

一次史料における時間的コンテキストを含む 社会関係記述モデルの提案と実践

小川 潤^{1,a)} 永崎 研宣² 大向 一輝¹

受付日 2021年5月18日, 採録日 2021年11月2日

概要: 歴史研究における知識グラフの活用は史資料のメタ情報記述を中心として徐々に進みつつある。しかし、一次史料の記述内容そのものを知識グラフとして記述する事例は限定的で、歴史研究における史料解釈の段階にまで活用できる構造化データはほとんど見出されない。本研究は、このような史料内容の記述を行うためのデータモデルの検討、およびデータ作成に取り組んだ。題材としたのはカエサル著『ガリア戦記』『内乱記』である。この史料は、カエサルによるガリア戦争および内乱の過程を記述したものであるが、多くの曖昧性を含んでいる。とくに時間情報については月単位、あるいは年単位で曖昧な場合が多く、従来のモデルにおける絶対的時間情報を用いた記述では対応しきれない。それゆえ本研究は、出来事の時系列に基づく相対的時間情報を導入するとともに、特定の時間的コンテキストにおける人物や事物を表現するための EntityInContext という概念を定義することで、時間的曖昧性を含む史料構造化のための手法を提案した。

キーワード: 知識グラフ, RDF, オントロジ, 社会ネットワーク, 古代ローマ史

A Data Model for Social Relationships with Temporal Context in Historical Primary Sources

JUN OGAWA^{1,a)} KIYONORI NAGASAKI² IKKI OHMUKAI¹

Received: May 18, 2021, Accepted: November 2, 2021

Abstract: In the context of historical studies, the use of knowledge graph has recently progressed mainly for describing metadata of primary and secondary sources. The attempt for describing primary-source contents itself, however, is so limited that we could seldom find well-structured dataset which is contributive enough in the process of source interpretation. We, in this article, after careful consideration on how to deal with the complexity of primary sources, suppose the new ontology following which we construct the dataset in practice. The sources we take in this article are Caesar's *Gallic War* and *Civil War*, the problematic sources as they have high degree of ambiguity. In particular, the time information given by the texts is so ambiguous that we could only know approximate season or year for most of the events mentioned. In this case, traditional methods based on the "absolute" date information are no longer valid. Therefore, we suppose a new method for describing the "relative" time information by introducing a new concept "EntityInContext" which, based on the chronological description of events, represents a person or a thing in a certain temporal context.

Keywords: knowledge graph, RDF, ontology, social network, ancient Roman history

¹ 東京大学大学院人文社会系研究科
The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan
² 一般財団法人人文情報学研究所
International Institute for Digital Humanities, Bunkyo,
Tokyo 101-0062, Japan
^{a)} htjk6513khhbk@gmail.com

1. はじめに

デジタル・ヒストリー分野における知識グラフ、あるいは Resource Description Framework (RDF) の活用は各所で進んでおり、史料や書誌のメタデータ記述はもちろん、

史料中に現れる種々の人名・地名といった歴史データにまつわるメタ情報記述にも利用されている。こうした試みは、メタ情報に基づく史料の検索可能性を高め、研究環境の整備に貢献するものであるといえるだろう。

他方、一次史料、とくに叙述史料の記述内容そのものを知識グラフとして表現しようとする試みはいまだ十分に進んでいない。記述内容とは、史料中で言及される行為主体間の関係や時間的コンテキスト、因果関係といった、意味連関を有する内容のことを指すが、実際の史料におけるこれら意味連関の記述は千差万別であるうえに、読解者の解釈による部分も大きいので、メタ情報に比して記述が困難である [1]。

しかし今後、デジタル・ヒストリーが大規模なデータ整備と共有、そしてそれに基づく分析手法構築を志向するなかで、従来のようなメタ情報に加えて、史料記述内容そのもののデータ化、およびそれを用いた歴史分析の実践を視野に入れる必要が生じてくるだろう。これはすなわち、既存のデータベースなど、主に歴史解釈の材料を提供するために構築されたデータに加えて、一次史料の内容を表現し、データ解析に基づく推論などの手法を介して歴史解釈過程そのものに貢献しうるデータを構築していくということである。こうした、史料内容の解釈に活用しうるデータを記述する手段としては、知識グラフによる構造化が有効であると思われる。

こうした試みの一環として、本研究では、カエサルの代表的著作であり古代史研究における一級史料でもある『ガリア戦記』および『内乱記』の知識グラフ化を試みた。データ記述に際しては原典であるラテン語テキスト [2], [3] とともに日本語の全訳 [4], [5] も参照した。この全訳は西洋古典分野では一次史料に対して忠実になされた翻訳として参照されている。記述対象としたのは主に人物や共同体が有した社会関係に関わる言及である。こうした、史料記述から明らかになる行為主体間の社会関係や交流に着目し、その構造的特徴から社会構造を分析する社会ネットワーク分析の手法は、歴史研究においても積極的に導入されている [6], [7]。それゆえ、社会関係に関わる史料の内容情報を知識データとして記述できれば、社会ネットワークを軸とした歴史解釈を大規模データの探索・解析に基づいて行うことが可能となり、歴史研究の実践に貢献することになる。

歴史史料を対象とするデータ構築に際しては、関係自体をいかに記述するかという点はもちろん、それらの関係の特徴や変化をどのように時間的コンテキストの中に位置付けるかが大きな問題となる。

本稿で扱う 2 つの著作は出来事の時系列に沿った叙述形式をとっているが、個々の出来事について必ずしも精確な時間情報が与えられているわけではなく、「冬」「夏」といった漠然とした時期に関する言及のみがなされる場合もある。また、「命令」や「対話」といった一回的な接触に関

しては、時間情報が明示されることはほとんどない。それゆえ、こうした記述を多く含む両著作を歴史研究に活用可能な知識グラフとして表現しようとするならば、時間情報が曖昧、あるいは不明な場合にも時間的コンテキストを十分に表現できる記述モデルを採用する必要がある。

以上をふまえ、本稿は以下の章から構成される。まず 2 章では、知識グラフを用いた歴史データの構築・分析に関わる関連研究を概観し、3 章では既存研究のうちでもとくに重要な Factoid Prosopography Ontology (FPO) のモデルを検討する。4 章において、史料中に言及される行為主体間の関係とその時間的コンテキストを記述可能な構造化モデルを検討したうえで、5 章では、実際の史料において言及される社会関係を知識グラフとして記述する過程を述べる。そして 6 章で、構築したデータの検索性や分析への応用可能性を示したうえで、7 章において本研究の意義および課題を総括する。

2. 関連研究

歴史研究における知識グラフ活用の事例としてよく取り上げられるのが、歴史上の特定の時代・地域における書簡などのやりとりを通じた社会ネットワークについてのデータ構築および分析である。社会ネットワークを扱うこうした研究は、グラフ構造で知識を記述する知識グラフの特性を最も有効に活用できる分野であるといえる。

17 世紀オランダにおける知識人階級が有した書簡ネットワークを扱う CKCC [8] は、書簡テキストに対するテキストマイニングを用いて知識人同士の関係を抽出し、それを知識グラフとして記述している。また類似の試みとしては、16–18 世紀のヨーロッパにおける知識人の書簡ネットワークを扱う Networking Archives や*1、中世の証書データベースである Regesta Imperii に収録された証書群から、自然言語処理に基づく構文解析によって人的ネットワークを抽出し知識グラフ化を試みた研究 [9] などがあげられる。

これらはいずれも、比較的形式の定まった書簡形式の史料を対象とする研究であり、関係抽出のために自然言語処理技術を用いている点で興味深いだが、その正確性や精密性においてなお課題を残している。

このような、書簡史料を対象とした研究のほかにも、様々な類型に属する史資料を対象としたデータ構築の試みがなされている。たとえば、博物館や美術館の所蔵品を歴史的な出来事と関連付けて記述することを企図した Agora [10] は、イベント記述のためのオントロジである Simple Event Model (SEM) [11] を用いてデータ構築を行った。ほかに、戸口調査簿や財務会計史料といった特定の史料類型に特化したモデルを提案し、データ構築を進める研究 [12], [13] も

*1 <https://gtr.ukri.org/projects?ref=AH%2FR014817%2F1>

存在する。これらの研究においては、史料の記述内容そのものに踏み込んだ記述もある程度なされている一方で、叙述史料のような特定の形式を持たない史料への応用に際しては課題が残る。

次に、史料類型ではなく歴史研究の手法論に基づくモデル提案とデータ構築の事例として、人物に注目し、その経歴を記述するバイオグラフィ・プロソポグラフィを取り上げる。

まず近世・近代ヨーロッパを対象とした研究としては、英国教会の聖職者を対象とした CCEd や^{*2}、16–18 世紀ドイツにおける学者層のキャリアを対象とする Professorial Career Patterns (PCP) [14] がある。このうち PCP は、独自のオントロジを定義してデータ構築を行っている。中世については、ビザンツ世界やイングランド、スコットランド、フランク王国などを対象とする多くのプロジェクトが存在し、古代についても、Standard for Networking Ancient Prosopography (SNAP) のようなデータモデル [15]、歴史学におけるプロソポグラフィ研究の成果を、知識グラフを用いて構造化データとして提供する Digital Prosopography of the Roman Republic (DPRR) といったプロジェクトが見出される^{*3}。

ヨーロッパ以外にも、中国史の分野における China Biographical Database (CBDB) など大規模なプロジェクトが存在し^{*4}、注目に値する。

このような研究の多くは独自の方針やモデルを採用しているため、その全容を把握し、相互に比較することは困難である。

一方で、プロソポグラフィ記述に広く適用可能なデータモデルとして、Factoid Prosopography Ontology (FPO) がある^{*5}。このモデルは、上述の西洋中世を対象とする諸プロジェクトで採用されるとともに、古代を対象とする DPRR においても参考にされている。また、一次史料に基づくデータ記述を前提として設計され、必ずしもプロソポグラフィ研究の枠にとどまらない史料内容記述への応用可能性を有する点で、本研究の土台ともなる重要なモデルである。それゆえこの FPO については、章を改めてより詳しく説明する。

3. Factoid Prosopography Ontology (FPO)

FPO は、1990 年代半ばに King's College London のプロジェクトとして発足し、現在ではフランク王国や中世イングランド、東方ビザンツ帝国史など様々な分野において広く採用されている RDF に基づくプロソポグラフィ記述

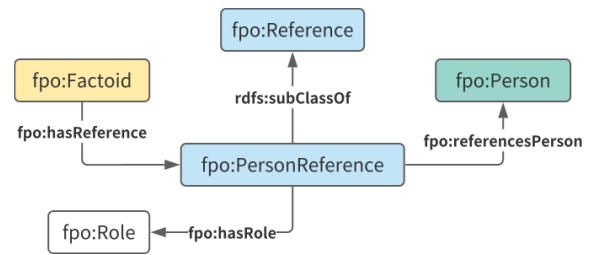


図 1 Factoid モデルにおける Reference 記述
Fig. 1 Description of Reference in FPO.

モデルである [16]。

このモデルの中心概念である Factoid は、一次史料中での人物や組織に関わる何らかの言及を表す。言及の中には地位、身分といった属性的情報とともに、複数の人物や組織の関係が含まれることがある。RDF ではこうした関係をプロパティによって記述することが多いが、FPO では、史料中での関係への言及そのものを Factoid のインスタンス、すなわちリソースとして扱う。そして、このリソースと、言及の対象である個々の人物や組織を表すリソースを関連づけるために Reference というクラス概念を導入し、Reference のインスタンスを通じて相互に連結する。

図 1 では、史料における何らかの関係への言及を表す fpo:Factoid がリソースとして記述され、fpo:hasReference によって、fpo:Reference のサブクラスである fpo:PersonReference を参照している。そして fpo:PersonReference は、普遍的な概念としての人物を表す fpo:Person のインスタンスに紐づけられている。

このように、関係（への言及）そのものをリソースとし、そこに関与する人物や組織への言及を表すリソースへと紐づけることによって、関係そのものに対して時間情報や種別などの情報を記述できるようになると同時に、特定の関係というコンテキストの中で言及される人物や場所に対しても、そのコンテキストに付随する情報を記述できるようになる。たとえば図 1 において fpo:PersonReference は、fpo:hasRole によって fpo:Role を参照しているが、これによって、特定の関係において有した役割によって人物が参照可能になったことを意味する。

関係そのものをリソースとして記述し、その中で果たした役割を軸に人物や組織を表現することのできるモデルは、FPO のほかにも存在する。たとえば、バイオグラフィ記述のためのモデルである Bio CRM [17] では、「ジョン・F・ケネディの婚姻関係」という関係を図 2 のように記述する。

図 2 左端の jfk_marriage が、「jfk の婚姻関係」を表すリソースである。このリソースが「婚姻関係」を表すことは rdf:type によって示され、その関与者として夫ジョン・ケネディ (john_kennedy_in_role_Spouse) と妻ジャクリーン・ケネディ・オナシス (jacqueline_kennedy_onassis_

*2 <https://theclergydatabase.org.uk/>

*3 <https://romanrepublic.ac.uk/>

*4 <https://projects.iq.harvard.edu/cbdb/home>

*5 <https://www.kcl.ac.uk/factoid-prosopography/ontology>

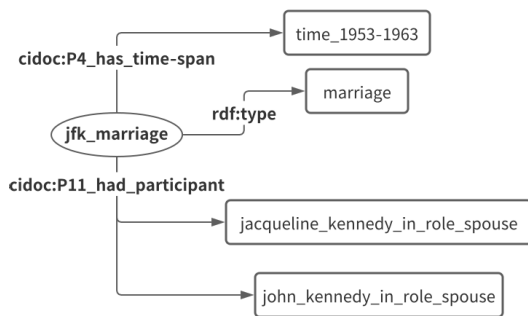


図 2 Bio CRM によるケネディの婚姻関係の記述

Fig. 2 Description of Kennedy's marriage by Bio CRM.

in_role_Spouse) が記述される。

ここで、両者の役割を表現するために、「特定の役割を持つ者としての人物」を1つのリソースとして記述する必要が生じている。このように、関係に関与した人物と、その人物が有した役割が不可分であるモデルでは、1人の人物が同一の関係において複数の役割を有していたり、役割についての解釈が分かれたりしてなどの要因で複数の役割を記述しておきたい場合には、その役割ごとに異なる人物のリソースを記述することになり、データが複雑化する恐れがある。また、一次史料においては、関係の中で人物や組織が果たした役割がつねに明示的に記されているわけではないため、場合によっては役割情報を付与することが困難である。さらには、一次史料中での人物名の表記など、役割以外の情報を記述する際にも問題が生じる。

これに対して、先に述べた FPO では、Factoid と Reference を用いて、まず特定の関係における人物や事物への言及其のものをリソースとして記述し、そのうえで、必要に応じて役割などの付随情報を与えて参照可能にしているため、複数の役割が存在する場合、あるいは役割が不明である場合も同一のモデルで対応することができる。また、役割以外の史料記述に関わる情報を付与することもできる。この点で、FPO は一次史料の構造化モデルとしてより適しており、本研究ではこの FPO を基にしてモデル拡張を行う。

だが、FPO にも依然として課題はある。それは、1 章でも述べた時間的コンテキストの記述に関わる問題である。

FPO における時間記述は、他の既存モデルと同様、Factoid に時間情報を属性として付与するという方法をとる。すなわち、ある出来事や関係が生じ、存在した時間点あるいは期間を、年月日や時分などの数値表現による絶対的時間情報をもって記述するということである。しかし、歴史叙述史料においてはつねにそのような時間情報が与えられるとは限らず、とくに古代史史料においては顕著である。もちろん、正確な時間情報が明らかでないとしても、可能期間 [18] を記述するという選択肢はあるが、そうした期間が年単位でしか確定できない場合も多い。

一方で、複数の出来事の順序については、多くの場合、

史料中に記述されている。時間的コンテキストを考えるにあたっては、これらの出来事の順序を保持することが重要であるが、これは絶対的時間情報のみによっては表現しきれない。そこで、出来事の順序という情報を有効に活用することで、時間的コンテキストの記述を試みたい。そのため本研究では、出来事の順序といった情報も表現可能な形で FPO を拡張し^{*6}、時間的コンテキストを表現可能な新たなモデルを検討、提案する^{*7}。

4. FPO の拡張モデル

本研究で提案する FPO の拡張は、以下の3点からなる。

- ① 出来事の時系列に基づく相対的時間情報の表現
- ② 相対的時間情報に基づく EntityInContext 定義
- ③ EntityInContext を含む関係記述の導入

本研究の意義の多くは②、③によることになるが、その前提として①の相対的時間情報の表現があるので、まずは①について述べる。

4.1 出来事の時系列に基づく相対的時間情報表現

絶対的時間情報に依拠することなく時間的コンテキストを表現可能な記述モデルを考えるならば、それに代わる時間的参照点を設定しなければならない。歴史史料における有用な参照点となりうるのは、史料中で言及される歴史的事象とその継起であろう。

歴史叙述である以上、事象の継起は何らかの形で記述されるものであるし、そもそも、何かを「物語る」という行為が、諸事象を一定の時間的コンテキストの内に配置することによって成り立つ以上、歴史叙述においてはこうした事象継起の記述が絶対的な時間情報と同等の重要性を持つ。

それゆえ提案モデルにおいては、こうした事象継起、より正確には史料中に言及される出来事の時系列を記述することで時間情報を表現する。すなわち、個々の出来事を表すリソース間を、時間的コンテキストを表現するプロパティによって関係づけるということである。具体的には、図 3 のようになる。

図 3 では、mentionedAsPrecedent プロパティによって、主語である KeyEventA が、目的語である KeyEventB に先立って生じたと史料に記述されていることが示される。また、mentionedAsSubsequent プロパティによって、主語の KeyEvent.C が、目的語の KeyEvent.A に続いて生じたと記述されている。2つのリソースが mentionedAsPrecedent で結ばれていれば、当然、逆向きに mentionedAsSubsequent も成立することになるため、これらのプロパティは owl:inverseOf によって制約されることになる。

^{*6} <https://github.com/johnBradley501/FPO/raw/master/fpo.owl>

^{*7} https://github.com/junjun7613/roman_factoid/releases/download/1.2.0/roman_factoid.owl

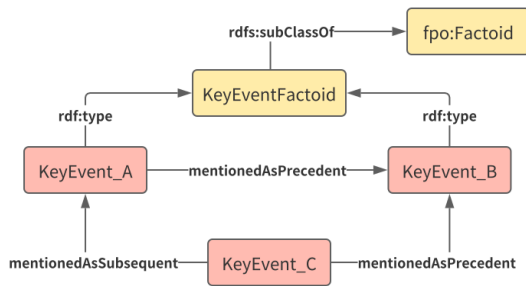


図 3 出来事の時系列記述
Fig. 3 Description of event sequence.

このような出来事の前後関係表現そのものは、CIDOC CRM などの既存モデルによっても提供されており*8, 新規性を有するものではない。だが、FPO をはじめとするバイオグラフィ・プロソグラフィを扱う既存モデルにおいては、こうした時系列記述は採用されておらず、ただ個々の出来事に絶対的時間情報を付与するのみであった。

これにより、個々の出来事に関する時間情報が曖昧な場合や不明な場合にも、史料中で言及される時系列に基づく相対的な時間情報表現が可能となる。

4.2 EntityInContext の定義

史料中で言及される人物や事物は、つねに一定のコンテキストの中でとらえられる必要がある。たとえば、2 人の人物が交渉を行ったという記述が史料中に存在する場合、それぞれの人物が当該時期にどこに滞在していたのか、誰と一緒にいたのか、あるいは、どのような組織に所属し、どのような地位を有していたのかといったコンテキスト情報を考慮する必要がある。それによればはじめて、両者が行った交渉の意義を何らかの意味連関の内に置き、歴史的解釈を施すことが可能になる。

このような時間的コンテキスト情報をふまえた関係記述を可能とするために導入するのが、特定のコンテキストにおける事物を表す EntityInContext である。EntityInContext はサブクラスとして、特定のコンテキストにおける「人物」を表す PersonInContext, 「場所」を表す LocationInContext を持つ。PersonInContext はサブクラスとして GroupInContext を、さらにそのサブクラスとして InstitutionInContext を持つ。このようなクラス階層は、FPO における fpo:Person, fpo:Group, fpo:Institution の階層に対応するものである。加えて提案モデルでは、史料中に現れる都市共同体や部族共同体を表す Community クラスを導入し、fpo:Group のサブクラスとして定義したうえで、それに対応する CommunityInContext クラスを定義している。EntityInContext およびそのサブクラスのインスタンスは、contextualAspectOf プロパティによって、没コンテキストな事物を表す fpo:Person や fpo:Location ク

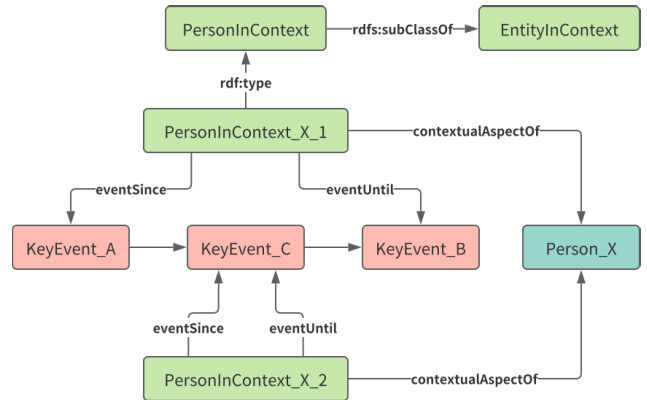


図 4 EntityInContext の記述
Fig. 4 Description of EntityInContext.

ラスのインスタンスを参照する。

EntityInContext は、上述の人物や場所、共同体など異なる要素を包括する上位概念であるため、それぞれを概念的に区別して表現することはできない。それゆえ、それぞれの要素についてのサブクラスを定義した。確かに、たとえば人物についての EntityInContext インスタンスを記述したい場合に、PersonInContext というサブクラスを定義しなくても、EntityInContext と fpo:Person の双方のインスタンスとして記述したり、fpo:PersonReference にリンクさせたりすることで、そのインスタンスが人物に関するものであることを示すことは可能である。しかし、前者については、既存の FPO において、コンテキストを持たない人物を表す fpo:Person のインスタンスと、時間的コンテキストを持つ事物を表すために定義した EntityInContext のインスタンスが同一視されないようにするため、別々のクラスとして定義した。後者についても、リンク先の fpo:Reference のサブクラスによって間接的にのみ EntityInContext インスタンスの種別を特定しうる状況は好ましくない。それゆえ、EntityInContext のサブクラスとして、一定のコンテキストにおける人物や場所といった概念を定義した。

また、EntityInContext のサブクラスを定義することで、それぞれに特有のコンテキスト情報を記述可能になる。たとえば、人物であれば所在地や地位、称号などが問題となる一方、場所であれば領域などの地理情報が問題となるだろうが、人物に固有の情報は PersonInContext に、場所に固有の情報は LocationInContext に記述するよう定義することが可能になる。

EntityInContext のインスタンスは、前節で説明した出来事の時系列記述に基づいて記述される。具体的には、図 4 のようになる。

たとえば、出来事 A から出来事 B の期間という特定の時間的コンテキストにおける人物 X をリソースとして記述したいとすれば、PersonInContext クラスのリソースとして PersonInContext_X_1 を記述する。この PersonInContext_X_1 は eventSince プロパティ

*8 <http://www.cidoc-crm.org/Entity/e5-event/version-7.1.1>

および eventUntil プロパティを用いて KeyEvent_A と KeyEvent_B に紐づけられる。また、PersonInContext は「普遍的なリソース」としての Person の一側面であると見なしうるので、PersonInContext_X.1 は Person_X を参照する。同様に、KeyEvent_C に関連するコンテキストを有する同一人物 X について、PersonInContext_X.2 を記述することもできる。

このように、出来事を軸として事物の状態の時間的変遷を記述するモデルは、これまでも他分野においていくつか提案されており [19], [20], 提案モデルにおいても同様のアプローチを用いている。

このように記述される EntityInContext のインスタンスは、時間的コンテキスト情報を表現するものであるとともに、史料に対するデータ作成者の解釈を表すものでもある。fpo:Reference は人物などについての史料記述そのものから得られる一方、EntityInContext の記述に際しては、データ作成者は史料中で言及される人物や場所、共同体がどのような時間的コンテキストの中でとらえられるのかを解釈し、その結果を1つのリソースとして記述する。既存モデルは、Factoid にも示されるように、史料に記された「事実」の集積を主としてきたが、本研究が提案する EntityInContext という概念は、解釈の記述を扱うことで、歴史分析に有用な時間的コンテキスト情報の表現手法を提供するものである。

4.3 EntityInContext を含む社会関係記述

前節における EntityInContext の導入は、史料中で言及される個々の人物や事物が有する時間的コンテキスト表現を可能にした。これをふまえて本節では、人物や事物の相互作用として現れる社会関係に対しても、時間的コンテキストを与えることを考える。社会関係に対して時間的コンテキスト情報を付与することの利点は交友関係や同盟など、不変ではなく時間によって変化する可能性のある関係、あるいは人物や事物が特定のコンテキストにおいて有した関係を表現できる点にある。

従来のモデルを用いてこうした関係を表現する際には、絶対的時間情報に基づいて記述される期間を関係に対して付与することになるが、すでに述べたように、絶対的時間情報は史料に記載されていない場合が多い。また、交友関係などの持続的な関係のほか、面会や対話などの一回的な社会関係についても、絶対的時間情報が明らかであることは少ない。

こうした問題を解決し、社会関係への時間的コンテキスト付与を実現するために提案モデルでは、従来、コンテキストを持たない「普遍的なリソース」を参照していた Reference リソースに、新たに定義した EntityInContext を参照させる。これを図示すれば、図 5 のようになる。

左端の Factoid_A は、史料中での人物 X および Y への

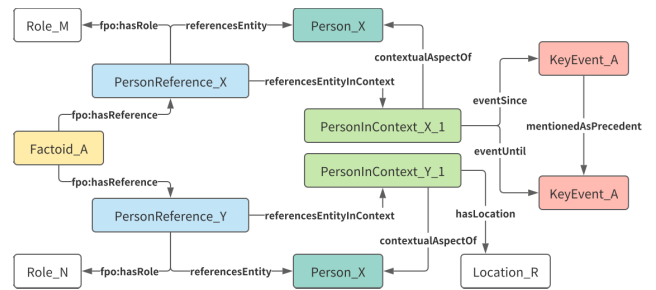


図 5 EntityInContext を用いた社会関係の記述
Fig. 5 Description of social relationship with EntityInContext.

言及を表す PersonReference_X, PersonReference_Y を参照している。FPO では、両者は直接的に Person_X および Person_Y を参照していた。しかし、それらはいずれも時間情報を含まないため、Factoid_A 自体に時間情報が記述されていない場合には、時間的コンテキストの表現は不可能であった。

それに対し提案モデルでは、PersonReference_X, PersonReference_Y をそれぞれ、PersonInContext_X.1, PersonInContext_Y.1 に紐づけることで、PersonInContext リソースが有する、出来事の時系列に基づく時間的コンテキストを Factoid_A に対して間接的に与えることが可能になる。たとえば、PersonInContext_X.1 というリソースは KeyEvent_A から KeyEvent_B の期間をもって定義されているため、Factoid_A が持つ時間的コンテキストも同様の期間を指すと解釈することができる。

また、PersonInContext_Y.1 に対して、この時期における滞在地の情報を Location_R として記述するなど、PersonInContext リソースが特定の時間的コンテキストにおける人物の状態を表現しうることは、前節で示したとおりである。こうした記述によって、Location_R に滞在していた時期の人物 Y が有した関係を抽出するといった、時間的コンテキストに基づく検索も可能になる。

なお、このような記述モデルは、従来のような絶対的時間情報記述と排他的な関係にあるわけではない。Factoid に対して従来の方法に基づいて時間情報を付与することも当然可能である。

5. 一次史料に基づくデータ作成

5.1 『ガリア戦記』『内乱記』について

本研究では、提案モデルを用いて、カエサル著『ガリア戦記』『内乱記』を知識グラフとして記述した。『ガリア戦記』は全 8 巻からなり、紀元前 58 年から同 51 年にかけてガリアを中心として行われたカエサルの軍事活動を、1 年につき 1 巻を費やして編年体で記述した著作である。なお、第 8 巻のみはカエサルではなく、同時代人ヒルティウスの筆になる。一方の『内乱記』はガリア戦争に続くカエサルとポンペイウス間の内戦の顛末を記述している。

すでに述べたように、いずれの著作も時間情報については曖昧な記述が多い。その一方で、編年体記述ゆえに、出来事の時系列については比較的明確に記されており、この点で、本研究が提案する拡張モデルを用いた構造化手法に適している。

実際のデータ作成では、まず史料中から出来事や社会関係について言及している箇所を抽出し、人物や事物についての情報を CSV 形式で入力したうえで、プログラムを用いて Turtle 形式の RDF データに変換するという手順をとった。史料から抽出される関係記述は膨大なため、そのすべてを取り上げることはできないが、以下に史料からいくつかの具体的な記述を引用し、そのデータ記述例を示す。

5.2 第2次ブリタンニア遠征時の社会関係記述

本節では、『ガリア戦記』から2つの社会関係記述を取り上げる。いずれも、カエサルによる第2次ブリタンニア遠征中の関係に言及する記述である*9。

『ガリア戦記』5巻10章2節

…タイントゥス・アトリウスのもとから騎兵がカエサルのもとへ着いた。その報告によると、前夜、大嵐が起り、ほぼ全艦船が被害を受け、海岸に打ち上げられたという。…

『ガリア戦記』5巻11章4節

また、ラビエヌスに書簡を送り、彼の指揮下の軍団兵を使って可能なかぎり数多くの艦船を建造させるようにした。

両者はいずれも、カエサルとその配下の人物との間のやりとり言及するものである。1件目からは、カエサルとともにブリタンニアにわたった副官アトリウスからカエサルへの報告 (report) という関性を、2件目からは、大陸に残っていた副官ラビエヌスに対するカエサルの命令 (command) という関係を見て取ることができる。

カエサルの第2次ブリタンニア遠征は、紀元前54年の7月から9月初めごろにかけてなされたと考えられているが、その正確な日付や個別の関係が生じた日付については明らかでない。それゆえ FPO をはじめとする従来のモデルにおいては、関係を表すリソースに対して、紀元前54年7月～紀元前54年9月といった可能期間を付与することになるであろうが、これでは第2次ブリタンニア遠征の正確な期間を表現することはできない。なぜなら、紀元前54年7月には遠征前の期間が、紀元前54年9月には遠征後の期間が含まれるからである。こうした問題は畢竟、絶対的時間情報が不明であることに起因するが、今回の提案

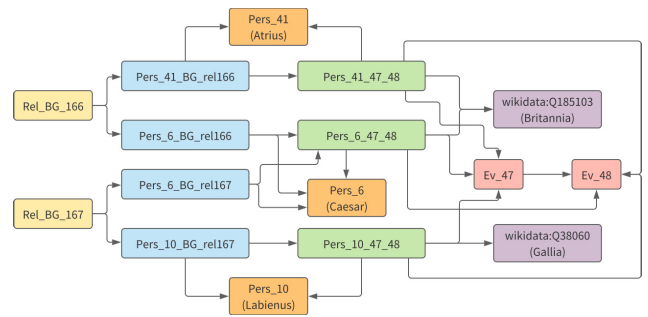


図6 カエサルが有した社会関係データ記述例
Fig. 6 Example data for Caesar's social relationships.

モデルを用いることで、絶対的時間情報は依然として不明ながらも、第2次ブリタンニア遠征という時間的コンテキストの表現は可能になる。

また、2件目のカエサルとラビエヌスの関係についてみれば、両者は当時、異なる場所に滞在していた。このような地理的コンテキストについても、従来のように関係を表す Factoid リソース自体に情報を付与していくモデルでは十分に表現することができない。というのも、両者が異なる場所に滞在していた以上、関係そのものに特定の地理情報を付与することは不可能だからである。だがこのような問題も、両者それぞれの時間的コンテキストを表すリソースを記述し、そのリソースに地理情報を付与することで対処可能である。

図6は、提案モデルに沿ってデータ化された上記2つの社会関係を示している。なお今回は、プロパティ表記は省略した。

図6の左端には、史料から抽出された2つの社会関係を表す Factoid リソース (黄) が記述されている。このリソースは、まず社会関係中に含まれる3人の人物についての計4つの言及を表す PersonReference リソース (青) を参照するが、このリソースは「普遍的なリソース」としての Person (橙) を参照する。ここでは、カエサル (Pers_6)、アトリウス (Pers_41) とラビエヌス (Pers_10) である。

提案モデルではこれに加えて、 PersonInContext リソース (緑) を記述し、 PersonReference から参照させる。たとえば、「カエサルのブリタンニアへの出航 (Ev_47)」から「カエサルのガリアへの帰港 (Ev_48)」の期間におけるカエサルを表す PersonInContext リソースは、 Pers_6_47_48 と記述した。これにより、絶対的時間情報が明らかでなくても、第2次ブリタンニア遠征時の各人物という時間的コンテキストは正確に表現されている。

また個々の PersonInContext リソースには、付加情報として滞在地を示すリソース (紫) などを付与することができる。ここでは、カエサル (Pers_6_47_48) とアトリウス (Pers_41_47_48) を表す2つのリソースはブリタンニア (wikidata:Q185103) を、ラビエヌス (Pers_10_47_48) を表すリソースはガリア (wikidata:Q38060) を目的語とし

*9 邦訳については、文献 [2], [3] の高橋訳を用いた。

て有している。ブリタンニアやガリアといった地理リソースについて、今回はウィキデータを参照したが、より古代史に特化した形での記述を求めるならば、Pleiades^{*10}のような歴史地理典拠データを参照することも可能である。

ここまで、カエサルとラビエヌスについての PersonInContext リソースを例に、滞在地というコンテキスト情報の記述を扱ったが、当然、コンテキスト情報は滞在地に限られるものではない。人物であれば、当時の地位や特定の集団への所属、共同体であれば構成員数や名称などが、時間によって変化するコンテキスト情報である。こうした情報も滞在地で示したのと同じように、EntityInContext,あるいはそのサブクラスのインスタンスに記述することになる。その際には、滞在地の事例のように外部データを用いてもよいし、適当な外部データが存在しなければ、独自に語彙を定義して記述してもよい。

史料中には今回取り上げた2例のほかにも様々な種類の社会関係が記述されている。それらについても Factoid を中心として、出来事の時系列に基づく EntityInContext に紐づける形で時間的コンテキストを表現する点に変わりはなく、基本的には本節で述べた記述手法を適用すればよい。

5.3 データ全体の概要

史料内容の具体的な記述手法は上述のとおりであるが、最後に、現時点でのデータ全体の概要を示しておく。まず現時点でデータ作成が完了しているのは、『ガリア戦記』全巻(1~8巻)および『内乱記』の1巻26節までである。

データ作成にあたってはまず、Community, EntityInContext, Reference, Event, Factoid, Group, Institution, Person, Role のそれぞれの項目について CSV ファイルを作成した。そのうえで史料の精読を通して出来事や関係についての記述を抽出して Factoid を作り、その中に人物や集団、組織が言及されていれば、普遍的なリソースとしての Person や Group, Institution, Community のインスタンスを作成し、Reference を通じて Factoid にリンクする。この際、人物や集団が何らかの役割を有していたと考えられる場合には、Reference のインスタンスに対して、Role のインスタンスを紐づける。また、重要なイベントがあれば、それを Event のインスタンスとして記述し、そこから EntityInContext のインスタンスを作って、先に記述した Reference のインスタンスから参照させる。このように諸々のインスタンスを CSV として記述したうえで、これを Turtle データに自動変換する。

作成した Turtle ファイルを Dydra にアップロードしたところ、総トリプル数は7,721となった^{*11}。リソース数は、Factoid が^s 356, Person が^s 110, Group が^s 6, Institution が^s 1, Community が^s 100, PersonReference が^s 600, GroupRef-

erence が^s 5, CommunityReference が^s 223, PersonInContext が^s 248, GroupInContext が^s 4, CommunityInContext が^s 167, Event が^s 60, Role が^s 39 の、計 1,919 である。作成した Turtle 形式のデータは、ウェブサイト上で公開している^{*12}。

6. データの検索と分析

作成したデータが実際にどのように歴史分析に活用可能であるのかについて、以下ではいくつかの事例を示しつつ、その有用性を示す。1章で述べたように、本研究では社会関係の記述を中心にデータを作成したため、以下で示す事例も基本的には社会ネットワークの抽出・可視化およびその分析となる。

6.1 時間的コンテキストをふまえた SPARQL 検索

作成したデータは、RDF のためのクエリ言語である SPARQL を用いて検索する。以下では、実際に作成データを歴史研究に活用する際に用いられる可能性のある検索条件をいくつか取り上げ、有効にデータを取得できることを示す。

まず検証すべき検索条件は、特定の期間を指定し、その間の社会関係を抽出するというものであろう。歴史研究においてはつねに一定の時間的コンテキストのうちで事象をとらえる必要があり、その意味で、時間的な幅を考慮した検索可能性は不可欠である。そして、その際に問題となる期間は、往々にして絶対的時間情報ではなく、何らかの出来事や事象との関連で設定される。そうした、出来事や事象を軸とした検索は、以下のようなクエリによって実現できる。

```
PREFIX fpo: <https://github.com/johnBradley501/
FPO/raw/master/fpo.owl#>.
PREFIX roman: <https://junjun7613.github.io/
roman_factoid/roman_factoid.owl#>.

SELECT DISTINCT ?relationship ?content
?name_of_entity ?role WHERE {
<Ev_47> roman:mentionedAsPrecedent* ?event.
?event roman:mentionedAsPrecedent* <Ev_48>.
?entityInContext (roman:eventSince | roman:eventUntil)
?event.
FILTER NOT EXISTS {
?entityInContext roman:eventUntil <Ev_47>}
FILTER NOT EXISTS {
?entityInContext roman:eventSince <Ev_48>}
?entityReference roman:referencesEntityInContextex
```

^{*10} <https://pleiades.stoa.org/>

^{*11} <https://dydra.com/junjun7613/roman>

^{*12} https://github.com/junjun7613/roman_factoid/releases/latest


```
?entityInContext.
?relationship fpo:hasReference ?entityReference.
?relationship roman:hasContent ?content.
?relationship fpo:hasReference ?reference.
?reference fpo:hasRole/roman:roleName ?role.
?reference roman:referencesEntity ?entity.
?entity roman:name ?name_of_entity.
}
```

このクエリは、前章で扱ったカエサルによる第2次ブリタンニア遠征という時間的コンテキストにおいて言及される社会関係、その内容、関連する人物・共同体、その役割を抽出するものである。冒頭において、始点となる出来事 (Ev_47) と終点となる出来事 (Ev_48) を与え、その間に存在するすべての出来事を、プロパティパスを用いて取得したうえで、それらの出来事に紐づけられた EntityInContext リソースを抽出している。この際に FILTER NOT EXISTS を用いているのは、Ev_47 を終点とする期間や Ev_48 を始点とする期間を除外し、Ev_47 と Ev_48 の間に「含まれる」EntityInContext リソースのみを抽出するためである。次に、抽出された EntityInContext リソースを参照する fpo:Reference リソースを検索し、それが言及される社会関係記述へと遡っている。

先に述べたように、第2次ブリタンニア遠征中の個々の社会関係そのものに正確な時間情報を与えることは不可能であるが、このようなクエリ検索を行うことで、少なくとも第2次ブリタンニア遠征という時間的コンテキストにおいてはそれらの関係を抽出することが可能になる。上記クエリによる検索結果は表 1 である。

表 1 をみると、抽出した社会関係のうちには、カエサル自身を含むものもあれば、含まないものもあることが分かる。このように、作成データを用いれば、カエサルの行動に基づく時間的コンテキストを活用しつつ、多様な社会関係を検索することができる。さらに、クエリを書き換えれば、カエサル自身を含まない周辺的な社会関係網を抽出するといった複雑な検索を行うことも可能である。また、検索において設定する期間は、クエリ冒頭の始点・終点を書

表 1 上記のクエリ検索結果 (一部抜粋)
Table 1 Results of the query above.

Factoid ID	内容	行為主体	役割
Rel_BG_166	report	Atrius	reporter
		Caesar	reported
Rel_BG_178	command	Britanni	commander
		Camtium	commanded
		Carnutes	commanded
		Grudii	commanded
		Levaci	commanded

き換えることにより容易に変更することができる。

次に、出来事ではなく人物や共同体といった行為主体、とくに特定の時間的コンテキストにおける行為主体を軸に検索を行いたい場合もあるだろう。たとえば、「ガリアを離れ、ブリタンニアに滞在していた時期のカエサルが有した社会関係」を抽出するといった検索である。この場合にも、絶対的時間情報による検索は不可能であるが、提案モデルにおいて定義した EntityInContext を用いれば検索は可能になる。上記の検索を行うためのクエリが以下である。

```
PREFIX fpo: <https://github.com/johnBradley501/
FPO/raw/master/fpo.owl#>.
PREFIX roman: <https://junjun7613.github.io/
roman_factoid/roman_factoid.owl#>.

SELECT DISTINCT ?relationship ?name_of_entity
WHERE {
?entityInContext roman:contextualAspectOf <Pers.6>;
roman:hasLocation <wikidata: Q185103>.
?entityReference roman:referencesEntityInContext
?entityInContext.
?relationship fpo:hasReference ?entityReference.
?relationship fpo:hasReference / roman:referencesEntity
?entity.
?entity roman:name ?name_of_entity.
}
```

後半部の構造は、先述のクエリとまったく同一である。異なるのは前半部で、「普遍的なりソース」としてのカエサル (Pers.6) を参照し、滞在地情報としてブリタンニア (wikidata:Q185103) を持つ EntityInContext リソースをまず検索している。このリソースはすなわち、「ブリタンニアに滞在していた」というコンテキスト情報を持つカエサルを表すものである。あとは、先のクエリ同様、関係を表すリソースへと遡って検索を行う。

6.2 社会ネットワーク分析への応用可能性

前節のクエリ検索によって抽出された社会関係データは、実際の歴史研究に活用される必要がある。そのような活用例の1つとして、以下では社会ネットワーク分析への応用を取り上げる。

図 7 は、第1次ブリタンニア遠征開始から第2次遠征開始に至る時期の社会ネットワーク (左) と、前節表 1 に出力した第2次ブリタンニア遠征時の社会ネットワーク (右) を並置したものである。中央の黄色いノードはカエサルを表しているが、彼を巡るネットワークが、そのアクタ数を含めて大きく異なる様子が見て取れる。このように、異なる時間的コンテキストにおける社会ネットワークであって

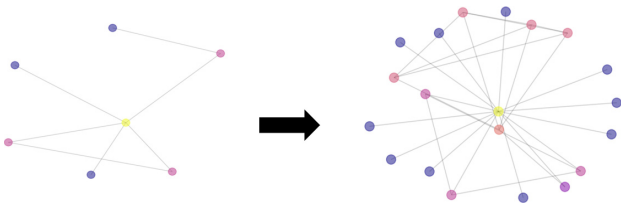


図 7 異なる時間的コンテキストにおける社会ネットワーク
Fig. 7 Social networks in different temporal contexts.

も、クエリ中のリソース指定をわずかに書き換えるのみで抽出することができる。

SPARQL の出力データにおいて、行為主体間の関係は述語によって直接的には表現されておらず、関係そのものを表すリソースを介して間接的に表現されているため、図のような社会ネットワークを描写するためには検索結果を加工する必要がある。今回は、Python のライブラリである SPARQLWrapper を用いて検索結果を JSON 形式で取得したうえで、社会ネットワーク描写に適するようにデータを整形し、NetworkX を用いて可視化・分析を行った。

本稿の目的はあくまで提案モデルと、それに基づいて作成したデータの妥当性を示すことにあるので、ネットワーク分析の詳細には踏み込まない——それゆえ、図 7 にもラベルを付していない——が、図のようなネットワークデータを抽出しえた以上、ネットワーク理論に基づき、時間的コンテキストや社会関係の内容・範囲を柔軟に変化させつつ分析を行うことが可能である。

7. 考察

7.1 手法論的な意義

本研究が一次史料において言及される具体的な行為までを含む社会関係を知識グラフとして記述し、共通規格に則って共有・検索が可能なデータを構築したことは、手法論的に大きな意義を持つ。前章で示したデータ検索および分析事例からも明らかのように、RDF として史料内容を記述することで、SPARQL による柔軟な検索が可能になり、様々なパラメータを操作しての探索的分析が容易になった。

そしてこの探索的分析可能性の範囲は、従来の歴史研究における社会ネットワーク分析の活用において大きな課題でもあった時間的コンテキストの表現にまで及ぶものであり、提案モデルを用いれば、絶対的時間情報に依拠することなく、随意に時間的コンテキストを指定しつつ、ネットワークの時間的変遷を抽出することができる。今回はブリタニアに滞在していた時期のカエサルを例にあげたが、同様にガリアやイタリアに滞在していた時期のカエサル、あるいは他の何らかの時間的コンテキストにおけるカエサルが有した社会関係を抽出し、それぞれのコンテキストにおける社会ネットワークの特徴や変化を分析することも容易である。

こうした、柔軟なコンテキスト情報に基づく検索可能性を

備えたデータは、往々にして曖昧な時間情報を含む歴史一次史料、とくに前近代史料の分析において非常に有用である。

7.2 課題

本研究が手法論的に大きな意義を持つことは上述のとおりであるが、課題もある。確かに、データに基づく探索的社会ネットワーク検索は可能になったが、現時点では、データの量と範囲の限界ゆえに、史料精読によって読み取ることのできる社会ネットワークを再抽出するにとどまっておらず、大規模データ分析に基づく新たな知見の発見という段階には至っていない。

こうした問題を解決し、歴史学における新たな知見の発見を促すようなデータを構築するためには、やはりデータ量 (= 知識量) の拡充が必要であろう。これは、第 1 にはその範囲を複数の史料へと広げ、広範な知識連関を構築すること、第 2 には、社会関係の記述だけでなく、それに関連する数量情報や地理情報、あるいは物体や周辺環境に関する情報などを包括的に記述するモデルを設計することを意味する。

第 1 の点については、本研究でも『ガリア戦記』と『内乱記』という連続する時期を扱う 2 つの史料を対象としており、非常に限定的ではあるものの、史料横断的な Entity-InContext の定義や社会関係の記述が可能となった。今後は、こうした史料横断的な知識グラフ化をさらに広範に進めていくことになる。

第 2 の点については現在、小説の内容記述のために設計されたオントロジ [21] を参考にしたモデル拡張を進めている。このオントロジは小説を対象とするがゆえに、社会関係に限定されない複雑な内容を柔軟に記述するモデルを提供している。一方で、FPO における Factoid や Reference、および本稿で導入した EntityInContext のような概念は定義されていないため、一次史料との対応関係、コンテキスト情報を十分に表現できない。この点で、歴史一次史料記述には不十分である。それゆえ、本稿で提案したモデルを基礎に、小説記述モデルが有する広範な内容記述可能性を取り入れていく。

また、こうしたデータ量の拡充を実現するためには、複数史料間で重複する記述の表現、研究者間で異なる解釈の表現といった問題についても考慮する必要がある。今後、それらの点についての検討を進めていく。

8. おわりに

1 章で述べたように、本研究の最終的な目的は、歴史研究における史料解釈の段階にまで貢献しうるデータを構築することである。そのためには、史料中に現れる人物や事物の関係はもちろん、時間をはじめとする様々なコンテキスト情報も知識として記述しなければならない。

本稿は、具体的な対象としては人物や共同体の社会関係

を知識グラフとして記述するにとどまらなかったが、時間的コンテキストの中で物事を記述するという、歴史史料を構造化するためのモデルの骨子ともいえる部分を設計することができた。以後は、モデルのさらなる改良とデータの拡充を着実に進め、歴史研究に資するデータの構築を実現したい。

参考文献

- [1] Meroño-Peñuela, A., Ashkpour, A., van Erp, M., Mandemakers, K., Breure, L., Scharnhorst, A., Schlobach, S. and van Harmelen, F.: Semantic technologies in Historical Research: A Survey, *Semantic Web*, Vol.6, No.6 (2015).
- [2] Caesar: *The Gallic War*, translated by H.J. Edwards, Loeb Classical Library 72, Cambridge (1917).
- [3] Caesar: *Civil War*, edited and translated by Cynthia Damon, Loeb Classical Library 39, Cambridge (2016).
- [4] カエサル, 高橋宏幸 (訳): ガリア戦記, 岩波書店 (2015).
- [5] カエサル, 高橋宏幸 (訳): 内乱記, 岩波書店 (2015).
- [6] 小川 潤: 歴史研究における社会ネットワーク分析の活用と可能性, *西洋史学*, Vol.269, pp.81–84 (2020).
- [7] Broekaert, W., Köstner, E. and Rollinger, C. (Eds.): The Ties That bind Ancient Politics and Network Analysis, *Journal of Historical Network Research*, 2020-4 (2020).
- [8] Ravenek, W., van den Heuvel, C. and Gerritsen, G.: The ePistolarium: Origins and Techniques, *CLARIN in the Low Countries*, Odiijk, J. and van Hessen, A., Ubiquity Press Limited, pp.309–316 (2017).
- [9] Opitz, J., Born, L. and Nastase, V.: Induction of a Large-Scale Knowledge Graph from the *Regesta Imperii*, *Proc. 2nd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritages, Social Sciences, Humanities and Literature*, pp.159–168 (2018).
- [10] van den Akker, C., Legêne, S., van Erp, M., Aroyo, L., Segers, R., van der Meji, L., van Ossenbruggen, J., Schreiber, G., Wiealinga, B., Oomen, J. and Jacobs, G.: Digital hermeneutics: Agora and the online understanding of cultural heritage, *Proc. 3rd International Web Science Conference*, pp.1–7 (2011).
- [11] van Hage, W.R. and Ceolin, D.: The Simple Event Model, In van de Laar, P., Tretmans, J. and Borth, M. (Eds.): *Situation Awareness with Systems of Systems*, Springer, pp.5–25 (2013).
- [12] Ashkpour, A., Meroño-Peñuela, A. and Mandemakers, K.: The Aggregate Dutch Historical Censuses: Harmonization and RDF, *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, Vol.48, No.4, pp.230–245 (2015).
- [13] 小風尚樹, 中村 覚, 永崎研宣: 構造化された財務記録史料データの分析手法の開発: イギリスの船舶解体業を事例に, *じんもんこん 2019 論文集*, pp.183–190 (2019).
- [14] Blanke, J. and Riechert, T.: Towards an RDF Knowledge Graph of Scholars from Early Modern History, E-published at arXiv:2009.06337v1[cs.DL] (2020).
- [15] Bodard, G., Cayless, H., Depauw, M., Isaksen, L., Lawrence, F. and Rahtz, S.: Standards for Networking Ancient Person Data: Digital Approach to Problems in Prosopographical Space, *Digital Classics Online*, Bd.3–2 (2017).
- [16] Pasin, M. and Bradley, J.: Factoid-based Prosopography and Computer Ontologies: Towards an Integrated Approach, *Digital Scholarship in the Humanities*, Vol.30, No.1, pp.86–97 (2015).
- [17] Tuominen, J., Hyvönen, E. and Leskinen, P.: Bio CRM: A Data Model for Representing Biographical Data for Prosopographical Research, *Proc. 2nd Conference on Biographical Data in a Digital World 2017*, pp.59–66 (2017).
- [18] 関野 樹: Linked Data におけるあいまいな時間の記述, *じんもんこん 2018 論文集*, pp.303–308 (2018).
- [19] Chawuthai, R., Takeda, H., Wuwongse, V. and Jinbo, U.: Presenting and preserving the change in taxonomic knowledge for linked data, *Semantic Web*, Vol.7, pp.589–616 (2016).
- [20] Ide, N. and Woolner, D.: Historical Ontologies, *Words and Intelligence II: Text, Speech and Language Technology*, Ahmad, K., Brewster, C. and Stevenson, M. (Eds.), Vol.36, Chapter 7, pp.1–18, Springer (2007).
- [21] 川村隆浩, 江上周作, 田村光太郎, 外園康智, 鶴飼孝典, 小柳祐介, 岡嶋成司, 村上勝彦, 高松邦彦, 杉浦あおい, 白松 俊, 張 翔宇, 古崎晃司: 第1回ナレッジグラフ推論チャレンジ 2018 開催報告: 説明性のある人工知能システムを目指して, *人工知能*, Vol.34, No.3, pp.396–412 (2019).



小川 潤 (学生会員)

2017年早稲田大学国際教養学部卒業。2019年東京大学大学院人文社会系研究科博士前期課程修了。同研究科博士後期課程在学中。専門は古代ローマ史。知識グラフ・RDFを用いた歴史データの構造化に関心を持つ。



永崎 研宣 (正会員)

一般財団法人人文情報学研究所主席研究員。2000年筑波大学大学院博士課程哲学・思想研究科満期退学。博士(文化交渉学)。東京外国語大学アジア・アフリカ言語文化研究所、山口県立大学を経て、一般財団法人人文情報学研究所の設立に参画し、現在に至る。人文情報学、仏教学の研究に従事。



大向 一輝 (正会員)

東京大学大学院人文社会系研究科准教授。2005年総合研究大学院大学複合科学研究科博士後期課程修了。博士(情報学)。同年国立情報学研究所助手、2007年同助教、2009年同准教授を経て、2019年より現職。人文情報学、ウェブ情報学、学術コミュニケーションの研究に従事。