができ、経験値が貯まるとレベルが 1 つ上がる。各 MISSION の最後の課題をクリアすることにより、次の MISSION の課題を選択できるようになる。宝箱の形式となっているのは、学習意欲を高めるため、ゲーミフィケーションの考え方を導入していることによる。

課題に挑戦すると、図 2 のような課題の詳細ウィンドウが表示される。課題のクリア条件は、次の 3 つのいずれかとなっている。

1. 課題に文字で解答する。
2. TETDM 上で特定の操作を行う。
3. 説明を最後まで読む。

1 は、図 3 のようなウィンドウに問題と解答欄が現れ、正しい解答を入力するとクリアとなる。2 は、指定されたツールをセットしたり、指定されたボタンを押すといった【クリア条件】として書かれている操作を行うとクリアとなる。3 は、表示される説明文を最後まで読む（最後のページまで進む）ことでクリアとなる。

### 2.1 TETDM の使い方とデータ分析

本節では、利用目的に依存しない利用者向けのチュートリアル、すなわち TETDM の使い方と、TETDM を使ったデータ分析を行うためのチュートリアル [中垣内 14] について述べる。MISSION の構成を、表 1 に示す。全部で 8 つのグループに分けられた 46 の課題から構成されている。

本チュートリアルシステムを利用した評価実験により、これまでに以下の結果が得られている。

- 物事の説明能力が高いならば、マイニングスキルが身に付きやすい可能性が高い、
- 物事への取り組みにおいて積極性が高いならば、マイニングスキルが身に付きやすい可能性が高い、
- TETDM の利用経験がない方が、マイニングスキルの伸びが大きい可能性が高い、
- プログラミングが得意、または物事の説明能力が高いことと、チュートリアルの達成所要時間が短いこととの間には相関がある。

### 2.2 文章推敲支援

本節では、利用目的を文章推敲に限定した利用者向けのチュートリアル [中垣内 15] について述べる。MISSION の構成を、表 2 に示す。全部で 9 つのグループに分けられた 36 の課題から構成されている。



図 1: チュートリアルのメインウィンドウ

本チュートリアルの開始に際しては、TETDM の基本的な操作方法を習得していることを前提としており、2.1 で述べたチュートリアルをクリアした、あるいはクリアした利用者と同等の利用スキルをもつ利用者が対象となる。

本チュートリアルにおいては、文章推敲に際して必要な視点によりミッションを構成し、各視点に対して用いられる TETDM 上のツールを利用した分析を行えるようになっている。本チュートリアルシステムを利用した評価実験の結果、本チュートリアルシステムを用いて学習した利用者は、学習後の文章推敲において、多くの MISSION に対応する視点レベルでの文章推敲を行うことができるようになった。

文章推敲という目的に対しては、ソフトウェアを利用しない推敲には限界があり、ソフトウェアなしでは特定の単語の頻度や、単語の出現位置を一望することはできない。しかし、複数のツールを独学で自由に使いこなせるようになるには多くの労力がかかるため、特定の目的に対する有効な利用方法を提示することには意味があると考えられる。

## 3. 開発者向けチュートリアル

本節では、TETDM 上に追加する独自ツールを作成するための開発者向けチュートリアルについて述べる。MISSION の構成を、表 3 に示す。全部で 5 つのグループに分けられた 16 の課題から構成されている。

本チュートリアルの開始に際しては、TETDM の基本的な操作方法を習得していることを前提としており、2.1 で述べたチュートリアルをクリアした、あるいはクリアした利用者と同

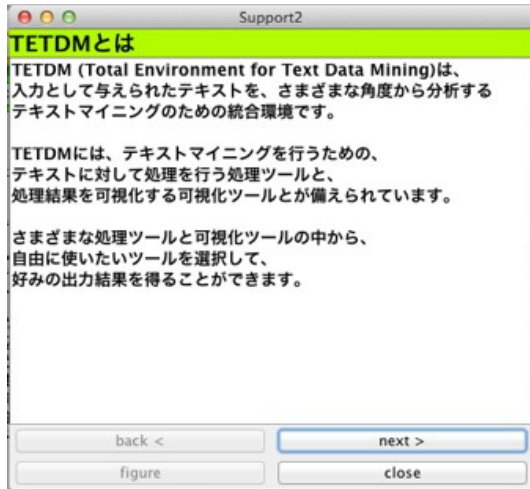


図 2: 課題の詳細ウィンドウ

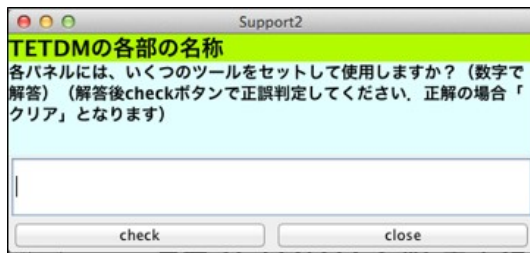


図 3: 課題の解答ウィンドウ

表 1: データ分析チュートリアル課題リスト

MISSION	課題内容
MISSION 0	TETDM の基礎知識: TETDM 内の各部の名称, 用語, 使用の流れ, 使用するテキストの準備, 起動, テキストの読み込み, キーワードの設定に関する基礎知識について合計 8 問.
MISSION 1	ツールの選択とセット: パネルの数の変更, ツールのセット, ツールの説明について合計 6 問.
MISSION 2	ツールの使い方の概要: ツールの操作について合計 6 問.
MISSION 3	結果を表示するツール: 結果を表示するツールについて合計 6 問.
MISSION 4	マイニングツールの操作: マイニングツールの操作について合計 5 問.
MISSION 5	可視化ツールの操作: 可視化ツールの操作について合計 5 問.
MISSION 6	マイニングツールと可視化ツールの操作: 可視化ツールの操作について合計 4 問.
MISSION 7	複数のパネルを使うツール: 複数パネルを利用した解析について合計 5 問.

に述べ 100 名以上の人により利用された。

2.2 の文章推敲のためのチュートリアルシステムは, TETDM のバージョン 1.0 において初めて実装が公開されるものであるため, 実際の利用はこれからとなる. 著者らの研究室においては, 文章の書き方を学ぶゼミを行っており, その中で活用して行くことを想定している.

3. のチュートリアル課題は, TETDM 上でツールを作成することを意図する学生が行ってきた. 先の利用者向けチュートリアルでも述べた, 研究や授業, 演習において TETDM 上で初めてのツール作成を試みる学生がこれまでに広く利用した.

いずれもチュートリアルなしでは, 同じ内容を個別に教える必要が生じ, また教えられる人がいなければ, 実際の利用に結びつけることが難しい可能性がある.

## 5. 実装予定のチュートリアルシステム

### 5.1 利用者向けチュートリアル

目的を特化しない利用者向けチュートリアルシステムについては, 効果的にテキストマイニングのスキルを身につけるための, さまざまな課題を用意していきたいと考えている. すなわち, ツールの利用方法について, 利用者が自分自身で試行錯誤して学習することも重要と考えられるが, 新しいツールの利用に際しては, 見落としや, 気づかない利用方法もあり, またそれらを含む単純な操作の組合せについて, 網羅的に試すことは困難と考えられる. そのため, いくつかの特に有効と考えられる利用方法を提示することには意味がある.

目的指向の利用者向けチュートリアルについては, 現在の文章推敲以外にさまざまな利用目的に特化したチュートリアルを作成することで, 一つの統合環境を, 幅広い目的に利用する手がかりとなることが期待できる.

図 4 に TETDM の利用目的とその関係を示す. TETDM のツールが実現する主な内容として, 情報理解, 知識獲得, 知識整理, 知識創発が挙げられる. 目的を特化しないチュートリアルに, これらを支援する内容を順次追加して行くことを想定し

等の利用スキルをもつ利用者が対象となる.

現在のチュートリアルの内容は, MISSION2 までで説明を行い, MISSION3 と MISSION4 は実際にツールを作成する課題となっているが, 作成する課題の詳細をシステム上で行うことが難しく, TETDM サイト (<http://tetdm.jp>) 上のチュートリアル課題を参照する形となっている. TETDM サイトにおいて, 実装手順が細かく解説されており, サンプルソースをダウンロードすることもできる.

TETDM 上で動作するツールの作成には, 定型的な記述を除けば 100 行未滿のプログラムから実装が可能な仕様となっているが, チュートリアルなしには, どこから手を付けていいのか皆目検討がつかないことも想定される. そのため, まずはチュートリアルに沿って, TETDM 上で動作するツールを一度作成できた, という経験を得るところまでをサポートすることが特に重要と考えられる.

## 4. チュートリアルの利用実績

2.1 のチュートリアルは, TETDM の利用者に, 最初に行ってもらうことでスムーズな TETDM の活用を期待するものとなっている. これまでに, 人工知能学会主催の AI ツールセミナーや, 広島市立大学主催の広島市民向け公開講座において, 本チュートリアルをベースに利用方法の説明が行われた. また, TETDM を利用して研究を行う大学生や大学院生, 研究室内のプログラミング演習ならびに TETDM を利用してツール作成を行う講義の演習課題などで広く利用された. 2.2 や 3. のチュートリアル課題を行う前提となっているため, これまで

表 2: 文章推敲チュートリアル課題リスト

MISSION	課題内容
MISSION 0	このチュートリアルの使い方:チュートリアルを進め方について合計 3 問.
MISSION 1	視点「文章の構成に関する内容」:文章の大まかな雰囲気・構成を掴む練習について合計 6 問.
MISSION 2	視点「全体的な段落間の関係」:文章中の段落の構成について,段落が適切な箇所,バランスで分けられているか,スムーズに読める文章を構成できているかを調べる手段について合計 6 問.
MISSION 3	視点「局所的な段落間の関係」:文章中の一部の段落の構成について,段落が適切な箇所,バランスで分けられているか,段落間のつながりが明確かを調べる手段について合計 3 問.
MISSION 4	視点「段落の表現に関する内容」:文章中の段落中の表現について,段落の主張点が明確か,主張点に対して一貫性があるかを調べる手段について合計 3 問.
MISSION 5	視点「全体的な文間の関係」:文章中の文の構成について,文章の主張点に対して一貫性のある文で,無駄なく構成できているかを調べる手段について合計 3 問.
MISSION 6	視点「局所的な文間の関係」:文章中の一部の文に着目して,文章の内容を把握する手段について合計 2 問.
MISSION 7	視点「文の表現に関する内容」:文章中の文に関して,読者が読みにくかったり,誤解を招く可能性のあるポイントを掴む手段について合計 5 問.
MISSION 8	視点「単語の表現に関する内容」:文章中の単語に関して,表現に問題がないかを調べる手段について合計 5 問.

ている.また具体的な応用に関わる目的を図中の下線で表しており,今回作成したチュートリアルの文章推敲は,最も左の文章作成に対応している.そのためそれ以外の,勉強や学習,スキル獲得,知識の伝達や伝承,創造活動に関わる具体的な目的に限定したチュートリアルシステムを構築して行くことが考えられる.

### 5.2 開発者向けチュートリアル

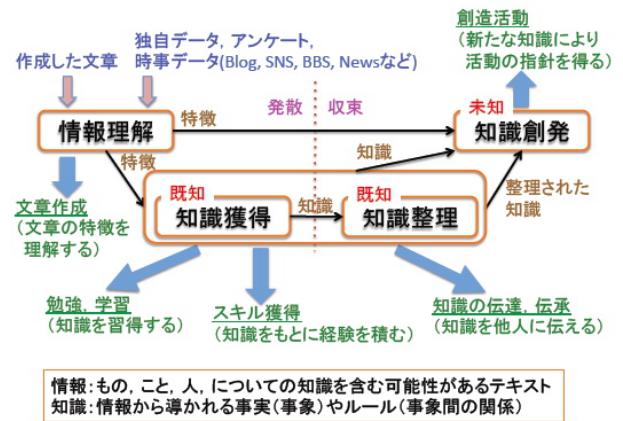
開発者向けチュートリアルについては,現在実装されている内容以外に,TETDM サイト上に,7つのツール作成の課題が用意されている.そのため,これらを含むチュートリアルシステムを実装することが挙げられる.開発者向けチュートリアルにおいては,TETDM の機能を有効に活用するプログラム例に基づいて,特に他の既存ツールとの有効な連携を図るプログラムを作成できるようになることが望まれる.

ツールの組合せを考えることで,複数のツールを統合するツールを作成するなど,ツール作成の幅が大きく広がると考えられる.すなわち,単一のツールの作り方だけでなく,連携して動作する複数のツールをまとめて作ることを含めたチュートリアルの作成が想定される.

また,TETDM に依存しない,一般の JAVA プログラミングや,プログラミングにおけるアルゴリズムの提供などにより,開発を支援することも考えられる.そのような,プログラム作

表 3: 開発支援チュートリアル課題リスト

MISSION	課題内容
MISSION 0	ツール作成の概要について合計 6 問.
MISSION 1	処理ツール作成の概要について合計 4 問.
MISSION 2	可視化ツール作成の概要について合計 4 問.
MISSION 3	処理ツール作成について合計 1 問.
MISSION 4	可視化ツール作成について合計 1 問.



成のための支援環境を整え,それらを活用するためのチュートリアルの作成も考えられる.

## 6. おわりに

利用者向けチュートリアル,ならびに開発者向けチュートリアルについて,これまでに実装した内容とその利用状況,ならびに今後実装を予定する内容について述べた.

さまざまな利用方法が期待される統合環境のチュートリアルは,利用目的に依存する部分と依存しない部分とに分けられるため,それぞれにおいて十分な利用者支援を行って行きたい.また,統合環境上で動作するツールの開発においても,さまざまな利用目的に対応したツール作りを支援できるチュートリアルを実装していきたい.

## 参考文献

- [砂山 14] 砂山渡,高間康史,西原陽子,梶並知記,串間宗夫,徳永秀和:統合環境 TETDM を用いたマイニングツールの開発と利用の実践,人工知能学会論文誌,Vol.29, No.1, pp.100 - 112, (2014)
- [中垣内 15] 中垣内李菜,川本佳代,砂山渡:TETDM を用いた文章推敲スキル育成のためのチュートリアルシステムの開発,第 9 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.29 - 36, (2015)
- [中垣内 14] 中垣内李菜,川本佳代,砂山渡:統合環境 TETDM を用いたテキストマイニング初心者スキル獲得支援,第 28 回人工知能学会全国大会, 1H5-NFC-01b-3, (2014)