

## テキスト史料の抜粋・分類・俯瞰を支援する歴史学研究支援システム

成瀬聡, 伊東幸宏, 小西達裕, 田村貞雄, 中谷広正  
静岡大学情報学部

本報告では, 対象をテキスト文書に限定し, 史料を電子化して保存すること, 史料の中から歴史学者の独自の視点から関心のある史料を選び出すこと, 選び出した史料を歴史学者の主観に基づいて整理すること, 整理された史料を1つの空間に配置して全体を俯瞰しやすく提示することを支援するシステムの動作を紹介する. なお, 今回のセッションでは, 構築したシステムのデモンストレーションを行う.

### Historical research support system which supports historians in extracting, classifying, and observing historical materials

Satoshi NARUSE, Yukihiro ITOH, Tatsuhiro KONISHI  
Sadao TAMURA, and Hiromasa NAKATANI  
Computer Science, Shizuoka University

Our purpose is constructing a system of supporting researchers in history. In this paper, we illustrate our system that supports historians in constructing text database, extracting historical materials in which they are interested, classifying extracted materials, and observing classified materials. We evaluate the effectiveness of our system by verifying the hypotheses that were set up by a historian.

テムは Java 言語を用いて記述されている.

#### 1. はじめに

本報告では, 史料データベースの検索, 史料の分類, 俯瞰を支援する歴史学研究支援システムを紹介する. 1867年頃, 伊勢神宮などのお札降りを発端として起こった「ええじゃないか」と呼ばれる事件に関する研究と, ほぼ同時期の「コレラ」の流行に関する研究を題材としてシステムの動作を示す.

#### 2. 支援システムの構成 [1]

システムは, 研究者の関心ある史料を抜粋する抜粋部, 抜粋された史料を研究者の主観によって分類する分類部, 分類された結果を研究者が俯瞰しやすいうように提示する俯瞰部の3つの部分で構成される. また, 文献データベースの構築を支援するツールも備えている. 本シ

#### 2.1 文献データベース構築支援ツール

OCR で読み取った歴史史料テキストから, 時間・場所に関する表現を半自動的に抽出し, 史料の時間・場所属性として属性付けられたテキストを HTML 形式に自動変換して史料データベースを構築する. 時間・場所表現の抽出には, 形態素解析ツール「茶釜」[3]を利用して

#### 2.2 抜粋部

抜粋部では, 研究者の関心のある史料を抜粋する. 関心のある史料を抜粋するためには, 歴史学者の関心を何らかの形で表現しておく必要がある. 本システムでは研究者が関心のある史料に多く見うけられる表現のパターンを指定

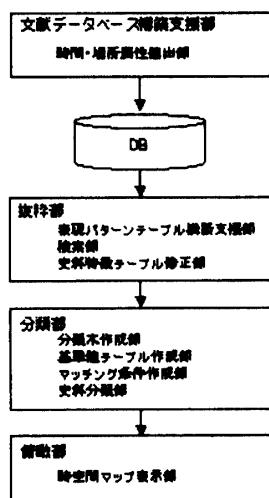


図 1. システム構成

することで自分の関心領域を指定する(図2)。ユーザにより作成された表現パターンテーブル中のパターンを各史料ファイルから検索する際にも、「茶筌」を利用している。

### 2.3 分類部

分類木作成部では、ユーザが分類の階層構造を表す木構造を作成するのを支援する。次に、この木の各ノードが表すグループに史料が属するか否かを判定する基準を作成する。判定基準も抜粋基準と同様に、キーワードの出現に基づく。木がすべて作られた後、選択された表現パターンテーブルで、抜粋されたファイルをトップノードから流し、各ノードに該当する史料ファイル郡が対応付けられた分類木を完成させる。

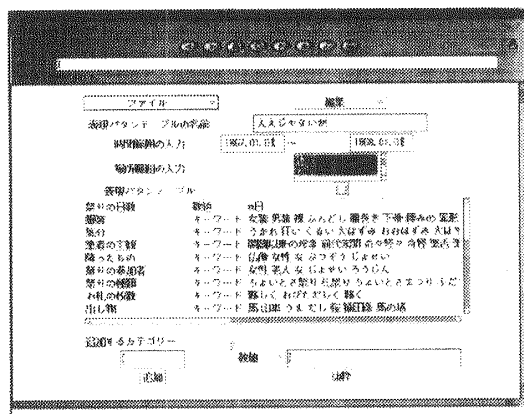


図 2. 表現パターンテーブル構築支援部

### 2.4 俯瞰部

俯瞰部では、分類結果を俯瞰しやすいように時間×地図の3次元空間にマッピングする(図3)。

### 3. 実験結果と評価

本システムの有効性を評価するために歴史学上意義ある仮説検証を行う。本研究では、「ええじゃないか」の広がりに関する研究[2]と、幕末期に猛威を振るった「コレラ」に関する研究に適用を試みた。その結果、歴史学者から仮説の論拠がより明確かつ、客観的に見渡せるようになったとの評価を得た。

### 4. まとめ

歴史学研究を支援するシステムを紹介した。これによって、今まで見えにくかった全体像や例外事象などの発見が容易になることが期待される。今後、更に多様な歴史学研究課題に対して本システムを適用してみる必要がある。

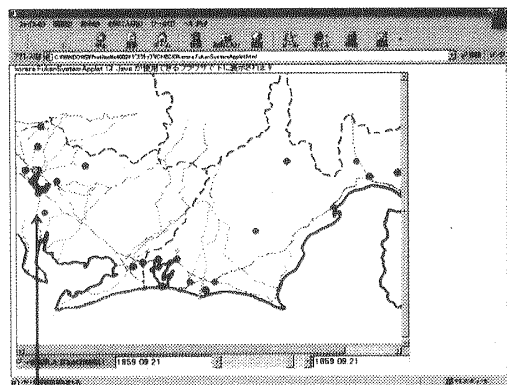


図 3. 時空間マップ表示部

### 参考文献

- [1] 伊東幸宏他:テキスト史料の抜粋・分類機能と分類結果の俯瞰機能による歴史学研究支援, 情報処理学会論文誌, Vol.40, No3 (1999)
- [2] 田村貞雄:ええじゃないか始まる, 青木書店 (1987)
- [3] 松本裕治他:日本語形態素解析システム『茶筌』 version1.0 使用説明書, NAIST Technical Report, NAIST-IS-TR97007 (1997)