

機関番号	研究種目番号	応募区分番号	小区分	整理番号
13901	05	1	90020	0002

令和7(2025)年度 基盤研究(B)(一般)研究計画調書

令和 6年 9月11日
1版

新規

研究種目	基盤研究(B)	応募区分	一般
小区分	図書館情報学および人文社会情報学関連		
研究代表者 氏名	(フリガナ)	コマミズ タカヒロ	
	(漢字等)	駒水 孝裕	
所属研究機関	名古屋大学		
部 局	数理・データ科学教育研究センター		
職	准教授		
研究課題名	異種データを活用した高精度な知識抽出と提供のための情報統合基盤の研究		
研究経費 (千円未満の 端数は切り 捨てる)	年度	研究経費 (千円)	使用内訳(千円)
			設備備品費 消耗品費 旅費 人件費・謝金 その他
	令和7年度	6,030	2,050 130 2,030 300 1,520
	令和8年度	5,300	1,000 150 2,030 600 1,520
	令和9年度	4,520	250 120 2,030 600 1,520
	令和10年度	4,150	0 0 2,030 600 1,520
	令和11年度	0	0 0 0 0 0
総計	20,000	3,300 400 8,120 2,100 6,080	
開示希望の有無	審査結果の開示を希望する		
研究計画最終年度前年度応募	--		

研究組織（研究代表者及び研究分担者）

	氏名（年齢）	所属研究機関 部局 職	学位 役割分担	令和7年度 研究経費 （千円）	エフオ ート （%）
研究代表者	コマミズ タカヒロ	名古屋大学 数理・データ科学教育研究センター	博士（工学）	4,030	20
	駒水 孝裕	准教授	研究統括		
研究分担者	イデ イチロウ	名古屋大学 情報学研究科	博士（工学）	500	5
	井手 一郎	教授	マルチメディアコンテンツ処理技術の研究開発		
研究分担者	イシカワ ヨシハル	名古屋大学 情報学研究科	博士（工学）	500	7
	石川 佳治	教授	効率的なデータ管理に関する研究開発		
研究分担者	ハタノ ケンジ	同志社大学 文化情報学部	博士（工学）	500	5
	波多野 賢治	教授	異種データ検索技術の研究開発		
研究分担者	カストナー マーク アウレル	広島市立大学 情報科学研究科	博士（情報学）	500	10
	KASTNER Ma rc Aurel	助教	マルチメディアコンテンツ処理技術の研究開発		
合計 5 名			研究経費合計	6,030	

1 研究目的、研究方法など

本研究計画調書は「小区分」の審査区分で審査される。記述に当たっては、「科学研究費助成事業における審査及び評価に関する規程」（公募要領参照）を参考にすること。

本研究の目的と方法などについて、5頁以内で記述すること。

冒頭にその概要を簡潔にまとめて記述し、本文には、(1)本研究の学術的背景や本研究の着想に至った経緯、研究課題の核心をなす学術的「問い」、(2)本研究の目的および学術的独自性と創造性、(3)関連分野の研究動向と本研究の位置づけ、(4)本研究で何をどのように、どこまで明らかにしようとするのか、(5)本研究の目的を達成するための準備状況、(6)本研究がどのような国際性（将来的に世界の研究をけん引する、協同を通じて世界の研究の発展に貢献する、我が国独自の研究としての高い価値を創出する等）を有するかについて具体的かつ明確に記述すること。

本研究を研究分担者とともにを行う場合は、研究代表者、研究分担者の具体的な役割を記述すること。

（概要）

オープンデータは得られた知見や主張の透明性・再現性を担保できるため、社会的に重要な動きである。一方で、生成系 AI がその有用性を実証し始め、人々の情報アクセス方法を変革しつつある。しかし、事実とは異なる情報を尤もらしく提示する「幻覚」の問題が指摘されており、RAG (Retrieval-Augmented Generation) がその対策として注目されている。グラフデータを情報源とする RAG (GraphRAG) はより広範な知識を生成系 AI に提供することができるが、テキスト以外のデータや異種データ混在時の挙動が明らかでない。本研究は、先行課題（科研費・基盤 (B) 21H03555) を踏襲し、テキストだけでなく画像や映像からなる異種オープンデータを Linked Open Data の枠組みで知識グラフ化する。その上で、RAG の仕組みをベースとした生成系 AI を用いた情報提供の枠組みを提案する。特に、知識グラフを情報源とした RAG の挙動を明らかにしつつ、正確な情報を提供するために、知識グラフ構造化と生成系 AI の適用方法を追究する。

（本文）

(1) 本研究の学術的背景、研究課題の核心をなす学術的「問い」

【学術的背景】

インターネットの普及や各種デバイスの大衆化により、電子的なデータは急速に増加している。データの急増と計算機デバイスの高性能化・低廉化に後押しされ、人工知能技術を筆頭にデータ科学技術は大きく進歩してきた。この背景には、オープンサイエンスやオープンガバメントなどの根底にあるオープンデータの動きがある。データをオープン化することで、得られた知見や知識、主張の透明性や再現性を担保できるため、オープンデータは社会的に重要な動きである。さらに、データの相互利用性を高めるために、Linked Open Data (LOD) という枠組みが World Wide Web を提唱した Tim Berners-Lee 氏によって提唱されている。

オープンデータの多くは構造化されたテキストであるが、画像や映像などのマルチメディアもまた重要な知識を含むデータである。例えば、人物や建物の画像は被写体の特徴を視覚的に表現し、自然現象や事件・事故を観測した画像や映像は起こった事象を視覚・聴覚的に記録できる。LOD はテキストで記述される知識を構造化するため、マルチメディアから抽出した（テキスト化された）知識は扱えるが、テキスト化が困難な機微な知識までは表現することができない。これに対し、マルチメディア自体を知識として LOD に取り込み、マルチメディアとテキストを相互利用できる環境が実現されれば、オープンデータを活用したアプリケーションの作成や異種データを扱う技術の研究開発をより促進することが期待される。そのためには、オープンデータ管理においてテキストをもとに構造化されたデータに限定した方法論から脱却しなければならない。

一方で、大規模言語モデル (LLM; Large Language Model) を皮切りに、生成系 AI (ChatGPT¹ など) と呼ばれるシステムがその有用性を実証し始め、人々の情報アクセスの方法を変革しつつある。さらに、マルチモーダル基盤モデル (MMFM; Multi-modal Foundation Model) と呼ばれる異なる種類のデータを同一の特徴空間上で扱えるようなモデル (CLIP [1] など) も登場し、テキスト以外の画像などのデータも同時に扱う技術が発展してきた。これらの発展により、チャットボットのような対話型のインターフェースを用いて、情報技術に関する知識を必要とすること

¹<https://chatgpt.com/>

【1 研究目的、研究方法など (つづき)】

なく、様々な高度なタスクを容易に実行できる環境が整ってきた。しかし、生成系 AI は、事実とは異なる情報をもっともらしく提示してしまう幻覚 (Hallucination) の問題が指摘されている [2]。生成系 AI が学習データに含まれない事実に関して幻覚をしばしば起こすため、それに対応するために RAG (Retrieval-Augmented Generation) [3] が提案され、幻覚対策としても研究されている [4]。しかし、RAG は情報源における関連要素の出現箇所が遠い場合に、情報源の内容をうまく活用できないことが指摘されており、情報源の内容をグラフ構造で表現し RAG を実現する GraphRAG [5] が提案されている。このことから、情報源の内容をグラフで表現する知識グラフは依然として重要な役割を果たす。一方で、GraphRAG はテキストに対する手法であり、異種データに対する有効性は明らかとなっておらず、幻覚などの問題が懸念される MMFM における RAG の方法論の確立は重要な役割を果たすことが期待される。

以上の背景から、本研究における学術的な「問い」は以下の通りである。

学術的「問い」

多様なメディアから知識を抽出・構造化し、正確な情報の提供方法が構築可能か

(2) 本研究の目的および学術的独自性と創造性

【本研究の目的：Linked Open Multimedia Data の構築と活用可能性の検証】

異種データを情報源とする RAG を実現するためには、知識グラフの構築・利用について以下の課題がある (右図)。これまでに、テキストから出現するエンティティ同士の関係を抽出する技術 (情報抽出 [6] など) や画像中から出現するオブジェクト同士の関係を抽出する技術 (シーングラフ生成 [7] など) などの情報抽出技術が研究されてきた。その中には、個体であるエンティティの関係を抽出する技術や個体ではなくエンティティの種類 (人や机など) の関係を抽出する技術があり、対象のデータや技術により抽出できる情報の粒度が異なる。このような知識グラフが RAG の情報源として有効であるかは明らかでない上に、どの程度精緻な知識グラフが適切であるかも定かではない。また、エンティティ同士の関係性が明確にテキストで記述されており、それに基づく知識グラフを用いることは有効である可能性は高いものの、テキスト以外のデータ (画像や動画) において、知識グラフの形態自体が有効であるかは定かではない。上述のように、マルチメディアの内容は必ずしもテキストで表現できるとは限らない。したがって、マルチメディアを直接的に知識グラフに取り込んだグラフ構造などのより高度なグラフ構造などの可能性を模索する必要がある。

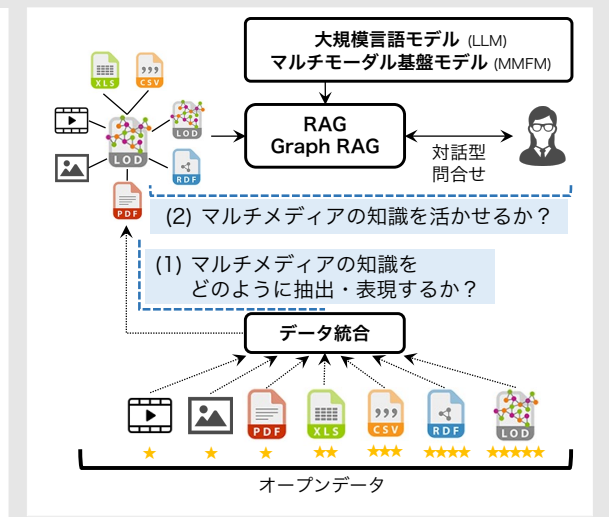


図 1: Linked Open Multimedia Data の構築と活用

図 1: Linked Open Multimedia Data の構築と活用

【学術的独自性と創造性】

本研究は、さまざまなフォーマットで公開されるデータを統合的に扱うためのデータ管理基盤に根ざすものである。特に、マルチメディアデータを知識グラフとして管理する点において、本研究は独自性を前の研究課題 (科学研究費補助金・基盤研究 (B) (21H03555)) から継承しているだけでなく、知識グラフの最適な構造を探る点でより発展的・創造的である。また、生成系 AI の可用性を更に拡大し得る点についても、創造的である。これらのことから、本研究が目的とするマルチメディアオープンデータの構造化およびその応用は LOD 研究だけでなくデータを統合的に扱うさまざまな研究や開発、および個々のデータ処理技術の研究開発において重要な分野となり、加えて、生成系 AI の可用性・信頼性の発展にも寄与することが期待できる。

【1 研究目的、研究方法など (つづき)】

(3) 本研究の着想に至った経緯と関連する国内外の研究動向と本研究の位置づけ

【本研究の着想に至った経緯】

オープンデータ利活用技術の発展は、データ公開後の再利用性の見通しを立てるために重要である。申請者は、異種オープンデータを効果的・効率的に管理・活用するために、前の研究課題「科学研究費補助金・基盤研究 (B) (21H03555)」にて、異種データの統合・管理および活用に関して、Linked Open Multimedia Data Management 領域を提唱し、データ統合およびデータ管理に関する研究を実施してきた。一方で、昨今の生成系 AI の台頭によるユーザの情報アクセスに対する変化から、アプリケーション開発部分において、生成系 AI の利便性との調和を検討するようになった。しかし、上述のように生成系 AI は、正確な情報取得という観点において、テキストで表現されたデータですら完璧には扱えておらず、画像や動画のような他メディアのデータを扱えるかどうかは、依然として未知の部分が多い。したがって、上記目的に対し、異種オープンデータからの知識グラフ構築を継続しつつ、生成系 AI (特に GraphRAG) において、構築した知識グラフを効果的に活用する方法について研究する本研究の着想に至った。

【関連する国内外の研究動向】

本研究の知識グラフ構築部分に最も関連する国際会議としては、セマンティックウェブに関する会議 ISWC がある。ISWC では、マルチメディアに関する議論はほとんどされず、標準的な LOD や知識グラフからの推論などの議論が中心となっている。また、LOD とテキストにおける横断的な研究は一時期行われていたものの、特に目立った成果は出ておらず、未だに取り組むべき課題とされている。他に関連する学会として、データ工学に関する国際会議である SIGMOD, VLDB やマルチメディア情報処理に関する国際会議である ACM Multimedia や Multimedia Modeling などでは、データ統合や異種データを活用した情報処理に関する議論が行われている。類似する研究は、自然言語処理分野 (国際会議 ACL や EMNLP) や機械学習分野 (国際会議 NeurIPS や ICLR) でも議論されているが、画像など他種のデータにおける研究は限定的である [8]。特に、ドメイン固有のデータや画像から得られる知識は極めて限定的 (局所的) であるため、知識の量や幅に違いがあるデータ間での効果的なデータ統合・活用方法は明らかにされていない。

他方で、本研究の知識グラフの活用部分に関して、生成系 AI の台頭により、テキストや画像などからの知識抽出は大幅に進歩した。一方で、機械的な抽出により、誤った知識が抽出されることが懸念されており、この課題に対する解法は依然として確立していない [9]。加えて、生成系 AI は、事実とは異なる情報をもっともらしく提示してしまう幻覚 (Hallucination) の問題が指摘されており [2]、RAG (Retrieval-Augmented Generation) [3] が幻覚対策としても研究されている [4]。RAG は非常に注目を集める技術であり、多くの研究者が取り組んでいる。これに対して、本研究で活用するデータはグラフ構造であり、グラフデータを RAG の仕組みに取り入れる GraphRAG [5] は新しく、異種マルチメディアデータにおける可用性は研究されていない。

【本研究の位置づけ】

本研究は、先行課題にて提唱した Linked Open Multimedia Data Management 領域において、異種マルチメディアデータの統合により活用可能なデータを統合・管理し、正確な情報を提供するための枠組みを研究するものである。特に、シングルモーダルおよびマルチモーダルの事前学習済みモデルを効果的に活用し、データ統合・管理・活用に適用する方法論を確立する。また、異種マルチメディアデータに含まれる情報をより正確に提供できる枠組みを RAG を用いた生成系 AI を基礎として、データ構築および RAG におけるデータ処理の2つの点で研究する。これらの研究は、上記領域において基盤的な技術であるとともに、生成系 AI においてその利用可能性を検証する研究であり、これらを研究・検証することで、異種オープンデータの利活用の新たな方法論を展開することが期待される。

【1 研究目的、研究方法など (つづき)】

(4) 本研究で何をどのように、どこまで明らかにしようとするのか

【本研究で取り組む内容】

本研究では、データ統合、データ管理、データ活用からなる Linked Open Multimedia Data Management 領域に照らし合わせ、各項目に対応して、下記の内容に分割して研究に取り組む。

《**データ統合：マルチメディアからの情報抽出**》異種データを統合するために、各メディアのデータからの情報抽出および異なるメディア間のエンティティ同定が必要となる。申請者はこれまでに、画像からの情報抽出としてシーングラフと呼ばれる画像内容のグラフ表現の抽出技術を画像集合に対して拡張してきた [10]。加えて、異種データを統合的に検索する研究として、錠剤写真と処方箋 (画像とテキスト) に横断して、処方箋内の錠剤名と画像中の錠剤を同定する方法を開発した [11]。これらの研究から、(1) 画像に対するシーングラフ抽出において、画像の内容を十分に表現するグラフデータが得られていないこと、また、それに起因して、(2) 異種データ統合において、語彙のミスマッチの問題が限定的な条件でのみ解決されており、一般的には未解決な課題であること、の2点が明らかとなった。したがって、本研究では、これらの解決に取り組み異種データ統合技術を発展させる。具体的には、(1) より詳細な情報を保持したシーングラフ抽出手法の構築、その上で、(2) 構築された詳細なシーングラフをもとにデータ統合手法を構築する。

《**データ管理：検索の効率化**》本研究で構築する知識グラフは巨大なものとなり、RAG での活用においては、その一部を検索して利用することになる。この検索において、基本的に、各データ点はベクトルによって表現される。深層学習の出現以来、このベクトルは大量のデータによって学習されたモデルによって生成されたものを用いることが一般化してきている。申請者は、これまでに検索性能向上のために、最新研究の再ランキングモデルの一部を事前計算化することで、同等の性能を10倍以上高速を実現した [12, 13]。これらの研究により、検索速度の向上はある程度達成されたものの、検索性能も同時に高める改善が必要であることとベクトル間の類似性が常にそのまま利用すべきでないことが示唆された。したがって、本研究では、これまでの高速化手法を基礎として、次のアイデアにより高性能な検索を目指す。検索における「類似」とベクトル空間上の「類似」は必ずしも同じ意図での「類似」とは限らないことに着目する。異なる観点で様々な再ランキング手法が提案されてきており、それぞれに独自の「類似」の視点が考案されている。そのような手法の良さを適材適所かつ相乗効果的に利用できる仕組みを研究する。これによって、高速で高性能な検索を実現し、RAG における情報源の取得効率・効果を向上させる。

《**データ活用：RAG における知識グラフの利用方策**》本研究では、データ活用において、生成系 AI における GraphRAG に注目する。上述のように、GraphRAG は新しい技術であるため、その性質のほとんどが明らかではなく、特にテキスト以外のデータを取り込んだ場合の性質は未知である。したがって、本研究ではまず画像を対象にする。画像における知識抽出の技術としてシーングラフ生成があり、長らく研究され、高度な手法が多数出現してきている。このことから、画像に対するシーングラフは利用しやすく、RAG における性能を計測しやすいことから、RAG における画像データの利用可能性を検証できる。次に、シーングラフにおいて、各ノードは物体のクラスだけでなく、画像中の位置を表す包囲矩形 (Bounding Box) を持つ。この包囲矩形から物体自体の画像を得ることができるため、この画像を RAG に利用する手法を構築し、物体のクラスよりも情報量の多いデータを取り入れた場合の RAG の振る舞いを検証する。これらの課程を経ることで得られた知見をより広範なデータに展開していく。このようにして、異種データを RAG に取り込み、ユーザに正確な情報を提供できるかどうかを検証する。

【本研究に取り組む体制と研究代表者・分担者の役割】

本研究において、異種データを効果的・効率的に扱うことが求められるため、これらの知見を有する研究者を分担者として加えることで、提案領域の研究を加速する。**研究代表者・駒水**は、研究全体の統括を行い、各要素技術開発の具体化と遂行の役割を持つ。加えて、データ統合における要素技術に関する研究開発および実装・実験を行う。**研究分担者・井手**は、画像を中心とした

【1 研究目的、研究方法など (つづき)】

マルチメディアコンテンツ処理に明るいことから、画像や映像を対象とした技術開発を担当する。研究分担者・石川は、データベース技術に明るく、データがベクトル表現されたときに効率的に検索する技術に関する知見を有するため、効率的なデータ管理に関する開発を担当する。研究分担者・波多野は、構造化テキストに対する検索技術を中心に、テキストデータ処理や知識データ処理に明るいことから、効果的なデータの統合および管理方法の開発を担当する。研究分担者・Kastnerは、画像とテキストを横断したデータ処理に明るいことから、マルチモーダルなデータ処理に関する技術開発を担当する。なお、システム開発や実験については、研究員や研究代表者・分担者の研究室の博士前期・後期課程学生を研究補助者として雇用し、適宜分担する。

(5) 本研究の目的を達成するための準備状況

本研究で取り組む三点について、上述のようにすでに関連する研究を進めてきている。したがって、関連研究調査や研究資源などの準備は最低限整っている。一方で、関連する研究分野 (特に、生成系 AI 関連) の進歩は目覚ましいため、関連研究の調査および研究方向性の精査は継続的に行う必要がある。しかし、これまでも同様の状況で研究成果を着実に出してきたことから、本研究の研究においても柔軟に最新の研究に適応することで、成果を出すことは可能と考える。以上のように、本研究で取り組む内容に向けた準備は十分にできている。

(6) 本研究が有する国際性

本研究は、世界的に可用性を示してきた生成系 AI の発展に寄与する研究である。本研究の観点は、直接的に生成系 AI モデルを考案することではないものの、その可用性を、特にマルチメディアデータ活用の観点において、向上させることを目的としている。加えて、知識の構造化に取り組むことで、人により解釈可能な形での知識データの構築の必要性を主張するものである。昨今の深層表現学習の発展より、知識は人の解釈の及ばない特徴空間での表現を基本としようとしている。一方で、その解釈可能性については、深層表現学習モデルがブラックボックス的な振る舞いをするために、疑問視する声もある。本研究は、解釈可能性について直接研究するものではないが、知識を解釈可能な形 (グラフ構造) で表現する可能性を捨てず、生成系 AI における利用可能性を広げることが期待できる点で、世界の研究の発展に貢献することが期待できる。加えて、異種データ横断的な分析や検索に関する国際ワークショップ ICDAR [14] などを主催し、継続する予定であるため、本研究を国際的に発信し、関連研究者と議論できる環境を整えてきており、国際的な情報発信の基盤も整っている。

[参考文献] ※下線は研究代表者または研究分担者

- [1] Radford, et al., “Learning Transferable Visual Models From Natural Language Supervision”, ICML, pp.8748-8763, 2021
- [2] Ji, et al., “Survey of Hallucination in Natural Language Generation”, ACM Comput. Surv. 55(12)-248, pp.1-38, 2023
- [3] Lewis, et al., “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks”, NeurIPS, 2020
- [4] Yan et al., “Corrective Retrieval Augmented Generation”, CoRR abs/2401.15884, 2024
- [5] Edge, et al., “From Local to Global: A Graph RAG Approach to Query-Focused Summarization”, CoRR abs/2404.16130, 2024
- [6] Zhou, et al., “A Survey on Neural Open Information Extraction: Current Status and Future Directions”, IJCAI, pp.5694-5701, 2022
- [7] Li, et al., “Scene Graph Generation: A comprehensive survey”, Neurocomputing 566: 127052, 2024
- [8] Müller-Budack et al., “Ontology-driven Event Type Classification in Images”, WACV, pp.2927-2937, 2021
- [9] Pan et al., “Unifying Large Language Models and Knowledge Graphs: A Roadmap”, CoRR abs/2306.08302, 2023
- [10] I. Phueaksri, M. A. Kastner, Y. Kawanishi, T. Komamizu, I. Ide, “Image-Collection Summarization Using Scene-Graph Generation With External Knowledge”, IEEE Access, Vol.12, pp.17499-17512, 2024
- [11] T. T. Nguyen, P. L. Nguyen, Y. Kawanishi, T. Komamizu, I. Ide, “Zero-Shot Pill-Prescription Matching With Graph Convolutional Network and Contrastive Learning”, IEEE Access, Vol.12, pp.55889-55904, 2024
- [12] T. Kato, T. Komamizu, I. Ide, “R-DiP: Re-ranking Based Diffusion Pre-computation for Image Retrieval”, DEXA, pp.233-247, 2024 (**Best Paper Award**)
- [13] T. Komamizu, “Visual Passage Score Aggregation for Image Retrieval”, MIPR, 2023
- [14] M. Dao, M. A. Riegler, D. Dang-Nguyen, H. Tran, R. U. Kiran, T. Komamizu, “ICDAR 24: Intelligent Cross-Data Analysis and Retrieval”, ICMR, pp.1332-1333, 2024

2 応募者の研究遂行能力及び研究環境

応募者（研究代表者、研究分担者）の研究計画の実行可能性を示すため、(1)これまでの研究活動（主要な研究業績を含む）、(2)研究環境（研究遂行に必要な研究施設・設備・研究資料等を含む）について2頁以内で記述すること。

「(1)これまでの研究活動」の記述には、研究計画に関連した国際的な取組（国際共同研究の実施歴や海外機関での研究歴等）がある場合には必要に応じてその内容を含めること。また、研究活動を中断していた期間がある場合にはその説明などを含めてもよい。

(1) これまでの研究活動

《研究代表者：駒水 孝裕》研究代表者・駒水は、データ統合における基礎技術、データ管理に関する研究に取り組んでおり、本研究遂行のための中心的な知見を有し統括するための知見を有し、研究分担者とも共同研究をすでに実施しており、本研究の代表者として牽引できる。研究代表者はこれまでに、画像検索 [1]（研究分担者・井手と共同）、不均衡データ分類 [2]、グラフデータベースにおける索引 [3]（研究分担者・波多野と共同）、法令文書処理 [4]、LOD の効率的な分析 [5] などの幅広い分野にて、Best Paper Award や Best Paper Runner-up を受賞するなど客観的に評価される研究成果を上げている。本研究に関する研究としては、以下のような成果を上げている。

- **LOD のデータ管理**：研究代表者は、LOD におけるエンティティの検索 [6, 7] や汎用的な分析システム [5] の研究を通して、オープンデータを活用する際に技術的な専門知識を必要としないアクセス・分析手法を実現した。[6] では、キーワード検索を用いて LOD 中のエンティティを検索する手法を提案し、ベンチマークにて最良の検索性能を実現した。[5] では、LOD 中の数値データを多次元分析するための、データ前処理を効率化した。また、研究分担者・波多野と共同研究において、グラフデータを効率的に検索するための索引構築に関する研究を行った [3]。
- **多様なデータに対する検索**：研究代表者は、構造化データの紐付けられたエンティティに対して、従来では異なる問合せ言語を別々に実行していたものを、単一言語での問合せを可能にした [8]。画像検索において、研究分担者・井手と共同で、効果的な画像検索手法を事前処理を工夫することで、検索性能を落とすことなく効率的な検索を実現する方式を提案した [1]。

《研究分担者：井手 一郎》研究分担者・井手は、画像を中心とするマルチメディアコンテンツのデータマイニング技術に関する知見を多数有し、研究分担者・Kastner と共同研究を実施してきている。本研究に係る技術としては、画像キャプションやシーングラフ生成に関する研究がある [9, 10]。同分担者は、画像キャプションに対して、心理学で用いられる印象粒度を表す心像性を取り入れ、心像性に合わせたキャプション生成 [9] や画像集合からの情報抽出および包括的なシーングラフ生成を実現している [10]。加えて、本研究の Linked Multimedia Open Data Management 領域において、今後必要とされる映像に対する技術開発の経験も有する [11, 12]。これらのことから、同研究分担者は本研究における画像処理技術に関する貢献だけでなく、提案領域の発展性についての議論を深めることができることから、研究分担者として適切である。

《研究分担者：石川 佳治》研究分担者・石川は、データベース分野の研究者であり、選択率推定 [13] やストリームデータ処理 [14] などを研究している。本研究に関連する研究として、ベクトル空間上での近傍検索を効率化する方法論に関する知見も有する [15, 16]。同研究分担者は、効率的なデータ管理に関する幅広い知見の提供や共同研究開発を通して、本研究の目的達成に貢献できるため、本研究の分担者として適切である。

《研究分担者：波多野 賢治》研究分担者・波多野は、テキストからの情報抽出や構造化データの効率的な管理に関する知見を有し、研究代表者・駒水とも共同研究を実施してきている。構造化データ（XML データ）を対象とした効率的な検索 [17] やテキストからの固有表現抽出と抽出された固有表現からの知識獲得に関する研究 [18]、大規模言語モデルを基本としたテキスト分類における効果的な手法 [19] に関する幅広い知見を有している。加えて、グラフデータに対する効率的な検索についても研究しており [3]、データ横断的な検索に必要な知見も有している。これらのことから、同研究分担者はデータ管理・検索に関する知見の提供や共同研究開発を通して、本研究の目的達成に貢献できるため、本研究の分担者として適切である。

【2 応募者の研究遂行能力及び研究環境 (つづき)】

《研究分担者：Marc A. Kastner》 研究分担者・Kastner は画像処理，特に画像と人間の感覚の関係性分析に知見を有し，研究代表者・駒水や研究分担者・井手とも共同研究を実施してきている。現状の多くの画像処理や言語処理に関するモデルは人間の感覚を十分に取り入れられていない。例えば，言語が想起するイメージを正確に表現することは容易ではない。同研究分担者は，画像キャプション生成や未知語からの画像生成などにおいて，人間の感覚を取り入れた研究を行っている [20, 21]。RAG における正確な情報提供において，マルチメディアデータに対する人の感覚を取り入れることで不正確な情報提供を是正することが期待される。したがって，同研究分担者は，人の感覚に寄り添ったデータ活用に関する知見の提供や共同研究開発を通して，本研究の目的達成に貢献できるため，本研究の分担者として適切である。

(2) 研究環境

研究代表者および分担者の研究施設には，それぞれの居室および研究室の実験室を利用する。それぞれの施設において，研究設備としてすでに数台の計算機が設置されているが，本研究とは別目的であるため別途購入する。メモリや GPU などの計算資源を十分に確保するために，名古屋大学のスーパーコンピュータ「不老」を積極的に活用する。ソフトウェア開発および実験のためのノート PC やタブレット PC，実験データ保管用のストレージなどは適宜購入する。

研究代表者と研究分担者の連携は，オンラインコミュニケーションツールを基本とし，必要に応じて対面での議論を行う。研究成果は研究代表者が集約するため，研究分担者は成果を研究代表者に共有する。連携体制はすでに確立しており，共同で研究を進める準備はできている。

〔参考文献〕 ※下線は研究代表者または研究分担者

- [1] T. Kato, T. Komamizu, I. Ide, “R-DiP: Re-ranking Based Diffusion Pre-computation for Image Retrieval”, DEXA, pp.233-247, 2024 (**Best Paper Award**)
- [2] T. Komamizu, “Towards Ensemble-based Imbalanced Text Classification using Metric Learning”, DEXA, pp.188-202, 2023 (**Best Paper Award**)
- [3] K. Kusu, T. Komamizu, K. Hatano, “Towards Efficient Data Access Through Multiple Relationship in Graph-Structured Digital Archives”, ICADL, pp.377-391, 2022 (**Best Paper Runner-up**)
- [4] T. Yamakoshi, T. Komamizu, Y. Ogawa, K. Toyama, “Japanese Legal Term Correction Using Random Forests”, JURIX, pp.161-170, 2018 (**Best Paper Award**)
- [5] T. Komamizu, T. Amagasa, H. Kitagawa, “SPOOL: A SPARQL-based ETL Framework for OLAP over Linked Data”, iiWAS, pp.49:1-10, 2015 (**Best Paper Award**)
- [6] T. Komamizu, “Random walk-based entity representation learning and re-ranking for entity search”, Knowl. Inf. Syst., Vol.62, Iss.8, pp.2989-3013, 2020
- [7] T. Komamizu, T. Amagasa, H. Kitagawa, “CROISSANT: Centralized Relational Interface for Web-scale SPARQL Endpoints”, iiWAS, pp.284-288, 2017
- [8] T. Komamizu, “SPARQL with XQuery-based Filtering”, ISWC, pp.69-74, 2020
- [9] M. A. Kastner, I. Ide, et al., “Estimating the imageability of words by mining visual characteristics from crawled image data”, Multim. Tools Appl. 79(25-26), pp.18167-18199, 2020
- [10] I. Phueaksri, M. A. Kastner, Y. Kawanishi, T. Komamizu, I. Ide, “Image-Collection Summarization Using Scene-Graph Generation With External Knowledge”, IEEE Access, Vol.12, pp.17499-17512, 2024
- [11] F. Nack, I. Ide, “Why did the Prime Minister resign? -Generation of event explanations from large news repositories-”, ACM Multimedia, pp.313-322, 2011
- [12] K. Doman, T. Tomita, I. Ide, D. Deguchi, H. Murase, “Event Detection based on Twitter Enthusiasm Degree for Generating a Sports Highlight Video”, ACM Multimedia, pp.949-952, 2019
- [13] Y. Wang, C. Xiao, J. Qin, R. Mao, M. Onizuka, W. Wang, R. Zhang, Y. Ishikawa, “Consistent and Flexible Selectivity Estimation for High-Dimensional Data”, SIGMOD, pp.2319-2327, 2021
- [14] D. Takao, K. Sugiura, Y. Ishikawa, “Approximate Fault-Tolerant Data Stream Aggregation for Edge Computing”, BDA, pp.233-244, 2021
- [15] K. Lu, Y. Ishikawa, C. Xiao, “MQH: Locality Sensitive Hashing on Multi-level Quantization Errors for Point-to-Hyperplane Distances”, PVLDB, Vol. 16, No. 4, pp. 864-876, 2023
- [16] L. Zhao, Y. Wang, J. Kato, Y. Ishikawa, “Learning Local Similarity with Spatial Interrelations on Content-Based Image Retrieval”, IEICE Trans. Inf. Syst. Vol. 106, No. 5, pp.1069-1080, 2023
- [17] 波多野 賢治, 他, “XML 文書検索システムにおける文書内容の統計量を利用した検索対象部分文書の決定”, 電子情報通信学会論文誌, J89-D(3), pp.422-431, 2006 (**平成 18 年度電子情報通信学会論文賞受賞**)
- [18] K. Kusu, N. Makino, T. Shioi, K. Hatano, “Calculating Cooking Recipe’s Difficulty based on Cooking Activities”, CEA@IJCAI’17, pp.19-24, 2017
- [19] Y. Kimura, T. Komamizu, K. Hatano, “An Automatic Labeling Method for Subword-Phrase Recognition in Effective Text Classification”, Informatica 47(3), 2023
- [20] M. A. Kastner, K. Umemura, I. Ide, et al., “Imageability- and length-controllable image captioning”, IEEE Access, vol. 9, pp. 162951-162961, 2021.
- [21] C. Matsuhira, M. A. Kastner, T. Komamizu, T. Hirayama, K. Doman, I. Ide, “Nonword-to-Image Generation Considering Perceptual Association of Phonetically Similar Words”, McGE@ACMMM, 2023

3 人権の保護及び法令等の遵守への対応（公募要領参照）

本研究を遂行するに当たって、相手方の同意・協力を必要とする研究、個人情報の取扱いの配慮を必要とする研究、生命倫理・安全対策に対する取組を必要とする研究など指針・法令等（国際共同研究を行う国・地域の指針・法令等を含む）に基づく手続が必要な研究が含まれている場合、講じる対策と措置を、1頁以内で記述すること。

個人情報を伴うアンケート調査・インタビュー調査・行動調査（個人履歴・映像を含む）、提供を受けた試料の使用、ヒト遺伝子解析研究、遺伝子組換え実験、動物実験など、研究機関内外の倫理委員会等における承認手続が必要となる調査・研究・実験などが対象となる。

該当しない場合には、その旨記述すること。

アンケート調査やインタビュー調査は予定されておらず、個人情報を伴うデータを扱わないため、基本的には該当しないが、研究遂行において個人情報や著作権に関わるデータを扱う場合には法令やライセンスに従いデータの利用を適切に行う。

研究代表者および分担者はそれぞれの所属機関にて倫理講習をそれぞれ受講しており、被験者を対象とした実験が必要となった場合においても、適切な審査の上で実施する事ができる。

以上に記す通り、研究代表者および研究分担者は「研究の人権の保護および法令等の遵守」のために必要な知識を備えており、十分に対応できると考える。

4 研究計画最終年度前年度応募を行う場合の記述事項（該当者は必ず記述すること（公募要領参照））

本研究の研究代表者が行っている、令和7（2025）年度が最終年度に当たる継続研究課題の当初研究計画、その研究によって得られた新たな知見等の研究成果を記述するとともに、当該研究の進展を踏まえ、本研究を前年度応募する理由（研究の展開状況、経費の必要性等）を1頁以内で記述すること。

該当しない場合は記述欄を削除することなく、空欄のまま提出すること。

研究種目名	課題番号	研究課題名	研究期間

該当しない。

年度	設備備品費の明細					消耗品費の明細	
	品名・仕様	設置機関	数量	単価	金額	事項	金額
R7	作業用ノートPC	名古屋大学	1	250	250	PC 周辺機器	55
R7	計算用 GPU	名古屋大学	1	1,000	1,000	Buffalo MiniStation SSD	75
R7	開発用ワークステーション	名古屋大学	1	800	800		
R7				計	2,050	計	130
R8	計算用 GPU	名古屋大学	1	1,000	1,000	Buffalo HD-WHA16U3/R1	150
R8				計	1,000	計	150
R9	データ管理用PC	名古屋大学	1	250	250	Crucial BX500 4TB	120
R9				計	250	計	120

設備備品費、消耗品費の必要性
各研究者はすでに研究設備を持っているが、本研究の用途とは異なる上に、経年劣化による追加の設備が必要となるため、作業および開発のための PC・ワークステーションを購入する。また、本研究は深層学習を行うために、名古屋大学のスーパーコンピュータ「不老」を用いるが、効果的に用いるために手元での開発用の GPU が必要となるため購入する。

基盤研究 (B) (一般) 1 1 - (1)

(金額単位 : 千円)

年度	国内旅費の明細		外国旅費の明細		人件費・謝金の明細		その他の明細	
	事項	金額	事項	金額	事項	金額	事項	金額
R7	国内会議 (東京, 2名, 2泊)	100	国際会議 SIGMOD (米国, 1名, 5泊) 参加	450	研究アシスタント雇用 (2人×3ヶ月)	300	国際会議参加費	400
R7	国内会議 (沖縄, 2名, 2泊)	180	国際会議 ACM Multimedia (アイルランド, 1名, 5泊) 参加	450			論文誌出版費	120
R7			国際会議 ACL (オーストリア, 1名, 5泊) 参加	450			名大「不老」利用費	1,000
R7			国際会議 ICMR (タイ, 1名, 5泊) 参加	400				
R7	計	280	計	1,750	計	300	計	1,520
R8	国内会議 (東京, 2名, 2泊)	100	国際会議 SIGMOD (米国, 1名, 5泊) 参加	450	研究アシスタント雇用 (2人×6ヶ月)	600	国際会議参加費	400
R8	国内会議 (沖縄, 2名, 2泊)	180	国際会議 ACM Multimedia (アイルランド, 1名, 5泊) 参加	450			論文誌出版費	120
R8			国際会議 ACL (オーストリア, 1名, 5泊) 参加	450			名大「不老」利用費	1,000
R8			国際会議 ICMR (タイ, 1名, 5泊) 参加	400				
R8	計	280	計	1,750	計	600	計	1,520
R9	国内会議 (東京, 2名, 2泊)	100	国際会議 SIGMOD (米国, 1名, 5泊) 参加	450	研究アシスタント雇用 (2人×6ヶ月)	600	国際会議参加費	400
R9	国内会議 (沖縄, 2名, 2泊)	180	国際会議 ACM Multimedia (アイルランド, 1名, 5泊) 参加	450			論文誌出版費	120
R9			国際会議 ACL (オーストリア, 1名, 5泊) 参加	450			名大「不老」利用費	1,000
R9			国際会議 ICMR (タイ, 1名, 5泊) 参加	400				
R9	計	280	計	1,750	計	600	計	1,520
R10	国内会議 (東京, 2名, 2泊)	100	国際会議 SIGMOD (米国, 1名, 5泊) 参加	450	研究アシスタント雇用 (2人×6ヶ月)	600	国際会議参加費	400
R10	国内会議 (沖縄, 2名, 2泊)	180	国際会議 ACM Multimedia (アイルランド, 1名, 5泊) 参加	450			論文誌出版費	120
R10			国際会議 ACL (オーストリア, 1名, 5泊) 参加	450			名大「不老」利用費	1,000
R10			国際会議 ICMR (タイ, 1名, 5泊) 参加	400				
R10	計	280	計	1,750	計	600	計	1,520

旅費、人件費・謝金、その他の必要性

本研究の国際性を担保するために、国際会議 ICMR にて国際ワークショップ ICDAR を継続的に実施する。また、RAG (Retrieval-Augmented Generation) について特に活発に議論される国際会議 ACL やマルチメディア情報処理に関する国際会議 ACM Multimedia, データ管理に関する国際会議 SIGMOD に研究動向調査および研究発表するための外国旅費が毎年必要となる。加えて、国内のコミュニティを醸成し、本研究の提案領域の普及および協力関係を拡大・強化するために国内学会での発表および意見交換のために国内旅費が必要である。
また、本研究をスムーズに遂行するために、代表者・分担者の博士前期・後期課程の学生を適宜研究アシスタントとして雇用する。
上述のように、本研究における計算資源として名古屋大学のスーパーコンピュータ「不老」を用いるため、その使用料を計上している。