

令和 3 年 6 月 18 日現在

機関番号：12102

研究種目：基盤研究(C) (一般)

研究期間：2017～2020

課題番号：17K00449

研究課題名(和文)ブラウジング型探索タスクの解明とその支援手法に関する研究

研究課題名(英文)Research on Browsing-type Information Seeking Tasks and Their Support Method

研究代表者

高久 雅生 (Takaku, Masao)

筑波大学・図書館情報メディア系・准教授

研究者番号：00399271

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：ブラウジング型の探索タスクを支援するための手がかりとして、Linked Open Data (LOD) データセットの活用を主軸に、分野固有のLODデータセットとして、教科書LOD、無形民俗文化財などを対象としたデータセットのモデル化と構築に取り組み、データセット公開、利活用手法の評価などを行った。また、分野横断型で利用可能な基盤手法として、DBpedia等を用いたセレンディビティ推薦手法や、ウィキペディアコンテンツにおける学術文献参照同定のための手法を開発した。開発した手法は、正解データに基づく精度評価やユーザ実験に基づく指標をもとに性能検証を行い、その効果について検証した。

研究成果の学術的意義や社会的意義

本研究で開発した成果は多様な領域で応用可能な手法であるとともに、一般的な研究のなかで活用できる。とりわけ、LODデータセットそのものはウェブ上で公開データとして受け入れられており、価値が高い。

研究成果の概要(英文)：As a clue to support browsing-type exploration tasks, we focused on the utilization of Linked Open Data (LOD) datasets. For field-specific LOD datasets, we designed, built and published datasets for textbook LODs, intangible folk cultural properties, etc. And then we evaluated utilization methods for each dataset. In addition, as basic methodologies for cross-disciplinary fields, we developed a serendipity recommendation method using DBpedia, and a method for identifying references to scholarly literature in Wikipedia content. Performances of these methods were verified by precisions based on a correct sample dataset and metrics based on user experiments.

研究分野：図書館情報学

キーワード：電子図書館 Linked Open Data 利活用 情報検索

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

1. 研究開始当初の背景

情報探索行動は図書館情報学における主要な研究領域の一つであり、探索者の情報ニーズとその探索過程をモデル化することは多くの研究者の焦点となってきた。また、近年のウェブ上では、探索空間に存在する文書やコンテンツの数が膨大になるため、このような環境下における探索者のタスクモデルや調査学習型探索 (*exploratory search*) といった枠組みも多く提案されている。このような状況において、探索者の持つ情報要求は複雑なプロセスを組み合わせたものとなり、探索時に用いるリソースが複合的になったり、複数の観点から探索を繰り返したりすることにより、課題解決に至るための目標を構成することとなる。

ブラウジング型探索は、このような試行錯誤を含む過程を支援するための戦略の一つとしてみることができ、単純なクエリの選択、検索戦術の選択に加えて、リソース内容に応じたコレクション全体の概観を閲覧することにより、探索対象の内容を適切に把握し、新しい観点で探索する手がかりを得ることが重要なポイントとなる。ブラウジング型探索のような試行錯誤を含む探索活動のプロセス全体を支援することは、探索者の助けとなる利益となるだけでなく、複雑な探索タスクの理解をも広げることが期待できる。

研究代表者はこれまで、調査学習型探索における探索タスクの特性や利用者属性が探索全体のプロセスに与える影響をみる研究や、コンセプトマップを用いた探索前後の探索者の知識変容の過程などに着目した研究を行ってきた。ただし、研究代表者自身を含む従来の研究は、探索ニーズに応じた探索プロセスを観察して分析する研究であり、直接に情報アクセスの利活用に向けた研究アプローチではなかった。とりわけ、ブラウジング型探索を支援するための情報アクセス手法として、どのような機能が必要か、どのような手がかりを用いるか、どのように評価するかを包括的に研究した例は見当たらない。

2. 研究の目的

本研究課題では、いくつかの具体的な資料コレクションを対象として、ブラウジング型探索を実現する情報アクセス手法を提案し、その有効性を評価する。本研究では、ブラウジング型探索を支援するために、任意の資料コレクションに対するブラウジング型探索を実現するためのアクセスの観点を設定する。この観点的設定には、リソース固有のものと分野横断可能なものの2種類があることを仮定した上で、これらの観点を通じたブラウジング型探索行動を分析し、その利活用に資する支援手法を明らかにする。

3. 研究の方法

研究にあたっては、資料コレクション固有のブラウジングを支援する観点には、新しい手がかりを用いる必要がある。本研究では、この手がかりに Linked Open Data (LOD) の考え方に基づいて公開されているエンティティを用いる。LOD はリンクする複数のエンティティ同士を結びつけるデータ構造をもち、ブラウジング型探索の要件である複数の資料群をたどりながら閲覧するための仕掛けと親和性が高い。適切なブラウジング型探索の支援となるよう、明示的な関係性をもたない LOD エンティティと資料コレクションとの関係性を推定して、その関係性を手がかりとした観点を提示することにより、ブラウジング型探索の支援とする。このためのコレクション内の情報抽出と LOD エンティティの紐づけを行う手法を開発する。

具体的には、グラフモデルにより構造化された LOD エンティティと、構造化されていないコレクション内の個別リソースとの間をリンク付けし、エンティティとして扱う手法とその適用範囲、適用のための粒度を検討する。さらに、開発した手法をいくつかの資料コレクションに対して適用し、その手法の有効性を明らかにする。

4. 研究成果

本研究課題の研究期間を通じた成果のうち主なものを以下に挙げる。

国立教育政策研究所教育図書館が所蔵する近代教科書の書誌情報約 7,500 タイトルに基づく LOD データセットを対象として、データセットスキーマの抽出と形式記述の方式を検討し、文書化と検証を自動的に行う方式を提案した。あわせて、任意の RDF/Turtle 形式のデータセットに対して、これをウェブ上で Linked Data 原則に基づくデータセットとして提供するためのツール `tt12html` を開発し、公開した。

上記成果をもとに、横展開を図るために、他の分野領域のデータセットとして、無形民俗文化財データセット IFCP を公開し、その有効性評価として機能要件に基づく利活用プロトタイプ評価を行った。

また、別の視点から書誌情報をとらえるための枠組みとして「セレンディピティ」に基づく推

薦手法について検討を進め、書籍の著者情報のリンク関係をもとにしたセレンディピティ推薦手法を提案した。提案手法は、LOD データセットと GoodReads の書誌情報約 25,000 冊のデータセットを対象としたユーザ実験を通じて評価し、新規性の尺度において比較手法に比べて有意な改善が得られた。

これらの成果に加えて、ウィキペディア英語版を対象とした百科事典記事内における学術文献参照関係のナレッジグラフを抽出する試みとして、学術文献参照データセットから初出時点を特定する手法を開発し、約 18 万記事に掲載された延べ 92 万件以上の DOI リンクを対象として評価し、全体で 93.3%と高い精度で特定できることを示した。

今後の展望として、ブラウジングタスクの特性を活かすための手がかりについての検討は進んだものの、実際のタスク支援の手法については今後の課題として残っている。これらの課題の解決に向けて、継続する研究課題の中で引き続き取り組んでいきたい。

5. 主な発表論文等

〔雑誌論文〕 計10件（うち査読付論文 3件 / うち国際共著 0件 / うちオープンアクセス 9件）

| | |
|---|-------------------------|
| 1. 著者名 高久 雅生 | 4. 巻 69 |
| 2. 論文標題 進化する情報検索技術 | 5. 発行年 2019年 |
| 3. 雑誌名 情報の科学と技術 | 6. 最初と最後の頁 182 ~ 188 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.18919/jkg.69.5_182 | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |
| 1. 著者名 高久雅生 | 4. 巻 2 |
| 2. 論文標題 「Europeanaの翻刻プロジェクトと日本の翻刻プロジェクト」に参加して | 5. 発行年 2018年 |
| 3. 雑誌名 デジタルアーカイブ学会誌 | 6. 最初と最後の頁 298-299 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.24506/jsda.2.3_298 | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |
| 1. 著者名 江草由佳, 高久雅生 | 4. 巻 68 |
| 2. 論文標題 教科書Linked Open Data (LOD) の構築と公開 | 5. 発行年 2018年 |
| 3. 雑誌名 情報の科学と技術 | 6. 最初と最後の頁 361-367 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.18919/jkg.68.7_361 | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |
| 1. 著者名 高久雅生 | 4. 巻 25 |
| 2. 論文標題 ウィキペディアタウンと図書館の連携が可能にする未来の知的環境 | 5. 発行年 2019年 |
| 3. 雑誌名 LRG : ライブラリー・リソース・ガイド | 6. 最初と最後の頁 36-50 |
| 掲載論文のDOI (デジタルオブジェクト識別子) なし | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |

| | |
|---|-----------------------|
| 1. 著者名 高久雅生 | 4. 巻 28 |
| 2. 論文標題 タスク重要: ユーザタスク指向のプラットフォーム設計と開発を目指して | 5. 発行年 2019年 |
| 3. 雑誌名 情報知識学会誌 | 6. 最初と最後の頁 363-366 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.2964/jsik_2019_015 | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |

| | |
|---|-----------------------|
| 1. 著者名 吉川次郎, 高久雅生, 芳鐘冬樹 | 4. 巻 30 |
| 2. 論文標題 DOIリンクに基づくWikipedia 上の参照記述における編集者の分析 | 5. 発行年 2020年 |
| 3. 雑誌名 情報知識学会誌 | 6. 最初と最後の頁 21 ~ 41 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.2964/jsik_2020_004 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |

| | |
|---|-------------------------|
| 1. 著者名 高久雅生 | 4. 巻 70 |
| 2. 論文標題 図書館業務に関連するプログラム開発事例とその諸要素 | 5. 発行年 2020年 |
| 3. 雑誌名 情報の科学と技術 | 6. 最初と最後の頁 192 ~ 196 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.18919/jkg.70.4_192 | 査読の有無 無 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |

| | |
|---|-------------------------|
| 1. 著者名 吉川次郎, 高久雅生, 芳鐘冬樹 | 4. 巻 30 |
| 2. 論文標題 Wikipediaに学術文献の参照記述を追加する編集の特定手法 | 5. 発行年 2020年 |
| 3. 雑誌名 情報知識学会誌 | 6. 最初と最後の頁 370 ~ 389 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.2964/jsik_2020_033 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている (また、その予定である) | 国際共著 - |

| | |
|--|-----------------------|
| 1. 著者名 江草由佳, 岡本真, 清田陽司, 高久雅生, 常川真央 | 4. 巻 35 |
| 2. 論文標題 図書館におけるテクノロジー活用を促進する「場」の創出: Code4Lib JAPAN コミュニティの取組み | 5. 発行年 2020年 |
| 3. 雑誌名 人工知能 | 6. 最初と最後の頁 766-774 |
| 掲載論文のDOI (デジタルオブジェクト識別子) なし | 査読の有無 無 |
| オープンアクセス オープンアクセスではない、又はオープンアクセスが困難 | 国際共著 - |

| | |
|---|--------------------|
| 1. 著者名 吉川次郎, 高久雅生, 芳鐘冬樹 | 4. 巻 31 |
| 2. 論文標題 Wikipedia上の学術文献の参照記述の追加に関する時系列分析 | 5. 発行年 2021年 |
| 3. 雑誌名 情報知識学会誌 | 6. 最初と最後の頁 3~19 |
| 掲載論文のDOI (デジタルオブジェクト識別子) 10.2964/jsik_2020_037 | 査読の有無 有 |
| オープンアクセス オープンアクセスとしている(また、その予定である) | 国際共著 - |

[学会発表] 計13件(うち招待講演 0件/うち国際学会 3件)

| |
|---|
| 1. 発表者名 Renlou Weng, Masao Takaku |
| 2. 発表標題 Author-Oriented Book Recommendation Using Linked Open Data for Improving Serendipity |
| 3. 学会等名 情報処理学会第137回情報基礎とアクセス技術研究発表会 |
| 4. 発表年 2020年 |

| |
|---|
| 1. 発表者名 高久雅生 |
| 2. 発表標題 軽量Linked Open Data公開ワークフロー: Poorman's ToolkitからXlsx2Shapesまで |
| 3. 学会等名 Code4Lib JAPANカンファレンス2020 |
| 4. 発表年 2020年 |

| |
|---|
| 1. 発表者名 Itsumi Sato, Masao Takaku |
| 2. 発表標題 Aggregation and Utilization of Metadata for Intangible Folk Cultural Properties Using Linked Open Data |
| 3. 学会等名 iConference 2021 (国際学会) |
| 4. 発表年 2021年 |

| |
|--|
| 1. 発表者名 Renlou Weng, Masao Takaku |
| 2. 発表標題 Book Recommender System Using Linked Data for Improving Serendipity |
| 3. 学会等名 第27回情報知識学会年次大会 |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 Yuka Egusa, Masao Takaku |
| 2. 発表標題 A Case Study of Japanese Textbook Linked Open Data: Publishing a Small Bibliographic Collection from a Special Library |
| 3. 学会等名 DCMI 2019 (国際学会) |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 叢艶, 高久雅生 |
| 2. 発表標題 唐詩の構造化に関する研究: テキストの差異提示機能の検討 |
| 3. 学会等名 じんもんこん2019 |
| 4. 発表年 2019年 |

| |
|---|
| 1. 発表者名 Renlou Weng, Masao Takaku |
| 2. 発表標題 Author-Oriented Book Recommendation Using Linked Open Data for Improving Serendipity |
| 3. 学会等名 第137回情報基礎とアクセス技術研究発表会 |
| 4. 発表年 2020年 |

| |
|---------------------------------------|
| 1. 発表者名 高久雅生, 小幡将司, 江草由佳 |
| 2. 発表標題 OPAC利用ログに基づく文献検索システムの試作と評価 |
| 3. 学会等名 第26回情報知識学会年次大会 |
| 4. 発表年 2018年 |

| |
|---|
| 1. 発表者名 吉川次郎, 高久雅生 |
| 2. 発表標題 英語版WikipediaにおけるDOIリンクの初出時点の分析: 研究分野を中心に |
| 3. 学会等名 情報メディア学会第17回研究大会 |
| 4. 発表年 2018年 |

| |
|--|
| 1. 発表者名 高久雅生 |
| 2. 発表標題 Linked Open Data on Github Pages: 教科書LODの経験から |
| 3. 学会等名 Code4Lib JAPANカンファレンス2018 |
| 4. 発表年 2018年 |

| |
|-------------------------------|
| 1. 発表者名 清水花菜子, 高久雅生 |
| 2. 発表標題 新書本を用いた学問発見支援手法の提案 |
| 3. 学会等名 第23回情報知識学フォーラム |
| 4. 発表年 2018年 |

| |
|--|
| 1. 発表者名 Yan Cong, Masao Takaku |
| 2. 発表標題 Prototype of Linked Open Data Model for Tang Poems |
| 3. 学会等名 Japanese Association for Digital Humanities Conference 2017 (JADH2017) (国際学会) |
| 4. 発表年 2017年 |

| |
|--|
| 1. 発表者名 佐藤いつみ, 高久雅生 |
| 2. 発表標題 Linked Open Dataを用いた年中行事・祭礼情報の構造化とその利活用 |
| 3. 学会等名 人文科学とコンピュータシンポジウム2020 |
| 4. 発表年 2020年 |

〔図書〕 計1件

| | |
|---------------------------|-----------------|
| 1. 著者名 根本 彰、齋藤 泰則、高久雅生 | 4. 発行年 2020年 |
| 2. 出版社 日本図書館協会 | 5. 総ページ数 349 |
| 3. 書名 レファレンスサービスの射程と展開 | |

〔産業財産権〕

〔その他〕

-

6. 研究組織

| | 氏名 (ローマ字氏名) (研究者番号) | 所属研究機関・部局・職 (機関番号) | 備考 |
|-------|---|---|----|
| 連携研究者 | 江草 由佳 (Egusa Yuka) (60413902) | 国立教育政策研究所・研究企画開発部教育研究情報推進室・ 総括研究官 (62601) | |

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
| | |