

# 歴史マイクロナレッジの構築と人文学のための対話型AIシステムの開発

## どんな研究？

- 歴史マイクロナレッジ：史料に記された複雑で多様な歴史情報をRDFを用いたLinked Dataとして記述するための方法論とツールを開発
- 生成系AIの歴史研究への応用：史料典拠に依拠して学術的に信憑性のある回答を生成する対話型AIシステムの開発

## 何がわかる？

- さまざまな歴史情報をデジタルデータとして広範に接続することで過去の世界をデジタル空間に再構成し、探索できるようにする
  - ➔ 歴史ビッグデータの構築
- AIを活用した新たな歴史研究の方法論を実践的に検証し、確立する
  - ➔ AI歴史学

## 背景・目的

### 歴史情報データ化の課題

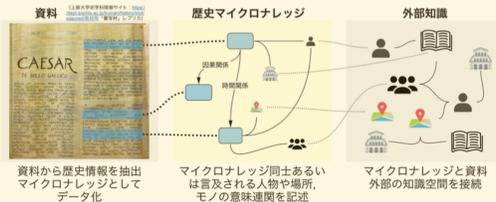
歴史地名や歴史時間情報に関する基礎データの整備や、歴史史料の来歴を記述するためのメタデータ・スキーマの整備は進みつつあるが...



肝心の「史料の中身」のデータ化については研究が不十分...

### マイクロナレッジとは？

史料における歴史的な出来事、関係性、状況、環境への「言及」をデータの基本単位として歴史知識をRDFで構造化するためのモデル



### 生成AIの利用

大規模言語モデル（LLM）をはじめとする生成AIの飛躍的発展によって、「史料の中身」のデジタル化や解析が可能になりつつあるが...

- 典拠の明示性と信頼性が担保されたAI利用環境の整備
- AIを組み込んだ歴史（人文学）研究ワークフローの検討と確立

## 研究内容

### HIMIKOエディタの開発

#### HIMIKO (Historical Micro Knowledge and Ontology) エディタとは？

歴史研究者による歴史マイクロナレッジ構築を支援するためのRDFエディタ。テキストや画像を閲覧しながら、直感的なグラフ描画とデータ入力によって複雑なRDF知識グラフを作成することができる。

#### 柔軟な知識グラフ描画 & 高い一次史料参照性



#### HIMIKO Editor for Image

- IIIF manifestによる画像のインポート
- IIIF Content State APIに基づく画像のキュレーション・アノテーション
- グラフノードとキュレーションの接続

#### HIMIKO Editor for Text

- TEI/XMLによるテキストのインポート
- 既存のTEIタグとグラフノードの接続
- 任意の領域指定によるグラフノードとテキストの接続

#### 基本/共通機能

- ノードとエッジの描画による知識グラフ構築
- JSON、TURTLE形式でのデータダウンロード
- JSONによるデータアップロード

#### 設定ファイル編集ツール

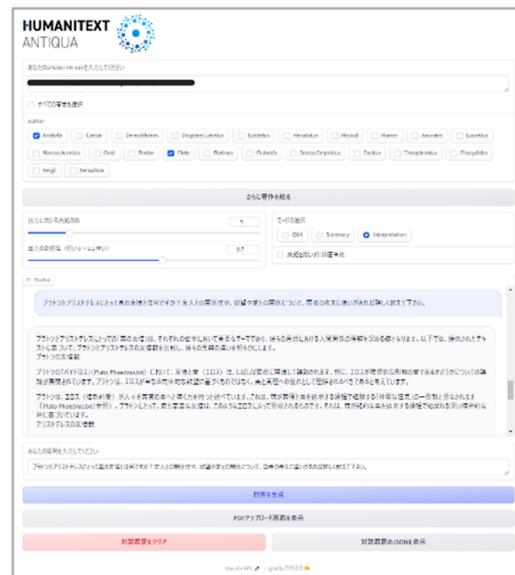
- プロジェクトごとに定義する語彙・オントロジーに基づいて設定ファイルを作成
  - ➔ 詳細なカスタマイズが可能に！



### 歴史研究とAI

#### Humanitext Antiquaの開発

名古屋大学・桜美林大学の研究者と共同で、西洋古典原典のデータベースを再整備し、信頼できる資料に基づく専門的解釈システムを開発。



#### コンセプト

- 学術的に信頼できるテキストデータに基づいて回答を生成
- 回答生成の典拠を明示するとともに、即時参照可能に（RAG）
- 多様な利用可能性をユーザに提供

#### 基本機能

- 回答に用いる著者・著作の選択
- 各種モードの選択
  - ・ RAG or Not RAG
  - ・ 「Q&A」「解釈」「対話」
- 典拠一覧の表示と、典拠テキストのインターフェイス上での閲覧
- 対話履歴のJSON表示
  - ・ 外部プログラムでの再利用性

#### 展望

- マルチモーダル化
- 研究実践の蓄積

### 碑文知識ベースの構築

科研費基盤研究（C）「碑文テキストと実体空間情報のデータ構造化と接続に基づく碑文知識ベースの構築」（24K15674、代表・小川）

