

歴史知識学における史料の知識化・構造化フレームワークの構築

赤石, 美奈 / AKAISHI, Mina

(雑誌名 / Journal or Publication Title)

科学研究費助成事業 研究成果報告書

(開始ページ / Start Page)

1

(終了ページ / End Page)

4

(発行年 / Year)

2017-06-23

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 23 日現在

機関番号：32675

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25330269

研究課題名(和文) 歴史知識学における史料の知識化・構造化フレームワークの構築

研究課題名(英文) A Framework for Historical Knowledge Organization

研究代表者

赤石 美奈 (AKAISHI, Mina)

法政大学・情報科学部・教授

研究者番号：60273166

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究では、統合された知識を分解・再編成することで新しい知識を生み出し、「歴史知識学」を発展させるためのシステムの設計・開発を行った。まず、膨大な編年型データを抱える学問分野である歴史学を対象に、歴史学者のニーズや思考プロセスについて考察を進めながら、編年型データ解析の基本システムの設計・開発を行った。次に、概念を表す言葉の意味の変化を捉えるために、概念の体系化を行い、時間変化に伴い動的に変化する概念体系構造を構築した。さらに、解析結果により得られた新たな視点に基づき、史料の構造化を行い、新たな知見を生み出す歴史編纂技術と統合することにより、『史料の知識化・構造化フレームワーク』を構築した。

研究成果の概要(英文)：A vast amount of human experience records is stored all over the world. For understanding and using such chronicles effectively, it is necessary to find several relations embedded in records. In this paper, we propose a chronicle analysis tool (CAT), which provides several views of records from different viewpoints. For such purpose, we introduce the notion of chronicle data and a visual analysis framework for embedded relations in chronicles. On the basis of these notions, we describe a chronicle analysis tool and show a variety of historical human networks from different viewpoints as examples of its implementation.

研究分野：情報科学

キーワード：歴史知識学 視覚化 編年型データ解析

1. 研究開始当初の背景

様々な分野において、時間属性を伴うテキストデータ(政治、経済などのニュース、会議録、ブログなどの編年型データ)が大量に蓄積されており、その有効活用が望まれている。しかしながら、情報を大量に集めてきても、それらに関連付けて総合的な結論や判断に結び付けることは容易ではない。様々な情報を関連付けながら読み解いてゆく必要がある。ところが、問題が複雑であればあるほど、必要な情報の種類や量が増大し、情報同士の関連性を見出すことはますます難しくなる。さらに、既存のマイニング技術は、問題を抱えているユーザ自身が操作するには難しく、ユーザは、解析結果のみを提示されるにすぎない。この結果、既に既知の情報や、解釈できない情報など、ユーザにとって価値の低い情報を得るために時間と労力を費やされる結果となる。そこで、複雑に関連する情報を、単純な関係の重ね合わせと考え、異なる視点から様々な関係を抽出し、複雑な問題を単純な問題に分解し、ユーザ自身が問題について試行錯誤しながら考えを深めていくことを可能とするフレームワークが必要であると考えた。

2. 研究の目的

近年、収集された歴史史料の蓄積・アクセス手法から、それを研究対象化し、最終的に歴史知識として文化社会に定着させるために、歴史学と知識工学を融合させた歴史知識学の必要性が認識されてきている。本研究では、統合された知識を分解し、再編成することで新しい知識を生み出し、「歴史知識学」を発展させていくことを目指し、既存のデータを知識化・構造化することで、新たな知を創出する『編年型データ解析フレームワークの構築』について提案する。

3. 研究の方法

まず、膨大な編年型データを抱える学問分野である歴史学を対象として、歴史史料(大日本史料)を用い、歴史学者のニーズや思考プロセスについて考察を進めながら、編年型データ解析の基本システムの設計・開発を行う。このシステムは、編年型データに対する汎用的な解析手段を提供することを目的とし、(i)データ全体の俯瞰、(ii)分析の着眼点の獲得、(iii)詳細分析という解析プロセスデザインに基づき、データを視覚化し、様々な観点からのデータパターンを提供する。次に、概念を表す言葉の意味の変化を捉えるために、概念構造の変化、概念間の関係の変化を扱う仕組みについて研究を行う。これに基づき、概念の体系化を行い、時間変化に伴い動的に変化する概念体系構造を構築する方法について明らかにすることにより、編年

型データ解析フレームワークを実現する。さらに、解析結果により得られた新たな視点に基づき、史料の構造化を行い、新たな知見を生み出す歴史編纂技術と統合することにより、『資料の知識化・構造化フレームワーク』を構築する。

4. 研究成果

本研究では、大量の編年型データを対象とし、そこに含まれるトピックや付与されたメタデータの経年変化を視覚化することにより、漠然と認識していたことを数値的に確認したり、新たな発見のための気づきを得ることを支援するための解析ツールを開発する。連続する時間に裂け目を入れることにより、本来、存在しない不連続面が生じ、事象に対する様々な見せ方を提示できると考える。そのようなシステムの設計・及び開発を行い、膨大な歴史史料をはじめとして、ニュースやブログ、Twitterなどの様々な編年型データに適用し、システムの有効性について検証した。

(1) 編年型データ解析プロセスモデルの構築

編年型データ解析のための汎用的な手段を提供することを目的とし、(i) データ全体の俯瞰、(ii) 分析の着眼点の獲得、(iii) 着眼点から切り口の獲得、(iv) 詳細分析という解析プロセスデザインに基づき、データを視覚的に解析する編年型データ解析モデルを構築し、システムを開発した。開発される編年型データ解析ツール(CAT: Chronicle Analysis Tool)は、様々な視点から、データを視覚化し、その分布やパターンの変化の発見を促すツールである。具体的には、scope、axis、focus等のパラメータを設定することで、情報の時系列分布状況を俯瞰し、解析の観点を見つけ、詳細分析に至る解析プロセスを支援するシステムである。

(2) 関係分析と関係変化可視化フレームワーク

～ に示すような、歴史史料の分野において有用な各種関係やパターンを定義し、それらの分析・可視化のフレームワークの提案、及びシステムを構築した。

分布パターンに基づく関係分析

図1に、1550年～1650年における“贈答”に関わるデータの分布状況を、“人名”を縦軸、時間を横軸として示す。ここでは、当時の有力武将と天皇との相関関係が分布パターンとして表れている。史料の分布パターンより、贈答関係に着目すると天皇(権威者)と有力武将(権力者)が常にペアとして現れ、“贈答”というキーワードが『権威者と権力者の関係』を示す指標となりうることを示唆している。このような間接的に概念を定量化

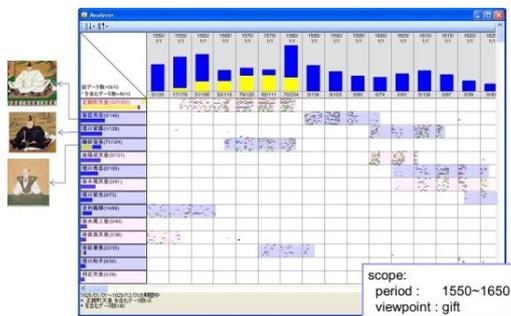


図1 「贈答」に基づく人間関係



図2 「贈答」「戦」に基づく人間関係の変化

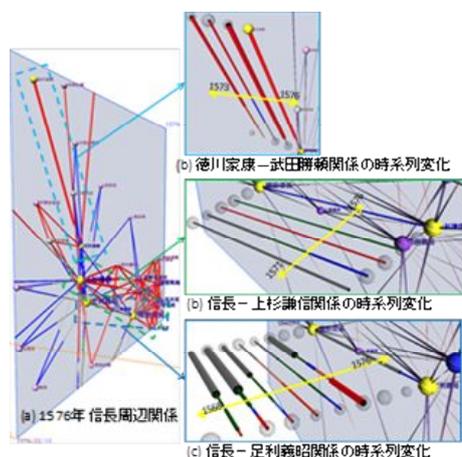


図3 人間関係の時間的変化

するための指標を分布パターンとして見つけるための手法について検討する。

多視点に基づく関係分析

例えば、人間関係のネットワークは、状況に応じて、様々に変化する。この状況として、得られる指標を視点として導入することにより、ある人物を“贈答”という視点や“戦い”という視点からとらえた時の人間ネットワークを抽出することが可能である。図2に、“贈答”と“戦”に関する人間関係の変化を示す。このように情報を構造化するための切り口として視点を導入し、多視点から、関係パターンを視覚化する手法について、検討し、さらに、これを人間以外の概念一般に

適用し、異なる視点からの概念構造ネットワークの抽出、及び視覚化手法について研究した。

時系列変化に基づく関係分析

概念を結び付けている関係の構造変化のみではなく、関係の種類や割合の経年変化を提供することにより、より深い考察が可能となると考える。このため、関係の種類や重要性、クラスタリング手法について検討した。図3は、人間関係の関係そのものが時間により変化する様子を可視化したものである。

(3)概念構造の経年変化

概念構造解析

長期間にわたるデータを扱う場合、社会構造の変化や生活、文化の発展に伴い、概念(意味)の変化や概念構造(他の概念との関係)の変化が生じる。動的に変化する概念や概念構造を扱うダイナミック・オントロジーを構築し、データからのボトムアップに構築される概念の体系化を試みた。

概念構造変化可視化フレームワーク

概念や概念構造の変化を視覚化するためのフレームワークについて研究・開発を行った。連続する時間に多角的に裂け目を入れることで現れてくる様々なパターンを視覚化し、新たな知見や様々な解釈を創発するシステムを開発した。

(4)歴史編纂基盤技術の構築

CATを用いた解析結果により得られた新たな視点に基づき、史料の構造化を行い、新たな知見を生み出す歴史編纂技術と統合することにより、『資料の知識化・構造化フレームワーク』を構築した。

5. 主な発表論文等

〔雑誌論文〕(計 2件)

赤石美奈、石川徹也、歴史資料を解析する-歴史知識学の創成-,人工知能学会誌, vol.31, No.6, pp.758-763, 2016 (査読無)

Mina AKAIISHI, A Meme Media based Application Framework for Context-driven Information Access, Communications in Computer and Information Science, vol.372, pp.1-10, 2013 (査読有)

〔学会発表〕(計 2件)

谷合裕子、赤石美奈、概念関係の時間的構造変化の可視化、第78回情報処理学会、2016.3.11 慶應義塾大学(神奈川県横浜市)

谷合裕子、赤石美奈、概念階層の時間的変化の視覚化、人工知能学会・ことば工学会、2015.12.19 千葉大学(千葉県千葉市)

〔図書〕(計 0件)

〔産業財産権〕

出願状況(計 0件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況(計 0件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

赤石 美奈 (AKAISHI, Mina)
法政大学・情報科学部・教授
研究者番号：60273166

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

堀 浩一 (HORI, Koichi)
東京大学・工学系研究科・教授
研究者番号：40173611

伊藤正彦 (ITO, Masahiko)

東京大学・生産技術研究所・助教
研究者番号：60466422

(4) 研究協力者

()