



Title	GraphRAG による日本語文学テキストの知識グラフ分析 : 江戸川乱歩『D坂の殺人事件』を事例として
Author(s)	黄, 晨雲
Citation	言語文化共同研究プロジェクト. 2025, 2024, p. 5-17
Version Type	VoR
URL	https://doi.org/10.18910/102219
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

GraphRAG による日本語文学テキストの知識グラフ分析 —江戸川乱歩『D 坂の殺人事件』を事例として—

黄 晨雯

大阪大学大学院人文学研究科
〒 560-0043 豊中市待兼山町 1-8
Email: huang.chenwen.hmt@osaka-u.ac.jp

概要 本稿では、Microsoft が 2024 年に提唱した検索拡張生成手法 GraphRAG を日本語文学に適用し、その応答生成および知識構造の可視化に基づく解釈支援の有効性を検討する。従来の RAG は外部ドキュメントを参照しながら生成を行う手法だが、非構造的情報に依存するため、複雑な文脈や人物関係、因果構造の把握には限界がある。これに対し、GraphRAG は知識グラフを統合することで、エンティティ間の関係を構造的に可視化し、文脈的に一貫した応答を実現する。

本稿では、江戸川乱歩の短編小説『D 坂の殺人事件』を対象とし、主題理解と犯人特定という二種類の問い合わせに対して GraphRAG を用いた応答分析を行った。さらに、Gephi および Neo4j によって生成された知識グラフを可視化し、登場人物間の関係性、構造的中心性、因果関係のネットワーク的把握を試みた。その結果、GraphRAG は日本語文学において意味的整合性のある応答を生成し、構造の視覚化を通じた物語解釈支援にも高い有用性を示した。

キーワード GraphRAG, ネットワーク可視化, 物語構造

Knowledge Graph Analysis of Japanese Literary Texts Using GraphRAG: A Case Study of Edogawa Rampo's “The Murder on D-Slope”

Chenwen Huang

Graduate School of Humanities, The University of Osaka
1-8 Machikaneyama-cho, Toyonaka, Osaka, 560-0043 Japan

Abstract This study applies GraphRAG, a retrieval-augmented generation method proposed by Microsoft in 2024, to Japanese literary texts and examines its effectiveness in supporting textual interpretation through response generation and the visualization of knowledge structures. While conventional RAG relies on referencing unstructured external documents for generation, it faces limitations in capturing complex contexts such as interpersonal relationships and causal structures. In contrast, GraphRAG integrates knowledge graphs to structurally visualize the relationships between entities and generate contextually coherent responses.

In this research, Edogawa Rampo's short story The Case of the Murder on D. Hill is used

as a case study. Two types of queries—one targeting thematic understanding and the other aiming at identifying the perpetrator—are analyzed using GraphRAG-generated responses. Furthermore, the resulting knowledge graphs are visualized using Gephi and Neo4j to explore character relationships, structural centrality, and causal connections through network-based representations.

The results indicate that GraphRAG successfully generates semantically coherent responses in Japanese literary contexts and demonstrates high utility in supporting narrative interpretation through structural visualization.

Keywords GraphRAG, Network Visualization, Narrative Structure

1. 研究背景

近年、大規模言語モデルの発展は、自然言語処理の分野に大きな変革をもたらしている。中でも、検索拡張生成 (Retrieval-Augmented Generation, RAG) は、生成モデルに外部情報源を参照させることで、より正確で文脈に適した回答や文章の生成を可能にする手法として注目されている。RAG は、モデルが事前に獲得した知識に加え、外部の情報検索機能を組み合わせることで、ユーザーの問い合わせに対し動的かつ柔軟に対応する能力を備えている。

一方で、RAG のアプローチにはいくつかの課題も指摘されている。特に、日本語のように文法構造が複雑で省略や曖昧さの多い言語においては、RAG が生成する回答の文脈理解に限界が見られる。また、RAG は主に非構造的なドキュメント検索に依存しており、問い合わせが抽象的あるいは高文脈依存の問い合わせに対しては、関連情報の適切な取得が困難となる場合がある。

こうした背景の中で、2024 年に Microsoft が提唱した「GraphRAG」という新たなアプローチが注目を集めている。GraphRAG は、RAG に知識グラフを統合した新たな生成手法であり、複雑な文脈処理への対応力が期待されている。

本研究では、この GraphRAG を日本語文学テキストに適用する試みとして、江戸川乱歩の短編小説『D 坂の殺人事件』を対象に、その分析手法と成果を検証する。具体的には、Python を用いた GraphRAG ライブラリによりテキストからエンティティとリレーションを抽出し、その結果を Gephi や Neo4j で視覚化・分析することで、知識グラフの構築と解析を行う。

本稿の目的は、GraphRAG が日本語文学テキストに対してどの程度有効に機能するのかを実証することであり、同時に課題と今後の応用可能性についても考察を加えることである。生成 AI の日本語処理の今後を見据え、構造化情報を取り入れた生成モデルの新たな展開を探ることが、本研究の意義である。

2. RAG と GraphRAG の理論的基盤

2.1 RAG (Retrieval-Augmented Generation)

RAG は、大規模言語モデルと外部情報検索を組み合わせた生成手法である。RAG の主な特徴は、事前学習された知識に加え、ユーザーのクエリに応じて外部知識ベースから関連情報を

検索し、それを元により文脈に即した正確な文章を生成できる点にある。

この手法は、社内向けチャットボット、医療相談、法務支援など、正確な情報の参照が求められる場面で特に有用である。一方で、従来の RAG は主に非構造的な文書の検索に依存しており、情報の粒度や関連性が限定されることが多い。たとえば、「昨年の売上高はいくらですか？」のような具体的な質問には強いが、「昨年の経営的成功は何か？」といった抽象的な質問には対応が難しい。

このように、RAG には高い柔軟性と有用性がある一方で、抽象的、かつ関係性の強い問い合わせに対しては限界が存在する。

2.2 GraphRAG の概要と研究動向

このような RAG の限界を克服するために、Microsoft は 2024 年に GraphRAG (Graph-based Retrieval-Augmented Generation) と呼ばれる新たな手法を提案した。GraphRAG は、RAG に知識グラフ (Knowledge Graph) を統合することで、文書間の構造的な関係性を明示化し、それを検索および生成の過程に活用するものである。

GraphRAG のアプローチは、大きく「インデックスフェーズ」と「クエリフェーズ」の二つの段階に分けられる。まず、インデックスフェーズでは、テキストから人物、場所、出来事などのエンティティとその関係性を抽出し、知識グラフを構築する。この過程では、ノードとエッジの情報がグラフ構造として記録される。次に、クエリフェーズでは、ユーザーのクエリに基づいてグラフ構造が探索され、関連するノードやコミュニティ情報が取得される。これらの情報をもとに、言語モデルが回答文を生成する。

このように構造化データを活用することで、GraphRAG は非構造的な検索手法と比較して、文脈理解や関係性の把握において高い精度を実現する。実際、FAQ システムや顧客サポート、研究支援ツールなどの実務領域において、一部企業による GraphRAG の試験的導入が進められている。

Han ら (2024) は、GraphRAG の構成要素および検索生成の過程を整理し、従来の RAG との比較を通じて、その特性や課題 (例: 知識グラフ構築時のノイズやスパース性) を明らかにしている。さらに、問答やクエリ要約といったタスクにおける RAG と GraphRAG の性能差や、両者の統合的活用の可能性についても検討している。一方で、具体的なテキスト分析タスクへの応用には言及されていない。

このように、現時点では GraphRAG を用いた実証的な学術研究はまだ限られており、その応用可能性にはさらなる検討の余地がある。本研究はその初期的な応用例として、日本語のテキスト分析タスクにおける GraphRAG の有効性を検証することを目的とする。

3. 分析対象と方法論

3.1 分析対象：江戸川乱歩『D 坂の殺人事件』

本研究の分析対象は、江戸川乱歩による短編推理小説『D 坂の殺人事件』である。本作は、名探偵・明智小五郎が初めて登場する作品として知られ、日本近代ミステリーの出発点の一つともされている。物語は、語り手である「私」が、友人の明智とともに事件現場となる古本屋に偶然立ち寄り、密室殺人の謎を追うという筋書きを有している。

『D 坂の殺人事件』を本研究の分析対象として選定した理由は、以下の三点に基づく。第一に、本作の日本語原文は構文的に複雑かつ多様であり、江戸川乱歩特有の文語的文体や繊細な心理描写を含んでいる。このような文体は、自然言語処理において一定の困難を伴うため、日本語における GraphRAG の性能評価に適した題材と考えられる。第二に、登場人物・空間・出来事の関係性が明確に記述されており、推理小説としての性質上、人物間の相互関係や時間的進行、行動と動機の因果関係が比較的明瞭に把握できる。そのため、知識グラフ構築に必要な情報構造を明示的に抽出しやすいという利点がある。第三に、本作品は著作権が切れており、青空文庫を通じて一般に入手可能である。このことは、研究の再現性や今後の応用展開を視野に入れた際の利便性という観点からも有用である。以上の理由から、本研究では『D 坂の殺人事件』を分析対象として採用する。

3.2 GraphRAG の実装と応用

環境構築においては、Python 3.10 上に GraphRAG パッケージを導入し、プロジェクトディレクトリの初期化を行った後、OpenAI の API キーを設定し、使用モデルとして gpt-4o を選定した。特に、日本語による出力の自然性を担保するため、プロンプト設計において「日本語で答えてください」と明示的に指定するなど、言語面での調整を加えた。

インデックス作成の工程では、対象となる日本語小説テキストからエンティティおよび関係性を抽出し、構造化された知識グラフを生成した。その後、この知識グラフを参照して、global モードおよび local モードの二通りによる質問応答を実施した。global モードでは作品全体に関わる構造的・意味的情報に基づいた包括的な応答が得られるのに対し、local モードでは特定の文脈に焦点を当てた詳細な情報の抽出が可能である。これにより、物語の主題に関する理解と、登場人物の行動・動機といった具体的要素の双方から分析を行うことができた。

3.3 知識グラフの可視化と分析ツールの活用

知識グラフの構造的特徴を可視化し、複雑なエンティティ間の関係性を俯瞰的に把握するために、Gephi および Neo4j の二つのツールを併用した。Gephi は、GraphRAG によって出力されたグラフデータを視覚的に表示するために用いられ、ノード間の距離、ネットワーク密度、

クラスタリング構造などの可視化に寄与した。一方、Neo4j はグラフデータベースとして機能し、クエリベースの探索やノード属性（例：Person, Location, Organization）の分類、特定条件に基づくサブグラフの抽出に有用である。これにより、GraphRAG の出力結果に対する柔軟かつ構造的な分析が可能となった。

上記の分析手法を通じて、本研究では物語理解における GraphRAG の有効性、特に「主題把握」と「犯人特定」という二つの典型的な問い合わせに対して、どの程度意味的整合性のある応答が得られるかを評価した。

4. 分析結果と考察

第4章では、GraphRAG を用いて江戸川乱歩『D坂の殺人事件』を分析した結果に基づき、日本語テキストにおける意味構造の把握と知識抽出の有効性について考察を行う。

分析対象としては、第3章で述べた二種類の問い合わせ、つまり抽象的な主題分析と、具体的な犯人特定をもとに、それぞれ global モードおよび local モードによる応答結果を比較検討する。ここでは、生成された回答の精度、知識グラフの構造的特性、そして Gephi や Neo4j による可視化が分析に与える効果の三点を中心に論じる。

4.1 小説の主題に関するグローバルクエリ分析

最初に、「この小説のテーマは何ですか？」という問い合わせに対して、GraphRAG は以下のようないくつかの重要なテーマが存在します。それらのテーマは物語の進行やキャラクターの関係性を通じて深く掘り下げられています。

1. 小説のテーマ

この小説にはいくつもの重要なテーマが存在します。それらのテーマは物語の進行やキャラクターの関係性を通じて深く掘り下げられています。

2. 犯罪と捜査

犯罪と捜査のテーマは小説の中心的な要素です。明智や彼、そして警察などの主要なキャラクターが関与する様々な事件や捜査が描かれています。これらの要素は、犯罪を解決することの複雑さや人間行動の理解の難しさを強調している。

3. 人間関係と相互作用

犯罪と捜査の文脈における人間関係と相互作用も重要なテーマです。明智と語り手、彼と私、そして様々な証人や容疑者との関係が物語に深みを与え、犯罪が個人的なつながりに与える影響を探求しています。

4. 心理学的探求

心理学的探求もまた重要なテーマです。明智のようなキャラクターが『心理学と犯罪』といった作品を引用し、人間の観察や記憶の不確かさを説明します。このテーマは、コミュニティの知的深さや心理学理論が捜査に与える影響を強調しています。

5. 社会問題と秘密

古本屋での事件や男と女が関与する殺人事件など、社会問題と秘密のテーマも見逃せません。これらの出来事は、隠された社会問題や虐待的な関係、秘密の欲望を明らかにし、複雑で多面的な物語を形成して

います。

6. 正義と法的手続き

正義と法的手手続きのテーマも重要です。警察、探偵、検察官が様々な捜査に関与することで、正義の追求や犯罪の法的な影響が強調されています。

7. 結論

この小説は、犯罪と捜査、人間関係、心理学的探求、社会問題と秘密、そして正義と法的手手続きといった多岐にわたるテーマを通じて、読者に深い洞察を提供します。それぞれのテーマが絡み合い、物語に豊かな層を加えています。

GraphRAG は、原文テキストに含まれる主題を自動的に分類し、それぞれに関連する文脈情報を探して解釈を付加する点で、高い有用性を示した。とりわけ、日本語における抽象的な表現や比喩的記述に対しても一定の精度で意味を捉え、それを明示的な言語として再構成している点は特筆に値する。また、テーマ間の重なりも適切に記述されており、知識グラフの構造によって各トピック間の関係性が視覚的かつ論理的に提示されることで、作品に内在する複雑な意味構造をより明確に理解できるようになっていた。

4.2 犯人に関するローカルクエリ分析

次に、「殺人を犯したのは誰ですか？」という問い合わせに対する GraphRAG の応答を検討する。

1. 殺人を犯したのは誰か？

殺人を犯したのは、旭屋の主人です。彼は物語の中で中心的な役割を果たしており、複数の証拠が彼を犯人として指し示しています。

2. 旭屋の主人の役割

旭屋の主人は、残酷なサディストとして描かれており、殺人罪を犯したと疑われています。彼は罪跡を隠すために便所を借りた男のことを話しており、その行動が彼の犯行を裏付けるものとされています [Data: Entities (93); Relationships (102, 127, 128, 130, 129)]。

3. 犯行の動機と手口

旭屋の主人は、サード卿や女のマゾッホとの関係を持ち、複雑な人間関係の中で犯行を行ったとされています。彼の行動は計画的であり、証拠を隠すために様々な手段を講じています。例えば、電燈のスイッチを操作して指紋を残さないようにしたり、便所を借りて証拠を隠そうとしたりしています [Data: Entities (93); Relationships (102, 128, 130)]。

4. 他の登場人物との関係

犯人である旭屋の主人は、被害者との間に悪意がなかったとされる関係を持っていました。被害者は犯人と同意の上で殺害された可能性があり、この点が物語の複雑さを増しています [Data: Entities (91); Relationships (100)]。また、明智は犯人の存在を否定し、自分の無罪を証明しようとしていますが、これも物語の緊張感を高める要素となっています [Data: Relationships (39)]。

5. 結論

以上の情報から、殺人を犯したのは旭屋の主人であることが明らかです。彼の行動や人間関係、そして証拠の隠蔽工作が彼を犯人として特定する要因となっています [Data: Entities (93); Relationships (102, 127, 128, 130, 129)]。

```
{} create_final_relationships.parquet.as.json 1 ×
118 {"source": "\\"法廷\"", "target": "\\"会員\"", "weight": 1, "description": "\\"会員は、法廷に持ち出される問題に関する記録を書くように指示
119 \"source": "\\"薔薇屋の主人\"", "target": "\\"死んだ細君\"", "weight": 1, "description": "\\"薔薇屋の主人は死んだ細君の夫であり、彼女の身
120 \"source": "\\"薔薇屋の主人\"", "target": "\\"話者\"", "weight": 1, "description": "\\"話者は薔薇屋の主人に対して聯想診断を行い、彼の心理
121 \"source": "\\"薔薇屋の主人\"", "target": "\\"自首\"", "weight": 1, "description": "\\"薔薇屋の主人は自首したことで物語に影響を与えている
122 \"source": "\\"大岡越前守\"", "target": "\\"心理的方法\"", "weight": 1, "description": "\\"大岡越前守は心理学が発達していない時代において
123 \"source": "\\"ボオ\"", "target": "\\"デュパン\"", "weight": 1, "description": "\\"デュパンはボオの作品『ル・モルグ』に登場するキャラクタ
124 \"source": "\\"ボオ\"", "target": "\\"ドイル\"", "weight": 1, "description": "\\"ドイルはボオの作品に影響を受けており、ホームズにデュパン
125 \"source": "\\"ドイル\"", "target": "\\"ホームズ\"", "weight": 1, "description": "\\"ホームズはドイルの作品『レジメント・ペーシェント』に
126 \"source": "\\"ホームズ\"", "target": "\\"レジメント・ペーシェント\"", "weight": 1, "description": "\\"ホームズは『レジメント・ペーシェン
127 \"source": "\\"僕\"", "target": "\\"旭屋の主人\"", "weight": 1, "description": "\\"僕は旭屋の主人が殺人者であると考えている。\"", "text
128 \"source": "\\"旭屋の主人\"", "target": "\\"殺人罪\"", "weight": 1, "description": "\\"旭屋の主人は殺人罪を犯したと疑われている。\"", "text
129 \"source": "\\"旭屋の主人\"", "target": "\\"便所を借りた男\"", "weight": 1, "description": "\\"旭屋の主人は罪跡を隠すために便所を借りた男
130 \"source": "\\"旭屋の主人\"", "target": "\\"サード卿\"", "weight": 1, "description": "\\"旭屋の主人 is connected to サード卿 through
131 \"source": "\\"旭屋の主人\"", "target": "\\"女のマゾッホ\"", "weight": 1, "description": "\\"旭屋の主人 discovered 女のマゾッホ next
132 \"source": "\\"旭屋の主人\"", "target": "\\"古本屋の細君\"", "weight": 1, "description": "\\"旭屋の主人 and 古本屋の細君 both have sa
133 \"source": "\\"古本屋の細君\"", "target": "\\"旭屋の細君\"", "weight": 2, "description": "\\"古本屋の細君と旭屋の細君は、共に生傷を持ち、同
```

図 1: GraphRAG 生成の関係データ（「旭屋の主人」に関する抜粋）

GraphRAG は、知識グラフ上に存在するエンティティ（旭屋の主人）とその関連性（犯行の動機、行動、他者との関係）を統合的に捉え、文中の因果構造を明快に説明する能力を示した。図 1 に示されるように、旭屋の主人は「殺人罪」（関係 127）や「便所を借りた男」（関係 128）、「女のマゾッホ」（関係 130）といった複数の関係を通じて描写されており、これらの情報は彼の犯行を示唆する重要な手がかりとなっている。

たとえば、関係 128 において旭屋の主人は「犯罪を隠すために便所を借りた男」として記述されており、これは物語内での行動（証拠隠滅）と結びつけられている。また、関係 130 では「女のマゾッホ」との発見的関係が記述され、加害と被害の単純な二元対立を超えた人間関係が浮かび上がる。さらに注目すべきなのは、これらの関係が自然言語で記述された説明文（description）と共に記録されている点である。たとえば、関係 127 には「旭屋の主人は殺人罪を犯したと疑われている」という記述が明示されており、モデルの解釈が単なる構造的リンクにとどまらず、意味的・語用論的含意を持つものであることがわかる。

このように、GraphRAG は従来のベクトル類似度依存の RAG とは異なり、知識の構造化と自然言語理解を組み合わせることで、物語内の因果関係や人間関係の重層的構造を的確に捉えることが可能である。

4.3 Gephi による知識グラフの可視化

GraphRAG によって自動生成された知識グラフは、物語内の登場人物、事件、場所などをノードとして構成し、それらの関係性をエッジで結ぶネットワーク構造を形成している。本節では、GraphRAG の出力ファイル `summarized_graph.graphml` を用いて、オープンソースの可視化ツール Gephi によりその構造を視覚化し、物語全体の意味的構造と登場人物の中心性を分析した。

データ - テーブル ×								
ノード	辺	設定	ノードの追加	辺の追加	検索/置換	スプレッドシートのインポート	テーブルのエクスポート	作業を継続
Source	Target	Type	Id	Label	Interval	Weight	フィルタ:	Source
・白梅軒	「明智小五郎」	無向	0		1.0	9d25989cccf6bb60650864...	「明智小五郎は白梅軒で...	
・白梅軒	「古木屋」	無向	1		1.0	9d25989cccf6bb60650864...	「古木屋は白梅軒の向か...	
・白梅軒	「小林軒事」	無向	2		1.0	0a61aca2a278f05bc05...	「小林軒事は白梅軒から...	
・白梅軒	「アスクリーム屋」	無向	3		1.0	0a61aca2a278f05bc05...	「アスクリーム屋は白梅軒...	
・白梅軒	「堺」	無向	4		1.0	0951434e00191d1f1f4e4895...	「堺は白梅軒から遠い」	
・白梅軒	「君」	無向	5		1.0	0951434e00191d1f1f4e4895...	「君は堺から遠い」	
・古木屋	「明智小五郎」	無向	6		1.0	9d25989cccf6bb60650864...	「明智小五郎は古木屋に...	
・古木屋	「古木屋の女房」	無向	7		1.0	9d25989cccf6bb60650864...	「古木屋の女房は古木屋...	
・古木屋	「剛子」	無向	8		1.0	5057039e0a03370707...	「古木屋の剛子は古木屋...	
・古木屋	「私」	無向	9		1.0	7d73914b538498ab-ed46...	「古木屋の私たちは古木屋...	
・古木屋	「彼」	無向	10		1.0	0f6532cf2fcd3e3b4d...	「彼は古木屋で売られた...	
・古木屋	「彼女」	無向	11		1.0	0f6532cf2fcd3e3b4d...	「彼女は古木屋で売られた...	
・古木屋	「上野の広小路」	無向	12		1.0	d73d62d62d07e1b1d...	「古木屋は上野の広小...	
・古木屋	「事件」	無向	13		2.0	229d95805062ab0a0fb5f7...	「古木屋の事件は古木屋に...	
・古木屋	「加屋」	無向	14		2.0	25243ab34bdc7144deef9...	「古木屋の加屋は加屋で接続して...	
・古木屋	「時計屋」	無向	15		1.0	25243ab34bdc7144deef9...	「古木屋は時計屋の間に...	
・古木屋	「裏子屋」	無向	16		1.0	25243ab34bdc7144deef9...	「古木屋は裏子屋の間に...	
・古木屋	「袋屋」	無向	17		1.0	25243ab34bdc7144deef9...	「古木屋は袋屋の間に...	
・明智小五郎	「古木屋の女房」	無向	18		1.0	5087820e0a13370707...	「明智小五郎と古木屋の...	
・明智小五郎	「宝物の山」	無向	19		1.0	ab4147228cc670f2a91...	「明智小五郎の宝物の山には...	
・明智小五郎	「男」	無向	20		1.0	60e2c52d62d07e1b1d...	「明智小五郎は男の手を...	
・明智小五郎	「腰り手」	無向	21		1.0	ec3852b9475a5de5e...	「腰の手は明智小五郎の腰...	
・明智小五郎	「障子」	無向	22		1.0	7d73914b538498ab-ed46...	「九の高」の無い早い時間に...	
・明智小五郎	「九月」	無向	23		1.0	b40af78fb99abc91989c...	「男の腰のさき間に立つ」	
・古木屋のお神さん	「加屋のお神さん」	無向	24		1.0	5976947768b798...	「同じともし手の腰に傷が...	
・谷本潤一郎	「本屋」	無向	25		1.0	061360191d9d8050a297...	「谷本潤一郎の作品」	
・谷本潤一郎	「明智」	無向	26		1.0	7d79c947d7c51c01e...	「谷本潤一郎は明智の一族」	
・谷本潤一郎	「『達上』」	無向	27		1.0	7d79c947d7c51c01e...	「『達上』は谷本潤一郎の...」	
・明智	「私」	無向	28		3.0	16330a5442d1b941247...	「明智と私(手札)は手札...	
・明智	「警察」	無向	29		2.0	a21b1e671b8899c40...	「明智は死を免れました後...	
・明智	「事件」	無向	30		4.0	9c760abf950ea4ced9f2...	「明智は、事件の詳細を聴...	
・明智	「智」	無向	31		1.0	165839477672b8ab050b...	「智は聴いていた」	
・明智	「彼」	無向	32		1.0	185ed179672b8ab050b...	「彼(明智)は部屋の足を...」	
・明智	「神田伯龍」	無向	33		1.0	0661af124d74781688...	「明智は頭のさきや声が...」	
・明智	「D坂の事件」	無向	34		1.0	999446767d8501adec...	「明智はD坂の事件につ...	
・明智	「君」	無向	35		1.0	187aab5e36a3cae43...	「明智は君の推理につい...」	

図2: エッジテーブル

Gephi では、知識グラフを構成するノードおよびエッジの詳細情報を表形式で確認することができる。エッジテーブル（図2）には、Source（始点ノード）、Target（終点ノード）、および両者間の関係の強さを示す Weight が記録されており、これらの情報が視覚的なネットワーク描画の基盤となっている。

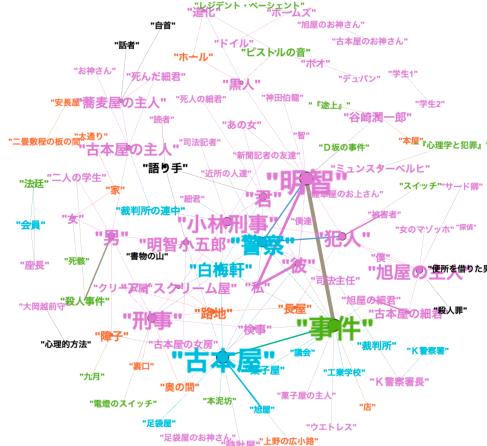


図3：全体ネットワーク図

図3は、Gephiによって可視化された知識グラフ全体の構造を示している。ノードのサイズは、そのノードに接続しているエッジの数に比例しており、サイズが大きいほど多くの関係がそのノードを中心に関連していることを意味する。また、エッジの太さ（weight）は関係の強度を示し、太い線ほど関係性が密接であることを表している。ノードの色は自動検出されたクラスタに対応しており、ネットワーク内のサブコミュニティ構造を視覚的に把握できる。図3で特に注目されるのは、「事件」「明智」「警察」「犯人」「旭屋の主人」といったノードがネットワークの中央付近に密集し、いずれも高い中心性と接続密度を有している点である。こ

これらのノードは、物語の中核的な構造を形成する重要な要素であることが視覚的にも確認できる。

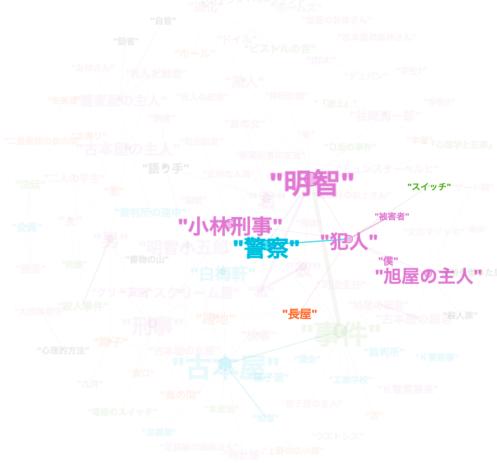


図4：犯人をめぐるネットワークグラフ

図4は、全体グラフから「犯人」に関するノードを中心として抽出・拡大したサブグラフである。この図からは、犯人が「明智」「警察」「小林刑事」「旭屋の主人」「被害者」「僕」などのノードと接続されていることが確認でき、特に旭屋の主人が主要な関係ノードとして現れており、物語中で犯人として疑われる立場にあることが構造的にも示唆されている。

このような知識グラフの可視化は、自然言語テキストの記述や構文解析だけでは捉えにくい人物間の関係性の強度や物語における構造的中心性を直感的に把握するうえで極めて有効であり、デジタル人文学における分析手法として大きな可能性を有している。

データ・テーブル×					
ノード	辺	設定	ノードの追加	辺の追加	検索/置換
		スプレッドシートのインポート			
Id	Label	Interval	source_id	description	type
「白梅軒」	「白梅軒」		0a67a2a52a278058bc549af7d58d8cc6...	白梅軒は、君が帰り道で立ち寄った場...	「ORGANIZATION」
「古本屋」	「古本屋」		0fa5352cf2f2ddcd3b3d031769e079911...	古本屋は、事件が発生した場所であり...	「ORGANIZATION」
「明智小五郎」	「明智小五郎」		50878206d0a133f7072f9298e7e92160...	明智小五郎は、主人公が白梅軒であり...	「PERSON」
「古本屋の女房」	「古本屋の女房」		50878206d0a133f7072f9298e7e92160...	「古本屋の女房は明智小五郎の副乳栄...」	「PERSON」
「障子」	「障子」		7db734bd538496ab8ed466872a4a6e4f...	障子は、専門家の間では無名と称され...	「GEO」
「九月」	「九月」		7db734bd538496ab8ed466872a4a6e4f...	「九月は、夏の終わりから秋の始まりに...」	「EVENT」
「古本屋のお神さん」	「古本屋のお神さん」		597b7c6aaefc9919875265a87c7c159c2...	「古本屋のお神さんは純善な人が、身...	「PERSON」
「旭屋のお神さん」	「旭屋のお神さん」		597b7c6aaefc9919875265a87c7c159c2...	「旭屋のお神さんもよく備をしており、...」	「PERSON」
「旭屋の御神さん」	「旭屋の御神さん」		06136b019d9d8050a29712b500ea2a8792...	「古本屋一郎は日本の著名な作家であり...」	「PERSON」
「谷崎潤一郎」	「谷崎潤一郎」		06136b019d9d8050a29712b500ea2a8792...	「谷崎潤一郎は、物語の中で登場人物が読めい...」	「GEO」
「本屋」	「本屋」		0661a124df74781a68b1254c20cc7bca...	「明智」は物語の登場人物であり、二階...	「PERSON」
「明智」	「明智」		0fa5352cf2f2ddcd3b3d031769e079911...	「明智」は、専門家の間では無名と称され...	「PERSON」
「私」	「私」		7d79c947d7c51cc04bd509f448ba...	「『途上』は谷崎潤一郎の小説で、作中...」	「EVENT」
「『途上』」	「『途上』」		0fa5352cf2f2ddcd3b3d031769e079911...	「探偵は犯事を解決する能力を持つ人物...」	「PERSON」
「探偵」	「探偵」		0fa5352cf2f2ddcd3b3d031769e079911...	「彼は、殺人事件の中心人物であり、犯...」	「PERSON」
「彼」	「彼」		0fa5352cf2f2ddcd3b3d031769e079911...	「0fa5352cf2f2ddcd3b3d031769e079911...」	「PERSON」
「本泥坊」	「本泥坊」		5963a919e0cd9299ac9e4b6e5d830d...	「本泥坊は古本屋で発生した事件であ...」	「EVENT」
「警察」	「警察」		ca32c2330a13731609a074cd24461ad...	「警察は指紋の検査を行なう組織です...」	「ORGANIZATION」
「細君」	「細君」		1bdb4bcfa5f332981d1b2e233cd3cd10...	「細君は細君の内に隠れて死んでいた女性で...」	「PERSON」
「二暮敷板の板の間」	「二暮敷板の板の間」		1bdb4bcfa5f332981d1b2e233cd3cd10...	「二暮敷板の板の間は、裏方に接して致...」	「GEO」
「安長屋」	「安長屋」		1bdb4bcfa5f332981d1b2e233cd3cd10...	「安長屋は、ごくありふれた間戸の建物...」	「GEO」
「死骸」	「死骸」		0b9a792b952c3ac3a93f93cb86fe996...	「死骸は、左側の壁寄りに、店の間の方...」	「EVENT」
「女」	「女」		0b9a792b952c3ac3a93f93cb86fe996...	「女」は男性に襲撃された被害者であり...」	「PERSON」
「大通り」	「大通り」		0b9a792b952c3ac3a93f93cb86fe996...	「大通りは表にあり、往来が絶えない場...」	「GEO」
「家」	「家」		0b9a792b952c3ac3a93f93cb86fe996...	「家は障子一重で、一人の女の女が惨殺され...」	「GEO」
「上野の広小路」	「上野の広小路」		d73dbf62d07ee41b42ce95ea8fe1b77...	「上野の広小路は、夜店がよく出る場所...」	「GEO」
「足袋屋のお神さん」	「足袋屋のお神さん」		d73dbf62d07ee41b42ce95ea8fe1b77...	「足袋屋のお神さんは、勝家に任んでお...」	「PERSON」
「時計屋」	「時計屋」		2524343b4bc7f144de9e201744349...	「時計屋」は事件の際に足袋屋のお...」	「PERSON」
「古本屋の主人」	「古本屋の主人」		3fa5d5ab5e541242d2d5aa047b5459...	「古本屋の主人は、事件の中心にいる...」	「PERSON」
「お神さん」	「お神さん」		3fa5d5ab5e541242d2d5aa047b5459...	「お神さんは、宿屋の足袋屋の人物で、時...」	「PERSON」
「K警察署長」	「K警察署長」		072a2e0073e931c94b1877c21dece186...	「K警察署長」は警察署の署長であり、...」	「PERSON」
「小林刑事」	「小林刑事」		072a2e0073e931c94b1877c21dece186...	「小林刑事は、指紋の検査を依頼された...」	「PERSON」
「裁判所の連中」	「裁判所の連中」		3fa5d5ab5e541242d2d5aa047b5459...	「裁判所の連中は警察からの急報で駆け...」	「ORGANIZATION」
「K警察署」	「K警察署」		3fa5d5ab5e541242d2d5aa047b5459...	「K警察署はK警察署が所属する警察...」	「ORGANIZATION」
「裁判所」	「裁判所」		072a2e0073e931c94b1877c21dece186...	「裁判所は法的な問題を扱う機関であり...」	「ORGANIZATION」
「司法記者」	「司法記者」		072a2e0073e931c94b1877c21dece186...	「司法記者は法的な事件を報道する記者...」	「PERSON」
「司法主任」	「司法主任」		072a2e0073e931c94b1877c21dece186...	「司法主任は法的な問題を担当する主任...」	「PERSON」

図5：ノードテーブル

一方、ノードテーブル（図5）では、各ノードに割り当てられた type（例：PERSON, GEO, EVENT）や description（自然言語による説明）などのメタ情報が確認できるが、これらの属性情報を活用した検索や条件付き可視化には、次節で述べる Neo4j による探索がより適している。

4.3 Neo4j による知識グラフの可視化

Neo4j は、グラフデータを効率的に格納、可視化するためのグラフデータベース管理システムであり、エンティティとリレーションを高度に構造化された形式で保持することが可能である。本研究では、GraphRAG が output したファイルを Neo4j にインポートし、知識グラフを属性ベースで動的に探索する。

Gephi によるネットワークが主に関係構造の可視化とクラスタ分析に特化しているのに対し、Neo4j では各ノードやリレーションに付与された属性情報をもとに、クエリベースで特定のサブグラフや関係構造を抽出することができる。また、Cypher と呼ばれるクエリ言語を使用することで、複雑なパターンマッチングや階層的構造の探索も容易であり、文学作品内の因果関係や登場人物の役割分担といった分析に有用である。

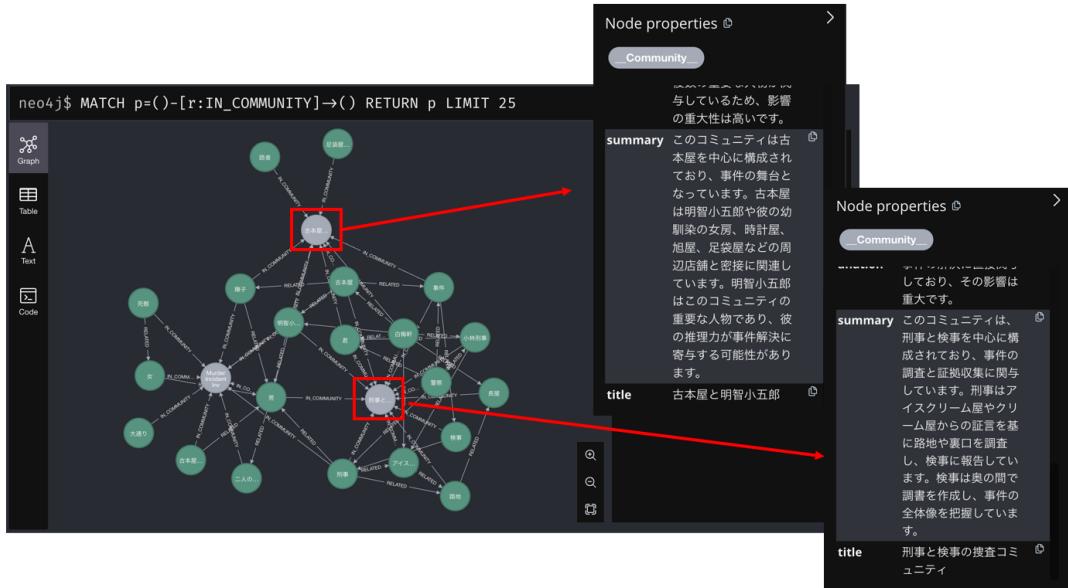


図 6 : Neo4j におけるコミュニティ構造

図 6 は、Neo4j 上で IN_COMMUNITY リレーションを用いて知識グラフのコミュニティ構造を抽出・可視化したものである。ここでは、登場人物や事件がどのような社会的・機能的なグループに属しているかを視覚的に把握することができる。

図中では、例えば「古本屋を中心に構成されたコミュニティ」と「刑事と検事の捜査に関連するコミュニティ」が確認され、それぞれのノードに付与された summary プロパティには、

そのコミュニティの構成要素や物語内での機能的役割が自然言語で記述されている。例として、古本屋を中心とするコミュニティには明智小五郎や旭屋、時計屋、足袋屋などの情報が含まれ、物語の舞台設定や犯行現場との関連性が明示されている。

一方、刑事を中心とするもう一つのコミュニティでは、アイスクリーム屋からの証言や捜査記録の整理が記されており、事件の調査と証拠構築における役割が強調されている。このように、Neo4j ではノードに付与された summary や title などの属性情報をもとに、各コミュニティの語り機能や構造的役割を定性的に把握することが可能であり、Gephi では実現しにくい意味論的補足を加える分析が可能となる。



図 7：人物ノードの関係構造

図 7 は、Neo4j において Person ラベルを持つノードを抽出し、犯人を中心とする人物ネットワークを表示したものである。図 7 では、「犯人」ノードが中心的な位置を占め、他の登場人物（警察、小林刑事、僕、彼など）と多数の関係性を有していることが視覚的に確認できる。

また、「旭屋の主人」に着目すると、そのノードには description 属性として「彼は殺人者であり、罪跡を隠すために便所を借りた男のことを話している」といった詳細な自然言語の記述が付与されている。これは、GraphRAG がテキスト中の人物描写や行動を解析し、意味的に解釈可能なプロファイルとして構造化していることを示している。

図 8 は、Neo4j において Geo ラベルを持つノードを抽出し、物語内に登場する地理的要素（場所）の構造を視覚的に表示したものである。図 8 では、「奥の間」や「裏口」「路地」などがノードとして確認され、それぞれの空間がどのような登場人物や事件と関係しているかが可視化されている。

たとえば「奥の間」は、図の中央付近に表示されており、「刑事と捜査のコミュニティ」や「犯人」などとの関係性を持つことが確認できる。右側の属性ビューに示されるように、このノードには description 属性として「奥の間は死体が発見された場所で、検事が陣取っていた部屋」といった事件における場所の役割が記述されており、物語内の犯罪構造を空間的に補完

する情報として機能している。このように、Geo ノードの探索を通じて、事件の発生場所や登場人物の行動軌跡を視覚的に把握することで、物語内の空間的ナラティブ構造をより立体的に読み解く手がかりが得られる。

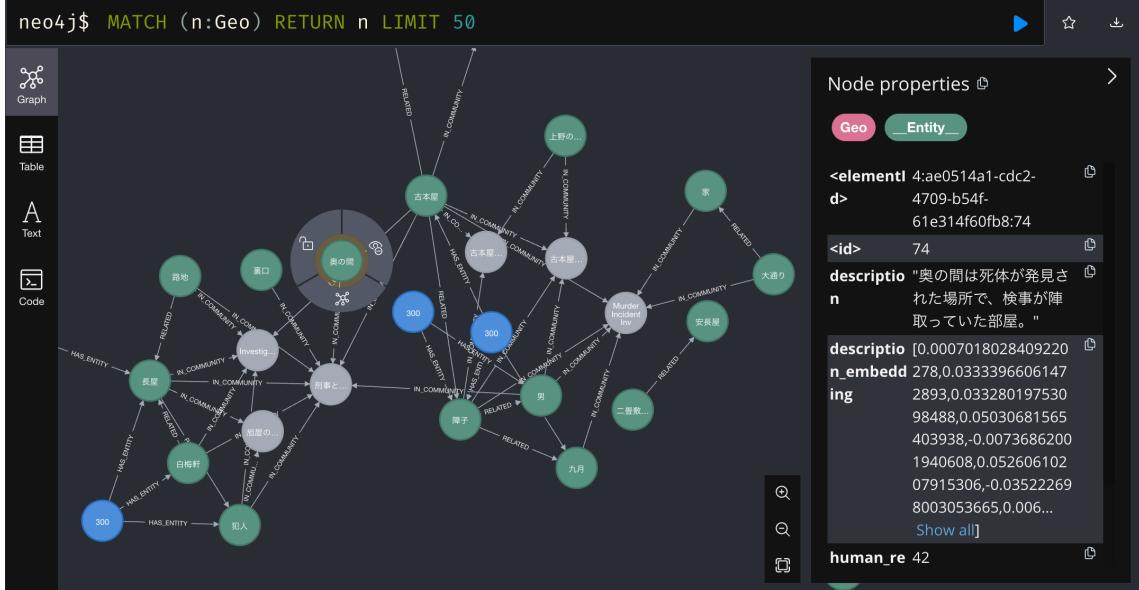


図 8：地理ノードの関係構造

本節で示したように、Neo4j は GraphRAG によって構築された知識グラフを対話的かつ属性ベースで探索する上で、極めて有効な手段である。特に、ノードやリレーションに付与された type、description、summary といったメタ情報を活用することで、登場人物の役割、事件との関係性、物語内での空間的な意味づけなどを細部まで精緻に読み解くことが可能となった。

5. おわりに

本研究では、GraphRAG を江戸川乱歩の短編小説『D 坂の殺人事件』に適用し、従来の RAG では把握が難しかった構造的・意味的関係の抽出に成功し、物語内のテーマ分析や犯人特定といった問い合わせに対して、高い精度と多角的な応答を示した。また、Gephi や Neo4j といった可視化ツールとの併用により、抽出された知識グラフの構造を視覚的に把握することが可能となり、文学作品の解釈支援手法としての有用性も確認された。

また、GraphRAG は理論的にはより長大なテキストにも適用可能であり、たとえば江戸川乱歩の連作シリーズや現代小説を対象とした分析において、登場人物の変遷や物語構造の展開、主題の深化といった要素を構造的に追跡する応用が期待される。そのためには、段階的なインデックス生成や部分的再構築といった処理の最適化が今後の課題となる。また、複数作品間の知識グラフを統合し、物語世界を横断的に分析する読解支援手法への発展も視野に入れられるだろう。

文献

1. Chandak, P., Huang, K., & Zitnik, M. (2023). Building a knowledge graph to enable precision medicine. *Scientific Data*, 10(1), 67.
2. Chai, Z., Zhang, T., Wu, L., Han, K., Hu, X., Huang, X., & Yang, Y. (2023). GraphLLM: Boosting graph reasoning ability of large language model. *arXiv preprint arXiv:2310.05845*.
3. Chen, J., Lin, H., Han, X., & Sun, L. (2023). Benchmarking large language models in retrieval-augmented generation. *arXiv preprint arXiv:2309.01431*.
4. Cvitkovic, M. (2020). Supervised learning on relational databases with graph neural networks. *arXiv preprint arXiv:2002.02046*.
5. Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. *arXiv preprint arXiv:2312.10997*.
6. Gupta, S., Ranjan, R., & Singh, S. N. (2024). A comprehensive survey of retrieval-augmented generation (RAG): Evolution, current landscape and future directions. *arXiv preprint arXiv:2410.12837*.
7. Han, H., Shomer, H., Wang, Y., Lei, Y., Guo, K., Hua, Z., Long, B., Liu, H., & Tang, J. (2025). RAG vs. GraphRAG: A systematic evaluation and key insights. *arXiv preprint arXiv:2502.11371*.
8. Han, H., Wang, Y., Shomer, H., Guo, K., Ding, J., Lei, Y., Halappanavar, M., Rossi, R. A., Mukherjee, S., Tang, X., et al. (2024). Retrieval-augmented generation with graphs (GraphRAG). *arXiv preprint arXiv:2501.00309*.
9. Han, H., Xie, Y., Liu, H., Tang, X., Nag, S., Headden, W., Li, Y., Luo, C., Ji, S., He, Q., et al. (2025). Reasoning with graphs: Structuring implicit knowledge to enhance LLMs reasoning. *arXiv preprint arXiv:2501.07845*.
10. Hu, Y., Lei, Z., Zhang, Z., Pan, B., Ling, C., & Zhao, L. (2024). GRAG: Graph retrieval-augmented generation. *arXiv preprint arXiv:2405.16506*.
11. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.-T., Rocktäschel, T., et al. (2021). Retrieval-augmented generation for knowledge-intensive NLP tasks. *arXiv preprint arXiv:2005.11401*.
12. Ma, X., Gong, Y., He, P., Zhao, H., & Duan, N. (2023). Query rewriting for retrieval-augmented large language models. *arXiv preprint arXiv:2305.14283*.
13. Peng, B., Zhu, Y., Liu, Y., Bo, X., Shi, H., Hong, C., Zhang, Y., & Tang, S. (2024). Graph retrieval-augmented generation: A survey. *arXiv preprint arXiv:2408.08921*.
14. Ram, O., Levine, Y., Dalmedigos, I., Muhlgay, D., Shashua, A., Leyton-Brown, K., & Shoham, Y. (2023). In-context retrieval-augmented language models. *arXiv preprint arXiv:2302.00083*.
15. Xia, F., Sun, K., Yu, S., Aziz, A., Wan, L., Pan, S., & Liu, H. (2021). Graph learning: A survey. *arXiv preprint arXiv:2105.00696*.