

LLM とオントロジーによる学習リソース拡充手法：A White-box Approach

Method for Expanding LOD Using LLM and Ontology: A White-box Approach

松浦 碧, 林 佑樹, 瀬田 和久

Aoi MATSUURA, Yuki HAYASHI, Kazuhisa SETA

大阪公立大学大学院情報学研究科

Graduate School of Informatics, Osaka Metropolitan University

Email: matsuura@ksm.kis.osakafu-u.ac.jp

あらまし：本研究では、Large Language Model (LLM)を用いた LOD の拡充アプローチにより、学習トピックの理解を促す学習リソースの構築手法について提案する。学習資源として LLM を用いる場合、その処理過程は Black-box となる問題がある。これに対し、LLM による Linked Data 形式の知識生成をオントロジーに基づいて行い既存の LOD と統合することで、学習視点の反映・網羅性・再現性・学習制御性・メンテナンス性・計算機可読性を備えた White-box な学習リソースの構築が可能となる。

キーワード：オントロジーに基づく LOD 拡充, Thinking Like a Historian Approach, Inquiry-based Learning

1. はじめに

オープンエンドな学習空間で、自分の興味に応じて主体的に学習を進めることで、学びの動機付けやエンゲージメントの高まりが期待できる。我々は、オントロジーと LOD (Linked Open Data) に基づいた適応的問い生成により、オープンエンドな学習空間においても多様な学習者の興味・学習内容を捉えた支援の実現を目指している。

Wikidata に代表される LOD は膨大かつ幅広い知識を計算機処理可能とする汎用リソースである。一方、学習領域固有の重要な関係性が必ずしも網羅されているとは限らない。そこで、Large Language Model (LLM) とオントロジーにもとづいた LOD 拡充アプローチにより、学習トピックの理解深化を促す学習リソース構築手法について検討する。

2. 学習リソースとしての LOD 拡充手法

2.1 特徴

本研究では LOD として Wikidata を採用し、「学習領域に応じた学びの視点」からこれを拡充する手法を検討する。

学習資源として LLM を用いる場合、その処理過程は Black-box となり、いわゆるハルシネーションや再現性の課題を考慮する必要がある。

本研究では学習トピックを学ぶ上で重要とされる学習活動、概念、関係性をオントロジーとして定義し、これに基づいて LLM が生成する知識を規定した知識生成を行い、生成した知識を Wikidata と統合することで上述の課題を低減する。

このアプローチの特徴として、学習リソースの観点から以下を担保できる点が挙げられる。

- 1) **学習視点の反映・網羅性**：学習トピックを学ぶ上で重要な概念、関係性を反映するとともに、網羅性を向上できる。
- 2) **再現性**：システムが利活用する知識が明示され、再現性が担保される。

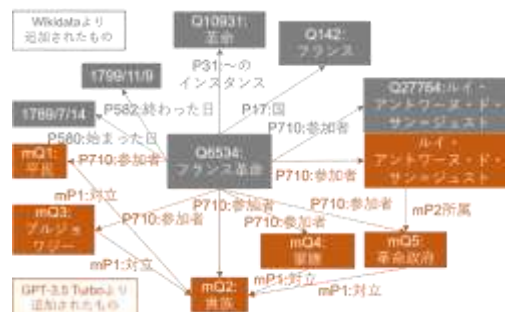


図1 拡充した Linked Data

- 3) **学習制御性**：知識生成主体をシステムは判別できるため、それに応じた学習制御が可能となる。
- 4) **メンテナンス性**：間違った可能性を内包する LLM が生成した知識であっても、生成主体が明示されることにより修正可能性を向上できる。

2.2 学習支援のためのデータ拡充アプローチ

Wikidata に含まれるノード・関係と LLM が生成した Linked Data のノード・関係が、同じ史実・関係を示す場合に id を対応づける必要がある。本研究では、同姓同名問題が生じる可能性を小さくするため、学習対象にしている史実に限定してトリプル集合を生成し、ラベル一致で id を対応づける。

より具体的な LOD の拡充手法を「フランス革命」について学ぶ観点から拡充する文脈で述べる。

1. 学習者が選んだ「(ラベルとしての) フランス革命」が「(概念としての) 革命」であることを、Wikidata に表現された instance-of 関係に基づいて捉える。
2. 「革命」を研究する歴史学者は、例えば当該の革命の「参加者」と「参加者間の関係」に着目して考えることが後述する TLH オントロジーに規定されている (図 2(C))。システムはこの概念定義に基づいて LLM (GPT-3.5 Turbo) にプロンプトを与え、Linked Data を生成する (図 1: 橙色)。例えば、フランス革命の「参加者」とし

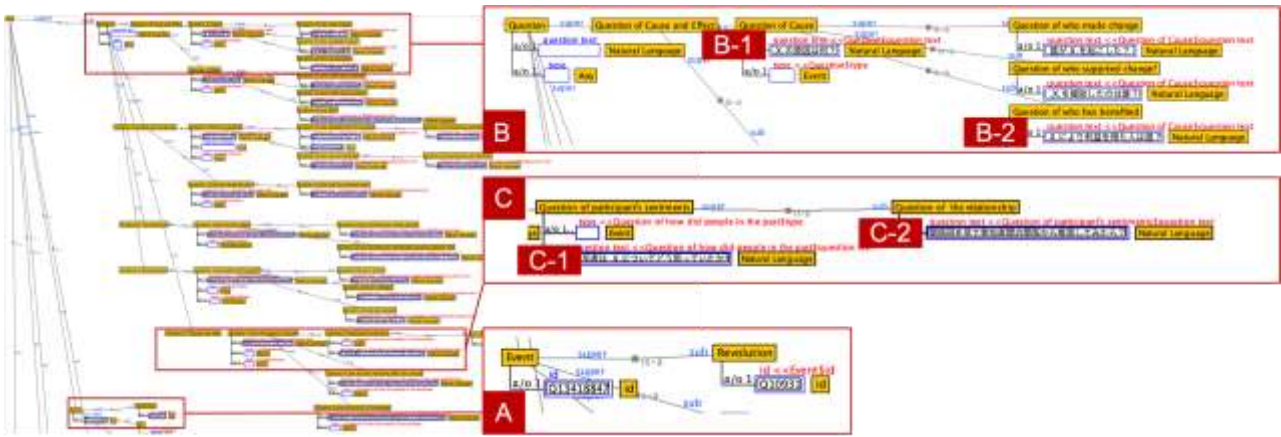


図 2 Thinking Like a Historian Ontology

て「貴族」「ブルジョアジー」を生成したり、「貴族」と「ブルジョアジー」の間に「対立」関係があることを生成する。

3. Wikidata 上のノード・リンクと生成した Linked Data のノード・リンクのラベル一致により、情報を統合・拡充する。

提案手法の1つの特徴として、Wikidata に元々ある人手による信頼性が高い情報と、LLM により拡充した誤りを含む可能性がある情報をシステムが区別できる点がある。このことで、利活用する情報の信頼性に対する自意識に基づいて、システムは例えば、(情報が信頼できない可能性がある)その根拠を探索させるといったように、学習者に対する働きかけを変化させることも可能になる。

3. 学習リソースを用いた学習支援

3.1 有意義な歴史学習を促す方法論

Mandell らは歴史学者がどのような史実探索をするかを体系化した Thinking Like a Historian Approach (TLH)を提唱している⁽¹⁾。ここでは、教科書に書かれたことを事実として暗記する学びから、批判的に探求しながら根拠に基づいた自分なりの解釈を伴う学びに転換することの重要性が説かれている。

TLH では「イベントの原因は何か?」のような学習活動の起点となる 22 個の問いが規定され、それらについて考察する際に意識すべきこととして例えば、「当時の人の視点、イベント以前に周辺で起きた出来事を考察する視点といった複数視点から情報を集める」、「出来事と人物間のつながりを説明する」といったことが体系化されている。

3.2 適応的な問いの提示手法

オープンエンドな空間での学びの適応的支援として、Jouault らは 2 つの LOD を統合した知識モデルを構築するとともに、それを意味処理可能とする歴史ドメインオントロジーを規定している。さらに、こうした概念と対応づけられた問い生成オントロジーを規定することで、巨大な知識ベースに基づいた問いの動的生成手法を提案している⁽²⁾。

本研究ではこの仕組みを援用し、2 章の提案手法

で拡充した LOD と TLH オントロジーに基づく問いの動的生成機構を開発した。

図 2 に TLH オントロジーに定義された問い概念の定義を示す。ここでは、3.1 節で述べた 22 個の問いが、それらについて考える際に意識すべき学習活動と対応付ける形で定義されている。具体的には、どのようなタイプの史実を学ぶときにどんな問いを考えるのが効果的であるか、そして、それについて理解を深める学習活動を規定している。このことで、いま学んでいることについて理解を深める問いを見出すことが難しい学習者に対して、TLH の考えに基づく有意義な探求を促す問いを提示することが可能となっている。

学習者が例えば「フランス革命」について学んでいる場面では、「革命はイベントの一種であり (図 2(A))、その原因 (図 2(B)) や参加者の心情 (図 2(C)) を想像することで理解が深まる」といった概念的つながりにもとづき、「フランス革命の原因は何か? (図 2(B-1))」、「参加者はフランス革命についてどう思っていたか? (図 2(C-1))」といった推論を促す問いを、適応的に提示できるようになっている。

4. まとめ

LLM による Linked Data 形式の知識生成をオントロジーに基づいて行い、既存の LOD と統合することで、2.1 節で述べた 4 つの特徴を有する White-box アプローチを実現した。なお、本稿の学習リソースを活用した学習支援システムについては参考文献 (3)を参照されたい。

参考文献

- (1) Mandell, N. & Malone, B.: “Thinking like a historian: Rethinking history instruction”, Wisconsin History Society (2008)
- (2) Jouault, C., Seta, K. & Hayashi, Y.: “Content-Dependent Question Generation Using LOD for History Learning in Open Learning Space”, New Generation Computing, Vol. 34(4), pp. 367-393 (2016)
- (3) 松浦碧 他: “オープンエンドな学習空間における史実探索学習支援システムの開発と評価”, 第 49 回教育システム情報学会全国大会 (2024, 掲載予定)