

# デジタルアーカイブとLinked Dataを用いた歴史学 研究支援に関する研究

著者	中村 覚
学位授与年月日	2017-03-23
URL	<a href="http://doi.org/10.15083/00076169">http://doi.org/10.15083/00076169</a>

## 論文の内容の要旨

論文題目 デジタルアーカイブとLinked Dataを用いた  
歴史学研究支援に関する研究

氏 名 中村 覚

### 1. 研究の目的

歴史学研究は、歴史資料（以下、資料）の目録データ作成や資料提供を行う「資料管理」、研究者が資料を用いて研究課題の解明に取り組む「資料研究」、資料や研究成果を広く一般に公開する「成果公開」の三つのプロセスから構成される。しかし、各々のプロセスで求められる成果物、それを生み出す方法論や活動主体の違いにより、これらのプロセスは独立して進められているのが一般的である。

本研究では、歴史学研究の発展に向け、上述したプロセス間の有機的な連携を支援することを目的とする。具体的には、デジタルアーカイブの活用による各々のプロセス支援に加え、プロセス間の成果物を **Linked Data** を用いて関連づけ、それらを相互に利用可能な手法の提案およびシステムの開発を行う。また、提案手法を複数の歴史学研究事例に適用し、有用性の評価を行う。

### 2. 提案する歴史学研究支援手法

**Linked Data** とは、分野やシステムを超えたデータ共有や再利用の支援を目的とし、Web上で機械可読な構造化データを公開する技術の総称である。この **Linked Data** を用いて、歴史学研究における成果物の共有と活用を支援する提案手法の概要を図1に示す。

本手法では、**Linked Data** を実現するための記述言語である **RDF (Resource Description Framework)** を用いて、歴史学研究を構成する「資料管理」「資料研究」「成果公開」プロセスの成果物を記述し、Web空間上で統合的に管理する。また、このデータを利活用するシステム群を開発し、各プロセスの要件に応じた支援を行う。さらに、各々のシステムが他のプロセスにおける成果物を関連づけて利用することにより、歴史学研究を構成するプロセスを有機的に関連づける。

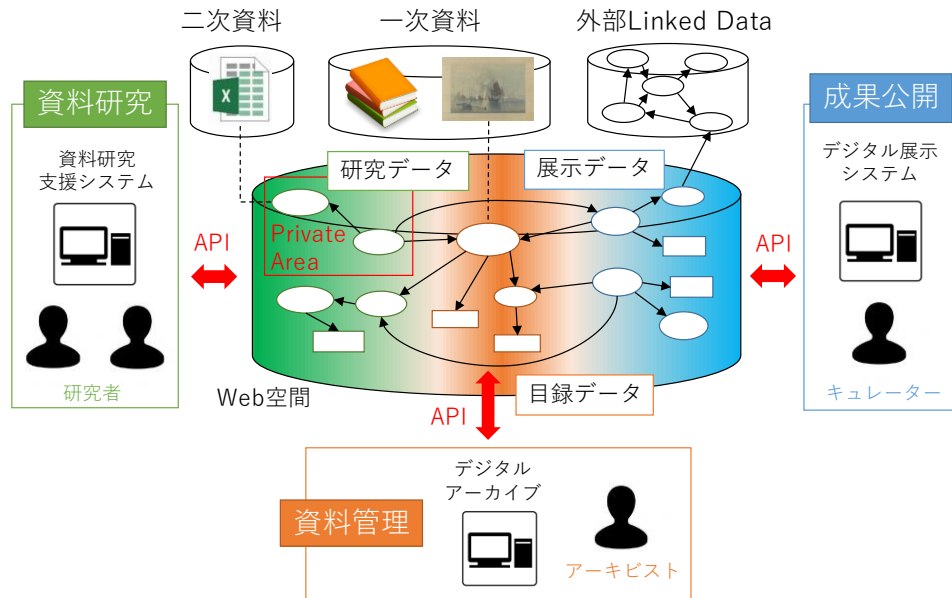


図 1 提案する歴史学研究支援手法

## 2.1. 「資料管理」プロセスの支援手法

「資料管理」プロセスは、管理対象とする資料の目録データの作成と、デジタルアーカイブの構築による資料提供の二つのプロセスから構成される。これらのプロセスの支援を目的とし、本手法では一般に MS Excel など管理される目録データを RDF データに変換する手順の定型化を行う。

また、Linked Data として公開された目録データを利用するアプリケーションの一つとしてデジタルアーカイブを位置付けることにより、アプリケーションとデータの疎結合を実現し、他のプロセスにおけるデータ活用を支援する。さらに、このデジタルアーカイブ構築において共通して必要となる機能群を提供するパッケージシステムを開発する。このパッケージシステムを利用することにより、構築コストの低減に加え、適用ドメインの違いに応じた機能追加やインターフェース開発に注力可能な環境を提供する。

## 2.2. 「資料研究」プロセスの支援手法

「資料研究」プロセスは資料の収集・整理・分析プロセスから構成される。本手法では、これらの各々のプロセスを支援する手法の提案、およびシステムの開発を行う。

資料の収集プロセスについては、Web API を利用して Linked Data 化された目録データに機械的にアクセスすることにより、資料の収集に要する労力を低減する。

整理プロセスについては、「資料管理」や「成果公開」プロセスへの成果物の活用を視野に入れ、研究者が蓄積するメモや二次資料などの研究データを資料のメタデータとして管理する手法を提案する。特に、資料研究の特徴である研究者の目的に応じた管理項目の多様性に対して、高い柔軟性を持つ RDF を用いることで、研究者による管理項目の追加や更新を可能とする。

分析プロセスについては、収集した目録データと、研究者が蓄積した研究データを活用す

ることにより、資料のメタデータに基づく定量的な分析を支援する。具体的には、研究者が設定した管理項目に基づき、資料の分類結果や、資料の出現頻度と関連年表を組み合わせた変遷史の可視化機能を提供する資料研究支援システムを開発し、研究者への新たな気づきの提供や、定量的な分析に基づく仮説の検証を支援する。

### 2.3. 「成果公開」プロセスの支援手法

「成果公開」プロセスでは、歴史になじみのない一般利用者への情報提供が目的となるため、対象者の興味や関心を喚起させる資料説明や背景知識を補完するキーワードの整理、地図や年表などを用いた視覚的な情報提供が求められる。このプロセスの支援に向け、デジタルアーカイブを用いて資料や研究成果を公開するデジタル展示を行う。本手法では、資料のキャプションやキーワード、展示テーマなどの展示データを RDF によって記述するためのスキーマ定義を行う。また、Wikipedia から情報を抽出して Linked Data として公開するプロジェクトである DBpedia との連携を通じ、キーワードの概要説明やサムネイル画像などの活用による一般利用者への背景知識の補完を支援する。

さらに、資料を年表や地図にマッピングして表示する可視化インタフェースを開発し、Web 空間の活用による展示スペースの制約を超えた多角的な資料提供を支援する。

## 3. 適用事例

### 3.1. デジタルアーカイブ構築による資料管理事例

柏図書館が所蔵する『平賀讓文書』、東洋文庫や東京大学文書館が所蔵する各種資料を対象とした資料管理事例への適用を通じ、提案手法の評価を行った。それぞれの組織が所蔵する目録データを RDF データに変換し、2.2 で述べたパッケージシステムを用いることにより、それぞれの組織の所蔵資料を Web 上で公開するデジタルアーカイブを構築した。この構築事例を通じ、複数ドメインへの適用実績に基づく提案手法の汎用性を示した。

また、上述した『平賀讓文書』や東京大学文書館の所蔵資料に加え、東京大学において公開されている『鴟外文庫』などの複数の資料群を対象としたデジタルアーカイブ構築を行った。各々の資料群に関する目録データを RDF データ化することにより、Web 空間上での相互利用が可能となり、複数の資料群に対する横断検索を可能とするポータルサイトを構築した。これらのデジタルアーカイブの構築実績から、本手法の有用性を確認した。

### 3.2. 資料研究事例

本適用事例では、『平賀讓文書』を対象とした戦艦の設計変更の分析や、東京大学文書館が所蔵する重要文化財の一つ『文部省往復』を対象とした明治初期の東京大学の設立史に関する分析を行った。3.1 の「資料管理」プロセスの成果物である Linked Data 化された目録データを活用することにより、デジタルアーカイブが公開する大規模なデータセットを研究対象資料として取り扱うことが可能となった。また、目録データとして提供されていない図面資料に記述された戦艦の煙突本数や、『文部省往復』に登場する東京大学の組織名称などを資料のメタデータとして新たに追加した。これにより、研究者の目的に応じた資料分類を可能とし、一次資料に基づく定量的な分析を実現した。特に『文部省往復』については、

資料の出現頻度に基づく変遷史の可視化を行うことにより、既知の学説とは異なる特徴を見出すことができた。

さらに、個人研究者を対象とした資料研究事例として、外交文書の送付先の決定過程に関する分析を行った。研究者が独自に収集した資料を Web 空間上で管理することにより、上述した事例と同様、研究者の目的に応じた研究データの蓄積や資料分析を可能とした。また、研究者による定性的な分析と情報技術者による定量的な分析を組み合わせることにより、定量的な分析に基づく研究者への気づきの提供や、研究者が定性的な分析に基づいて立案した仮説に対する定量的な検証などを実現した。この結果、研究課題とした外交文書の送付先決定者の特定や、送付先決定過程の変化に関する知見を得ることができた。

これらの適用事例を通じ、提案手法が資料研究プロセスにおいて、新たな知見の発見に寄与することを確認し、本手法の有用性を示した。

### 3.3. デジタル展示を用いた成果公開事例

3.1 で開発したデジタルアーカイブを用い、資料および研究成果を公開するデジタル展示を行った。展示に必要なデータ作成については、DBpedia が公開するデータセットとの連携により、展示資料のキーワードに関する概要説明や位置情報、年代情報を半自動的に補完することが可能となった。また、これらの情報を利用し、関連年表と展示資料を関連づけて展示する可視化インタフェースなどを構築することができた。この結果、過去に開催された企画展示などをシステム上で再現することができたことに加え、Web 空間を活用した永続的な成果公開を可能とした。さらに、『平賀讓文書』、東洋文庫、東京大学文書館を対象とした異なるドメインにおけるデジタル展示を実現し、本手法の有用性と汎用性を確認した。

## 4. 歴史学研究を構成するプロセス間の連携に関する考察

これまでの適用事例を通じ、「資料管理」プロセスの成果物を「資料研究」「成果公開」プロセスへ活用することができた。これに加え、他のプロセス間における成果物の相互利用例として、例えば「資料研究」プロセスで整理した翻刻結果や同定した年代情報などを目録データに反映させることにより、「資料管理」プロセスにおける目録データの追加や修正を行うことができた。また、地図や年表を利用した資料展示インタフェースを資料検索手法の一つとして利用することで、デジタルアーカイブの機能改良につなげることができた。これらの結果、提案手法が歴史学研究を構成するプロセス間の連携に寄与することを確認した。

## 5. 研究の結論

本研究では、歴史学研究を構成する「資料管理」「資料研究」「成果公開」プロセスの成果物を Linked Data によって統合的に管理する手法の提案、およびデジタルアーカイブを成果物の活用基盤として利用するシステム群を開発した。複数の組織を対象とした適用事例を通じ、提案手法が各々のプロセス支援に加え、プロセス間の成果物を関連づけて利用することにより、歴史学研究を構成するプロセスの有機的な連携の実現に寄与することを確認した。