

時間名による時間参照基盤の構築 —Linked Data を用いた期間の記述とリソース化

関野 樹（国際日本文化研究センター）

時間の参照には、年月日だけでなく、「江戸中期」などの時代区分、「戦後」などのできごとに基づく時代名などがしばしば用いられる。本研究では、これらの時間を参照するためのさまざまな名称を時間名リソースとしてデータ化し、それらが表す期間を Linked Data 技術を通じて取得する仕組みを構築した。

Time reference platform based on time names - Description and definition of time name resource for Linked Data Tatsuki Sekino (International Research Center for Japanese Studies)

Time names such as historical periods (e.g., “mid Edo period”) and names based on an event (e.g., “after the World War II”) are often used for time reference instead of calendar periods represented by dates, months and years. This study described these time names as resources for Linked Data, and examined procedure to decide the time period indicated by the time names, based on Linked Data mechanism.

1. はじめに

時間を参照するためのデータとして、暦法に基づく年月日の代わりに「江戸中期」、「縄文前期」などの時代区分や、「戦後」、「バブル期」といった時代を表す名称、さらには、建築物や人物・団体などの存在期間や履歴（例：「〇〇氏の在任中」）がしばしば用いられる。これらの時間参照として用いられるさまざまな用語（以下、総じて「時間名」と称す）は、歴史や考古をはじめとする人文分野のデータでも幅広く利用されている。

ところが、研究分野や組織の間で時間名の定義や用法は統一されておらず、分野や組織の枠を越え、同じ名称の時間名であったとしても、それらは異なる期間を指している可能性がある。結果として、時間名の使用は、分野や組織の枠を超えたデータ連携の障害となってきた。

こうした問題に対して、「時間名ではなく年月日を用いるべきである」といった意見も少なくない。しかしながら、時間名の多くは、何らかのできごと（為政者の交代やエポック・メイキングなど）や記述対象の変化（土器、絵画、建築の様式の変化など）といった、時間以外の根拠に基づいて定義されており、記述対象を時間軸に沿って区分または分類する役割も担っている。したがって、一概に時間名を年月日に置き換えることは、適切ではない。様々な時間名を識別し、それらを時間軸上で相互に比較するための仕組みが必要である。

そこで本研究では、これらの時間名が表す期間を、その定義に沿いながら時間名同士の関係として記述し、得られたデータ（時間名リソース）を

Linked Data 技術を通じて公開・利用するための枠組みを構築する。

2. 時間名が表す期間の表現

時間名が表す期間の根拠となるイベントや記述対象の変化は、漸進的に起こることが少なくない。このため、時間名が表す期間の始点や終点は、幅を持ったあいまいな値にならざるを得ない。本研究では、このような端点があいまいな期間を扱う新たな理論と手法[1]を用いることで、時間名が表す期間の適切な表現を試みる。

2-1. あいまいな期間の表現

本研究が用いるあいまいな期間の論理では、始点と終点の時間範囲は、それぞれ、時間点 Pb から Rb、および、時間点 Re から Pe で表される（図1）。これらの始点と終点の時間範囲は、それぞれ別の期間を使って表すことも可能である。つまり、始点の範囲が期間 B、終点の範囲が期間 C である期間 A を定義することができる。さらに、期間 B

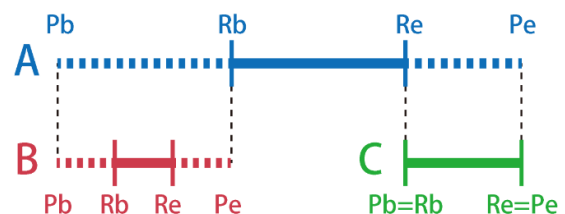


図1 あいまいな期間の概念。始点および終点の範囲（点線）を別のあいまいな期間で指定できる。

Figure 1 Concept of uncertain time interval. Ranges of the beginning and end points (dotted lines) can be specified by other uncertain time intervals.

やCもまた、別の期間で定義されたあいまいな期間であっても構わない。この場合は、期間BやCの時点PbからPeが期間Aの始点や終点の時間範囲となる。期間を使って新たなあいまいな期間を再帰的に定義できる性質は、この論理の重要な特徴である。これにより、時間名が表す期間を、定義に沿いながら時間名同士の関係として記述することが可能となる。

2-2. RDF データ

本研究では、時間名が表す期間や関連する情報を Resource Description Framework (RDF) データ [2]として記述する。これにより、時間名に関する情報を、Linked Data 技術により相互利用可能な時間名リソースとして扱うことができる [3]。これらの時間名リソースは、Unified Resource Identifier (URI) [4]により識別されるため、どの分野、または、どの組織の定義による時間名かを明確に区別できる。これにより、旧来より問題となってきた同じ名称の時間名同士の混乱も、解消が期待できる。

あいまいな期間を RDF データとして記述するために、HuTime オントロジが構築されており [5]、そのスキーマが公開されている (<http://resource.hutime.org/ontology/>)。この HuTime オントロジは、W3C 勧告の時間オントロジである OWL-Time [6]の拡張として設計されている (図2)。

あいまいな期間は、クラス `hutime:UncertainTimeInterval` (プレフィックス `hutime:`は、<http://resource.hutime.org/ontology/>を表す) のインスタンスとして記述される。あいまいな期間を記述するための4つの時点 (Pb, Rb, Re, および, Pe) を指定する方法は2通りある (図2)。Pb点を例にすると、1つは、`hutime:hasPossibleBeginning` を介して、OWL-Time の時点 (`time:Instant`) (プレフィックス `time:`は、

<http://www.w3.org/2006/time#>を表す) または期間 (`time:Interval`) のインスタンスを用いて指定する方法である。もう一つは、プロパティ `hutime:hasPossibleJdBeginning` を介して、ユリウス通日 (double 型のリテラル値) で指定する方法である。ユリウス通日は、紀元前 4713 年 1 月 1 日正午からの通算日数であり、連続した実数で時間軸上の位置を表現できるため [7]、可視化や比較などの処理が容易である。この利点を生かして、時間情報システム HuTime でも時間の内部表現としてユリウス通日が用いられている [8]。

4つの時点の代わりに、始点または終点の範囲を別の期間を使って記述することもできる。この場合は、プロパティ `hutime:hasRangeOfBeginning` と `hutime:hasRangeOfEnd` を介して、あいまいな期間 (`hutime:UncertainTimeInterval`) または、OWL-Time の期間 (`time:ProperInterval` など) で指定する。

2-3. 暦期間の参照

HuTime 暦 Linked Open Data (暦 LOD) [9]は、さまざまな暦の年号、暦年、暦月、暦日を RDF リソースとして提供する仕組みであり、そのデータは、HuTime オントロジに則って記述されている。各リソースは、あいまいな期間を表す `hutime:UncertainTimeInterval` クラスのサブクラスとして定義されているため、新たに定義されるあいまいな期間の始点または終点の範囲として、暦 LOD のリソースを指定することができる。

また、暦 LOD の各リソースでは、その時間範囲の始点と終点のユリウス通日の値を `hutime:hasPossibleJdBeginning` 等のプロパティを介して取得することができる。したがって、時間名リソース間のリンクをその定義に沿って辿り、最終的に暦 LOD のリソースが得られれば、その時間名が表す期間の始点と終点に関する4つのユリウス通日の値を取得することができる。

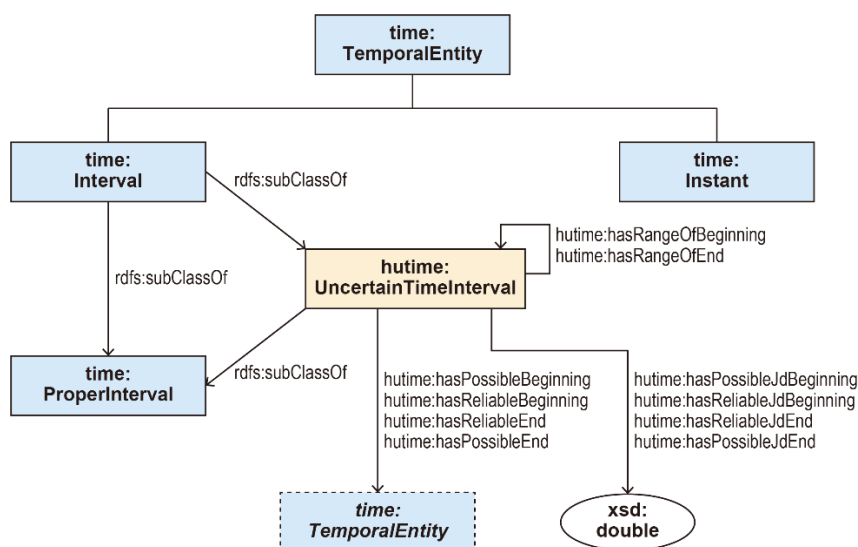


図2 HuTime オントロジ.
あいまいな期間を示す `hutime:UncertainTimeInterval` クラスは、OWL-Time のクラスと互換性を持つ。
Figure 2 HuTime Ontology.
`hutime:UncertainTimeInterval` class representing uncertain time interval has compatibility with classes of OWL-Time.

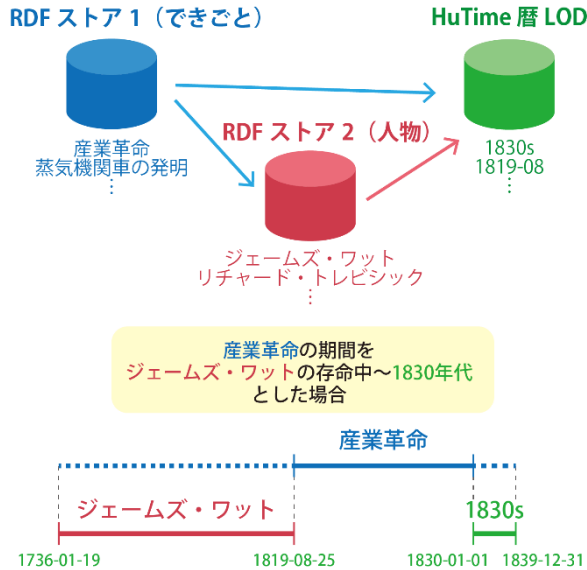


図3 時間名リソース「産業革命」の実験的な定義。期間の始点と終点の時間範囲（点線）は、外部のリソースを参照している。

Figure 3 A experimental definition of “Industrial Revolution.” Beginning and end ranges (dotted line) refer external resources.

3. 時間名リソース

3-1. 時間名が表す期間の記述

例として、時間名「産業革命」が表す期間を、外部のリソースを参照しながら記述する(図3)。周知のとおり、産業革命はある日突然始まったわけでもなければ、終わったわけでもなく、明確な始点や終点の日付を決めることができない。このため、この例では、実験的な定義として、産業革命は「ジェームズ・ワットの存命中にはじまり、1830年代に終了した」としている。その上で、時間名「産業革命」が表す期間の始点および終点の範囲として、人物の生没年を提供する人物リソースと暦 LOD の2つの外部のリソースを参照している。

この時間名リソース「産業革命」とその参照先の各リソースの RDF データの例が図4である。産業革命リソース (ex1:産業革命) の始点の時間範囲として、人物リソースの ex2:ジェームズ・ワットを、プロパティ `huntime:hasRangeOfBeginning`

を介して参照している(プレフィックス ex1 および ex2 は架空のもの)。同様に、終点の時間範囲として、暦 LOD の `hcal:101.1/10y/1830s` を、プロパティ `huntime:hasRangeOfEnd` を介して参照している。さらに、ex2:ジェームズ・ワットの始点(出生)および終点(死没)の暦日は、暦 LOD の暦日リソースを参照している。このように、リソース間の階層的なリンクにより、時間名が表す期間を、その定義に沿った形で RDF データとして記述することが可能となる。

HuTime オントロジに則って記述された時間名リソース「産業革命」は、新たな期間の記述に利用することもできる。例えば、時間名「近代」が産業革命から始まるものとして定義するのであれば、プロパティ `huntime:hasRangeOfBeginning` を介して ex1:産業革命を参照することで、時間名リソース「近代」をその定義に沿った形で記述することができる。

3-2. 時間軸上の位置の取得

図4で示されているリソース間のリンクを辿って、産業革命リソース (ex1:産業革命) が表す期間の時間軸上の位置を取得する。始点と終点が時間範囲として表されるため、両者の時間範囲の端点、つまり、4つの時間点の位置を求めることが必要となる(始点範囲: Pb~Rb, 終点範囲: Re~Pe) (図1参照)。また、表示や比較などの処理を前提とすれば、各点の位置は、時間軸上の位置を実数で表すユリウス通日の値として取得することが望ましい。図4aの産業革命リソース (ex1:産業革命) 自身はユリウス通日の値を保持していないものの、定義に沿って外部リソースへのリンクを辿ることにより、ユリウス通日の値を取得することができる。

(始点範囲の取得)

始点として参照している外部リソース ex2:ジェームズ・ワットが表す期間(生存期間)(図4b)を取得すると(1)、始点(出生)と終点(死没)の暦日が暦 LOD のリソースを参照していることが分かる。さらに、リンクを辿ってこれらの暦 LOD のリソース(図4c)が表す範囲を取得すると(3, 4)、ex2:ジェームズ・ワットの期間の

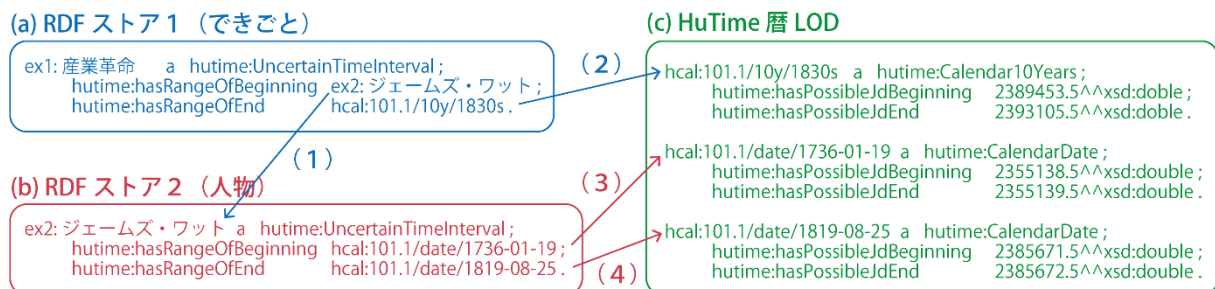


図4 時間名リソース「産業革命」と参照されているリソースの RDF データ (Turtle 形式)。

Figure 4 RDF data (Turtle format) of time name “Industrial Revolution” and referred resources.

始点（出生）と終点（死没）の時間軸上の位置を表す 4 つのユリウス通日の値は、Pb: 2355138.5, Rb: 2355139.5, Re: 2385671.5, Pe: 2385672.5 となる。なお、ここではジェームズ・ワットの出生および死没の時刻が分からないため、その始点と終点は、1 日間の幅を持った時間範囲として表現される。

次に、得られた ex2:ジェームズ・ワットの期間の値（ユリウス通日）を、ex1:産業革命の期間の始点の範囲として設定する。この時、ex2:ジェームズ・ワットの期間が取り得る最大限の範囲を ex1:産業革命の始点の範囲とする。したがって、ex2:ジェームズ・ワットの値の Pb（最小値）と Pe（最大値）の値を ex1:産業革命の始点の時間範囲として用いる。結果として、ex1:産業革命の始点の時間範囲は、Pb: 2355138.5, Rb: 2385672.5 となる。

（終点の時間範囲）

終点として参照している暦 LOD のリソース（図 4c）へのリンクを辿り（2）、その時間範囲を表すユリウス通日の値を得る。始点と同様の手順により、ex1:産業革命の終点の時間範囲は、Re: 2389453.5, Pe: 2393105.5 となる。

3-3. 検索・解析への時間名の活用

あいまいな期間に関する理論[1]では、あいまいな期間同士を比較し、その前後関係や包含関係を Allen の時間関係[10]に基づいて検証する方法が確立されている[11]。これにより、本研究が提案する枠組みに則って記述された時間名リソースは、検索や解析に活用することもできる。

たとえば、「産業革命中に起きたできごと」を検索する場合、図 4 のリソース ex1:産業革命を検索条件として指定できる。検索プログラムは、前節で述べた手順でユリウス通日の値を得た後、データベース内のレコードが持つユリウス通日の値と比較し、該当するレコードを検索結果として返す。なおこの例では、データベース内の各レコードについて、比較の処理の前に、HuTime Calendar API [12]などにより年月日がユリウス通日に変換されていることが前提である。

このような時間名を使った検索は、年月日の入力が必要になるという簡便さ以外の意義もある。一つは、単なる時間名同士の文字列マッチングではなく、時間名が表す期間に基づいて検索が行われていること。もう一つは、検索条件としてリソース ex1:産業革命が指定されことにより、この「産業革命」は、誰がどのように定義した期間であるかを明確にできることである。

4. 実証実験

3-2 節が示す手順で、実際に時間名が表す期間のユリウス通日の値を取得できるかを検証した。検証に用いたプログラムは、Web ブラウザ上で動作するもので、JavaScript により実装された。検証は、図 4 のリソース ex1:産業革命の RDF デ

ータを取得し、そこに記述されている期間の定義に沿ってリンクを辿りながら、始点と終点の時間範囲を表す 4 つの時間点のユリウス通日の値を取得した。この過程で、参照する各リソースのデータを取得するたびに、経過時間を記録した。これらにより、3-2 節が示す手順で RDF データが取得され、その結果として、想定通りのユリウス通日の値が取得できることを確認した。

図 4 のリソース ex1:産業革命とリソース ex2:ジェームズ・ワットの RDF データは、RDF ストア 1（できごと）、RDF ストア 2（人物）に見立てたインターネット上の別々のサーバに収容した。これらのサーバには、RDF データ（Content-Type: application/rdf+xml）が取得できるよう、CORS（Cross-Origin Resource Sharing）[13]に則って必要な設定を行った。また、暦 LOD は、すでに稼働・公開されているものをそのまま利用した。

経過時間の計測は、10 回試行し、その平均値を求めた。Web ブラウザのキャッシュの影響を排除するため、試行ごとに Web ブラウザを再起動した。

表 1 に検証結果を示す。まず、ex1:産業革命を定義した RDF データが取得された（経過時間の起点）。次に、ex1:産業革命の始点の時間範囲として指定された ex2:ジェームズ・ワットの定義が RDF ストア 2（人物）に見立てたサーバから取得され（1）、続いて、終点の時間範囲として指定された hcal:calendar/101.1/10y/1830s が暦 LOD から取得された（2）。さらに、ex2:ジェームズ・ワットの終点 hcal:calendar/101.1/date/1819-08-25 および始点 hcal:calendar/101.1/date/1736-01-19 を示すリソースが暦 LOD から取得された（4, 3）。

表 1 各リソースの取得経過。時間は、リソース ex1:産業革命を取得した時点からの経過時間（ミリ秒）。

Table1 Progress of access to the resources. The time is the elapsed time (milliseconds) from the time when resource ex1:産業革命 was accessed.

経過時間 (ミリ秒)		参照先のリソース	図 4 内の 参照番号
平均	標準 偏差		
0	—	ex1:産業革命	
6	5	ex2:ジェームズ・ワット	(1)
321	10	hcal:101.1/10y/1830s	(2)
329	9	hcal:101.1/date/1819-08-25	(4)
333	9	hcal:101.1/date/1736-01-19	(3)

ex2:ジェームズ・ワットの始点と終点の取得順序が逆ではあるものの、結果として、これらのアクセスの順序は 3-2 節の記述と同じであり、想定通りの手順が実行されたことが示唆された。なお、暦 LOD からのデータの取得に時間がかかっている原因として、303 コンテンツ・ネゴシエー

ション[14]によるアクセス回数の増加や、暦に関するデータの出力に要する暦 LOD サーバ上での計算時間が影響していると考えられる。

検証用のプログラムが最終的に出力した ex1: 産業革命の期間を表す 4 つの時点のユリウス通日の値は、Pb: 2355138.5, Rb: 2385672.5, Re: 2389453.5, Pe: 2393105.5 であった。これは、3-2 節で示した結果と同じであり、4 つの時点のユリウス通日の値が想定どおり取得できたことを示している。

これにより、図 3, 4 のように時間名の期間を定義し、3-2 節で示した通り、この定義に基づいてリンクを辿ることにより、その時間軸上の位置を決定できることがインターネット上の実環境で実証された。

5. 時間名リソースの拡充と課題

5-1. 時間名リソースの拡充

本研究の提案する手法では、公開されている時間名リソースの数が多いほど、相互連携が進み、その有用性が高まる。HuTime オントロジに基づいたデータが構築されることが最善ではあるものの、既にインターネット上に公開されている数多くの時間情報を活用するための手段を提供することも、当座の対策として有効である。

(クライアント側での対応)

インターネット上の時間情報をそのまま読み込んで、クライアント側で処理することにより、時間名リソースとして利用する。このような仕組みは、時間情報システム HuTime のライブラリにも組み込まれる予定である。

たとえば、カンマ区切りテキスト (CSV) [15]などの表形式のデータを時間名リソースとして扱うには、期間の始点と終点を示す列名、および、それらの値の暦法の種類を指定することにより、各レコードを、HuTime オントロジに則った時間名リソースの RDF データとして利用することができる (図 5)。

表形式のデータだけでなく、RDF データも、期間の始点と終点を特定するプロパティ、および、それらの値の暦法を指定することで、同様の処理が可能である。

Period	From	To
時間名A	昭和39年10月1日	平成20年12月14日
時間名B	昭和60年10月1日	平成24年3月16日



```
ex:時間名A a htime:UncertainTimeInterval;
  htime:hasRangeOfBeginning hcal:1001.1/date/昭和39年10月1日;
  htime:hasRangeOfEnd hcal:1001.1/date/平成20年12月14日.
```

図 5 表形式のデータを時間名リソースとして利用するためのデータ変換。

Figure 5 Data conversion to utilize table data as a time period resource.

(サーバ側での対応)

上述の既存のデータをそのまま活用する方法は、簡便である反面、クライアント側での処理やネットワークトラフィックの増大により、スループットの低下を招く。このため、データを提供するサーバ側で、あらかじめ、HuTime オントロジに則ったデータを取得できるようにデータを加工しておくことが考えられる。特に、期間を表す 4 つの時点のユリウス通日の値をあらかじめ求めておき、その情報を付加しておく、リンクを辿らなくても時間軸上の位置を取得できるため、処理速度の向上が期待できる。HuTime プロジェクトでは、このような既存の RDF データや表形式のデータにユリウス通日の値に関する情報を付加するためのサービスを提供することも計画している。

5-2. 利用に向けた課題

(時間名リソースの検索とスキーマ)

今後、さまざまな機関や分野から時間名リソースが公開された場合、それらの中から適切な時間名リソースを見つけるための仕組みが必要になる。解決策としては、対象とする期間、研究分野、記述言語、公開元などによる横断的な検索が想定される。

本研究は、検討対象を時間名が表す期間に絞って進めてきたが、このような検索を実現するには、研究分野や公開元などの時間名に付随する情報全体をカバーするスキーマの設計が急務である。

(既存の規格との整合性)

Dublin Core [16]をはじめとする既存の多くの規格では、時間は ISO-8601 形式[17]のリテラル値で指定することが規定されており、時間名リソースのようなリソースを指定することが想定されていない。OWL-Time が W3C 勧告となったことにより、今後、これらの規格においても、リソースによる時間の指定が可能となるように改訂されることが期待される。前述のとおり、時間名リソースが基礎とする HuTime オントロジは、OWL-Time との互換性を有する。このため、これらの規格において、OWL-Time に則ったリソースでの時間の指定が可能となれば、時間名リソースも同時に利用可能となる。

6. 結論

本研究では、「江戸中期」や「バブル期」などの用語、および、建物や人物などの時間に関連付けられる用語を時間名リソースとして RDF データ化し、Linked Data を通じて時間参照に利用する枠組みを実現した。これにより、研究分野や組織ごとに定義が異なる時間名を、URI で識別しながら、データ構築や解析に用いることが可能になった。

これらの時間名リソースが表す期間の始点と終点は、HuTime オントロジ[5]に則り、別の時間

名リソースを参照しながら記述された。これにより、できごとなどの情報に基づく時間名本来の意味に沿って、時間名が表す期間を記述することができる。これは、時間名を単純に年月日に対応付ける従来の方法とは異なるアプローチである。

さらに本研究は、この時間名の定義に沿ってリンクを辿ることにより、時間名が表す期間の時間軸上の位置が取得できることを実証した。これにより、時間名を時間参照に利用することができるようになる。また、リンクを次々に辿りながら時間軸上の位置を示す値（ユリウス通日）を取得する本研究の手法は、Linked Data 技術の応用という点でも新たな試みである。

結果として本研究は、「時間情報は年月日を用いて記述しておけば良い」という考え方に一石を投じ、時間名を本来の意味や目的に沿って有効に活用する道筋を開いた。今後、本研究の枠組みを活用した時間名リソースが数多く公開され、時間情報の基盤となってゆくことが期待される。

なお、本研究で検証用に実装されたプログラムは、修正およびテストの後、時間情報システム HuTime[8]に導入され、ソースコードと共に公開される (<https://github.com/hutime/WebHuTime>)。

謝辞

本研究は、JSPS 科研費 15H01723 「セマンティック・クロノロジー：時間軸に沿った知識の可視化と利用に向けた基盤構築」の助成を受けたものである。また、部分的に、JSPS 科研費 16H01897, 16H01830, 16H02918, 17H00773, 18H03588 の助成を受けている。

参考文献

- [1] Sekino, T. Using Uncertain Time Intervals in Linked Data. *International Journal of Geoinformatics* Vol. 15, No 2, 2019, p. 15-23.
- [2] World Wide Web Consortium (W3C). Resource Description Framework (RDF): Concepts and Abstract Syntax. <https://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> (参照 2019-10-27).
- [3] Bizer, C, Heath, T, and Berners-Lee, T. Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, Vol. 5, No. 3, 2009, p. 1-22.
- [4] Berners-Lee, T., Fielding, R. and Masinter, L. RFC3986, Uniform Resource Identifier (URI): Generic Syntax. <https://tools.ietf.org/html/rfc3986> (参照 2019-10-27).
- [5] Sekino, T. HuTime Ontology to Represent Uncertain Time Intervals. *Proceedings of Workshop "Ontologies for Linked Data in the Humanities", Digital Humanities Conference 2019*.
- [6] World Wide Web Consortium (W3C). Time Ontology in OWL. <https://www.w3.org/TR/owl-time/> (参照 2019-10-27).
- [7] Dershowitz, N. and Reingold, E.M.: *Calendrical*

Calculations, Cambridge University Press (2007).

[8] 関野 樹. Web HuTime - 時間情報のための Web プラットフォーム. *情報処理学会シンポジウムシリーズ* Vol. 2016, No. 2, 2016, p. 125-132.

[9] Sekino, T. Basic linked data resource for temporal information. *Proceedings of the 2017 Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC)*, 2017, p. 76-82.

[10] Allen, J. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 1983, Vol. 26, No. 11, p. 832-843.

[11] Sekino, T. Data description and retrieval using periods represented by uncertain time intervals. *Journal of Information Processing*. *In review*. (preprint: arXiv:1905.04611)

[12] 関野 樹. 暦に関する Web API - 暦法の変換と期間の計算. *情報処理学会シンポジウムシリーズ じんもんこん* 2017, 2017, p.23-28.

[13] World Wide Web Consortium. Cross-Origin Resource Sharing. <https://www.w3.org/TR/2014/REC-cors-20140116/> (参照 2019-10-27).

[14] World Wide Web Consortium. Cool URIs for the Semantic Web. <https://www.w3.org/TR/cooluris/> (参照 2019-10-27).

[15] Shafranovich, Y. RFC4180, Common Format and MIME Type for Comma-Separated Values (CSV) Files. <https://tools.ietf.org/html/rfc4180> (参照 2019-10-27).

[16] Dublin Core Metadata Initiative. DCMI Metadata expressed in RDF Schema Language. <https://www.dublincore.org/schemas/rdfs/> (参照 2019-10-27).

[17] International Organization for Standardization (ISO). ISO 8601-1:2019, Date and time — Representations for information interchange — Part 1: Basic rules. <https://www.iso.org/iso-8601-date-and-time-format.html> (参照 2019-10-27).