

技術史の知識基盤表現作成ツール

A Knowledge Representation Tool for Technology History

安孫子 一敏*・中平 勝子*・三上 喜貴*
ABIKO Kazutoshi・NAKAHIRA T. Katsuko・MIKAMI Yoshiki

知識表現, XML, 年表, 視覚化
Knowledge Representation, XML, Timeline, Visualization

要旨

現代の技術は多岐にわたり専門性も高い。このため、技術の一分野をとってみても全体を見渡すことは難しく、多角的な視点から技術史の分析が行えるようにするためには個々の分野や領域の知識を共有可能とする知識基盤の形成が必要である。本稿では、知識基盤表現ツール作成の第一段階として入力支援ツールの設計と実装を行った。そして、技術史について記された文章を用いて既存エディタ評価と新たに実装した入力支援機能の動作確認を行った。

1. 目的

現代の技術は多岐にわたり、専門性も高い。このため、技術の一分野をとってみても全体を見渡すことは難しい。本研究では多角的に技術史の分析が行えるようにするため、技術史の知識基盤を表現するツールを作成する。実現には歴史事象記述の枠組みの下、入力支援ツール・共有DB・可視化ツールが必要なる。本稿では、知識基盤表現ツール実装の第一段階として、効率的に技術史の入力が行え、且つデータの流通性を保つ手助けをする入力支援ツールの設計と実装を中心に報告する。

2. 知識基盤表現作成ツール

図1に本研究の最終目的である知識基盤表現作成ツールの全体像を示す。ツールは入力支援ツールの部分と可視化ツール部分から構成され、データはDBによって共有される。

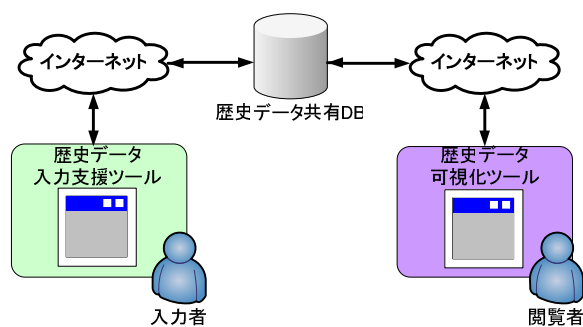


図1 知識基盤表現作成ツール全体像

入力支援ツールは歴史データを入力・編集・保存する機能を持っており、データ入力時にユーザの作業を補助する機能も有している。この機能はユーザのデータ入力を助け、入力されたデータの整合性を保つ目的がある。

可視化ツールは共有DBに保存された歴史データを取得・加工・表示する機能を持っており、DBが保持するデータをユーザが望む形で提示する。

3. 歴史事象の記述方法

歴史事象記述の枠組みは松井文献¹⁾によって提案されたものを採用し、microformats^{注1)}を用いて記述していく。microformatsはXHTMLの属性値に一定の規則に沿った名前付けを行い、文書からメタデータを抽出できるようにするものである。その後、メタデータは歴史の資料情報で汎用形式となっているRDF(Resource Description Framework)へ変換する。

歴史事象記述の枠組み概念モデルを表1に示す。本研究では、表1の定義を基にmicroformatsを用いて歴史事象を記述していく。イベント名は歴史事象の要約を示し、人物は歴史事象に登場する人物を示す。時間は歴史事象の開始した時間と終了した時間を示す。開始・終了時間の正確な時間が分からない場合には、それぞれの時間を大まかな範囲で指定することができる。場所は、歴史事象が起こった地理的場所を示す。場所指定には、過去現在の地名や緯度経度を用いる。原因はその歴史事象が起こる原

*長岡技術科学大学

*Nagaoka University of Technology

因となった歴史事象を指し示す。証拠はその歴史事象の存在を証明する資料を示す。

表 1 概念モデル

要素名	内容
イベント名	歴史事象を端的に表すラベル
時間	歴史事象が起きた時間
場所	歴史事象が発生した場所
人物	歴史事象の登場人物
原因	歴史事象が起きる原因となった歴史事象
証拠	歴史事象の存在を裏付ける証拠

4. 入力支援ツール設計と開発

4.1. 概要

歴史教材入力支援ツールの概要を図2に示す。①ユーザは歴史を記述した自然文のドキュメントを文献やWebから取得する。②取得したドキュメントを入力支援ツールに入力する。③ユーザは入力支援を受けながら歴史事象記述の枠組みに沿ってデータを編集していく。編集のイメージは、図2に示すように定義したタグを用いてキーワードを抽出していく形となる。④歴史事象入力を完了したドキュメントはXML文書として保存する。⑤XML文書をRDFへ変換した後、DBへ登録する。

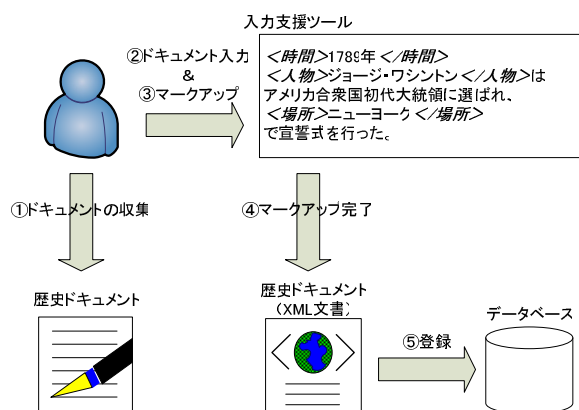


図 2 入力支援ツール概要

4.2. プロトタイプ

既存研究 文献1)において、歴史事象記述の枠組みを検証するために microformats の語彙によるマークアップ機能を追加したエディタが作成されている。エディタは PHP の CMS (Contents Management System) である WordPress 注2) とオープンソースエディタの OpenWYSIWYG 注3) を組み合わせて実装して

いる。本研究ではこのエディタをベースとし、入力支援機能を追加してツールを作成していく。

4.3. 新規開発する入力支援機能

● 人物名補完

自然文では人名を幾度も使用する場合、人物名の省略が行われる。それは文章の読みやすさや書き記し易さを向上させるために行われるが、コンピュータへ認識させる場合には問題が出てくる。姓や名の省略が行われたままにデータの登録が行われると、コンピュータは同一人物を指している場合でも区別できない。これではデータの流通性が阻害されてしまうため、ユーザの入力履歴を提示することで表記の揺れを吸収し人物名を補完する機能を追加する。

● 場所名補完

歴史ドキュメント文章中には過去現在問わず歴史事象が起こった地名が記録されている。その地名は、過去現在で同一である場合と異なる場合がある。この場合、地図上同じ位置に存在していてもデータ上違う場所として登録されてしまう。従って、地理情報を参照して緯度経度の情報も併せて入力できる機能を追加する。

● 年号名変換

過去から現在まで様々な暦が使用されてきた。しかし、歴史事象を記述する際には一意的な基準となる暦が必要となる。その基準として西暦を採用した。その西暦にもユリウス暦・グレゴリオ暦が存在し、暦の採用年も国ごとに異なり一意的な基準とは言い難い。しかし今回は実装を優先させるため、西暦にはグレゴリオ暦を採用し異なる年号を西暦年へ変換する機能を追加する。

4.4. 詳細設計

入力支援ツールのフローを図3に示す。ツールのメイン画面はエディタ編集画面である。基本フローは歴史の自然文をエディタに入力後、文章や単語をドラッグしエディタのボタンクリックによって要素のマークアップをするという流れで行われる。文章や単語の修正が必要な場合には、マークアップ作業中に行われる。

新規に実装された人物名補完・場所名補完・年号名変換の入力支援機能も同様のフローを辿る。異なる点は、DB との通信処理が追加される。文章や単語のドラッグ後、エディタのボタンクリックによ

て DB へクエリを発行する。この時入力支援に必要な情報を DB から取得し、ユーザへの入力支援機能を実現する。人物名補完機能は、入力履歴 DB との非同期通信を基に入力候補をユーザに示す。ユーザは入力候補の中に最適なものがあれば選択し、無ければ修正する。場所名補完機能は、地理情報 DB との非同期通信を基に地理情報をユーザへ提示する。ユーザは地理情報から緯度経度を選択し、地名が誤っていた場合には修正する。年号名変換機能は、年号変換 DB との同期通信を基に年号の自動変換を行う。ユーザは変換された年を確認し、必要があれば修正を行う。マークアップの完了した文書は歴史ドキュメント保存 DB へ保存される。

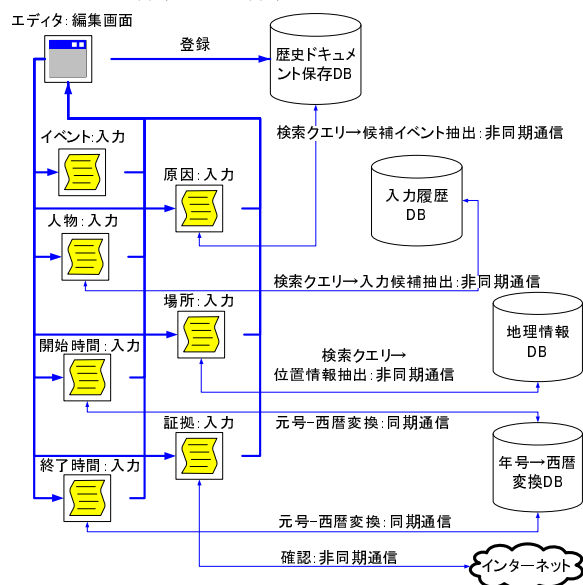


図 3 入力支援ツールのフロー

4.5. 実装

入力支援機能の実装にはクライアントサイドの

技術である Ajax とサーバサイドの技術の PHP を用いた。Ajax を採用した理由は、ベースとなるプロトタイプのエディタ (OpenWYSIWYG) が JavaScript で構成されていて、親和性が高いためである。PHP を採用した理由は、WordPress が PHP で構成されているためである。JavaScript と PHP の連携には HTML_AJAX 注 4) を用いた。HTML_AJAX は PHP PEAR のパッケージで、PHP サーバに存在するクラスをクライアントから実行できる利点を持つ。これを利用してデータをサーバからクライアントへ送信するプロセス、入力履歴・年号変換 DB へのクエリ発行、取得データの加工処理を行っている。地理情報 DB には Google Maps API 注 5) のサービスを利用して、地理情報の入力支援を実現している。

5. 実験

5.1. マークアップ機能評価

作成した入力支援ツールの有効性の評価を行う。評価には国立科学博物館の編成した系統化調査報告 文献 2) を用いて、入力支援ツールによるマークアップ実験を行った (図 4)。図 4 の編集例ではイベント名・人物名・時間の microformats 語彙によるマークアップが行われている。

5.2. 入力支援機能の確認

新規に実装した入力支援機能の確認を行うに当たり、図 5 と図 7 の例には系統化調査報告 文献 2) を使い、図 6 の例としては系統化調査報告 文献 3) を用いた。

人物名補完機能の実行例を図 5 に示す。人物『ルドルフ・ディーゼル (Rudolf Diesel)』が人物名として先にマークアップされている。その後同じ人物

```

<SPAN class=event title=ディーゼル機関運転の成功>
  <SPAN class=dtstart title=1897>1897 年 (明治 30 年) </SPAN>, ドイツ人<SPAN class=person title="ルドルフ・ディーゼル (Rudolf Diesel)">ルドルフ・ディーゼル (Rudolf Diesel) </SPAN>は 5 年にわたる実験の結果, ついにディーゼル機関の運転に成功し世界の注目を浴びた。 </SPAN>
  熱効率の高さと多様な燃料が使用できる点がほかの原動機との大きな違いだった。ディーゼル機関はまず陸上の発電や動力源に, 蒸気機関の代替として利用され, 続いて船舶推進用, 鉄道車両用, 自動車用などに用途が広がっていく。
  <SPAN class=event title=日本へのディーゼル機関導入>
    わが国へディーゼル機関が入ってきたのは<SPAN class=dtstart title=1907>1907 年 (明治 40 年) </SPAN>ころで </SPAN>
  <SPAN class=event title=ディーゼルエンジン国産 1 号>
    国産 1 号は<SPAN class=dtstart title=1917>1917 年 (大正 6 年) </SPAN>
    に誕生したことが記録されている。 </SPAN>
  
```

図 4 入力支援ツールによる編集例

の『ルドルフ・ディーゼル』をマークアップする際には、人物名補完機能が働き入力履歴から候補が提示される。ユーザが提示された候補を選択しマークアップを実行すると文章上では変化はない。しかし、XHTML のメタデータとしては『ルドルフ・ディーゼル (Rudolf Diesel)』が登録される。

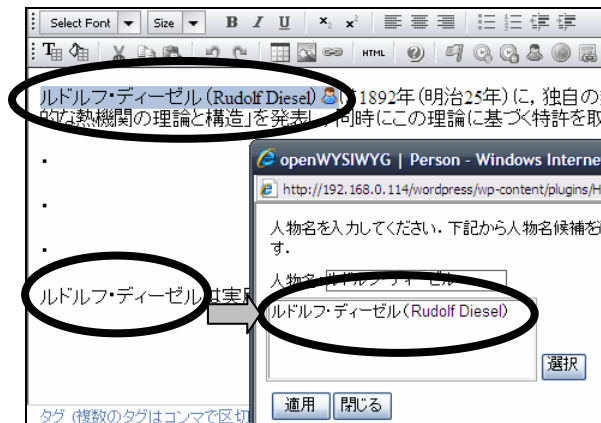


図 5 人物名補完機能の実行例

場所名補完機能の実行例を図 6 に示す。場所名『種子島』をマークアップする際、場所名補完機能が働き Google Maps から『種子島』の地図が表示される。そして、地図をクリックすることで緯度経度が取得できる。取得した緯度経度は場所名と共にメタデータへ登録される。

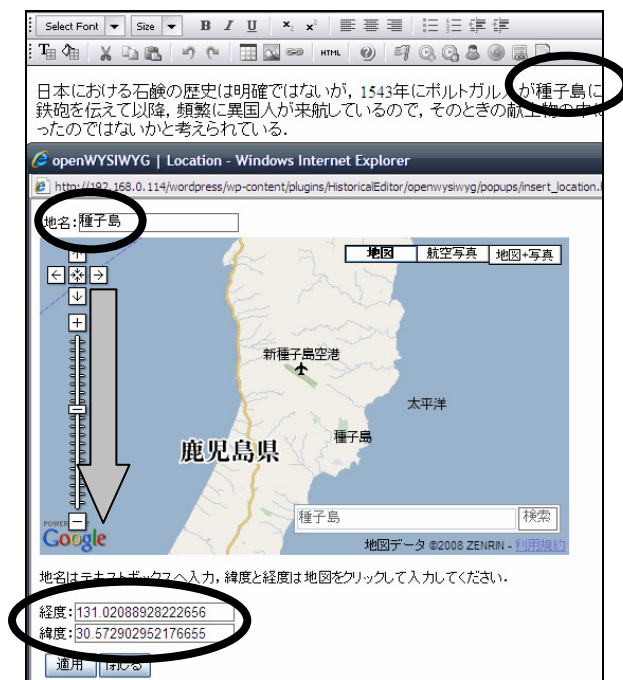


図 6 場所名補完機能の実行例

年号名変換機能の実行例を図 7 に示す。年号『明治 30 年』をマークアップする際、年号名変換機能が働き和暦が西暦へ変換される。メタデータへは西暦で登録されることになる。

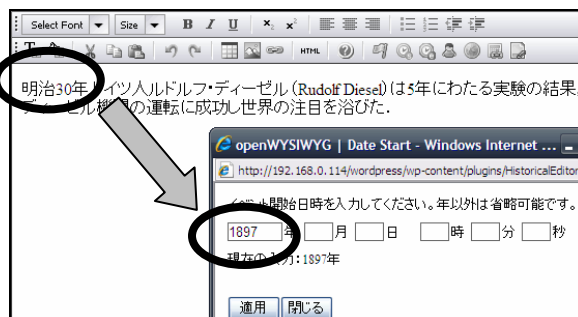


図 7 年号名変換の実行例

6. おわりに

今回、既存のエディタを技術史の文章に適用、マークアップを行いその有用性を調べた。また、エディタの更なる利便性向上を目的に人物名補完機能、場所名補完機能、年号名変換といった入力支援機能の実装を行った。特に年号名変換の入力支援は暦の基準として西暦を採用したが、入力支援機能の有用性検証を急ぐためユリウス暦やグレゴリオ暦を考慮せずに実装した。今後、実用段階においてユリウス暦やグレゴリオ暦といった暦の問題は無視できなくなるため、早期の解決を試みたい。

注

- 1) microformats, <http://microformats.org/>
- 2) WordPress, <http://ja.wordpress.org/>
- 3) OpenWYSIWYG, <http://www.openwebware.com/>
- 4) HTML_AJAX, http://pear.php.net/package/HTML_AJAX/
- 5) Google Maps API, <http://code.google.com/intl/ja/apis/maps/>

参考文献

- 1) 松井正志, 中平勝子, 三上喜貴「XML を用いた技術史の知識基盤表現」『FIT2007 第 6 回情報科学技術フォーラム』, pp. 425-427, 2007 年.
- 2) 佐藤一也「4 サイクルディーゼル機関の技術系統化調査」『国立科学博物館 技術の系統化調査報告 第 12 集』, pp. 1-81, 2008 年.
- 3) 中曽根弓夫「石鹸・合成洗剤の技術発展の系統化調査」『国立科学博物館 技術の系統化調査報告 第 9 集』, pp. 1-58, 2007 年.

(2008年9月30日原稿受理, 2008年11月4日採用決定)