

令和 6 年 6 月 17 日現在

機関番号：12102

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K12543

研究課題名（和文）アプリケーション・プロファイルの記述手法とその共有プラットフォームの構築

研究課題名（英文）Development of Application Profile Description Method and Its Sharing Platform

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） 3,300,000円

研究成果の概要（和文）：本研究は、メタデータの相互運用性を高めるため、その記述項目の検証や文書化に用いられるアプリケーションプロファイル（Application Profile）を簡易に記述でき、共有し、再利用可能とする新たな手法を提案するとともに、その提案手法に基づくプラットフォーム構築のためのツール群を開発し、公開した。具体的な手法としては、RDFシェイプ表現に基づいてLinked Dataのグラフ表現を直接検証可能なメタデータ語彙を構造記述し、容易に文書化し、メタデータ項目を追記変更しやすい工夫を加えた方法を提案し、複数の実用的なLinked Dataデータセット構築において実現可能であることを示した。

研究成果の学術的意義や社会的意義

従来、デジタルコレクション群の整理・組織化と提供に用いるメタデータ語彙利用の枠組みは、理論的な面からも実践的な面にも、簡便に実践可能で効果的な方法論はほとんど示されてこなかったのに対し、本研究での提案手法は、アプリケーションプロファイルの記述法の効果的な標準化と簡易化を実現することにより、より実践的なアプリケーションプロファイルの構築と提供ができるようになった。特に本研究の成果により、アプリケーションプロファイルの記述法と提供方法を同時に実現するやり方とツールを提供することにより、これまでよりも簡易に実践可能な方法を示した。

研究成果の概要（英文）：To improve the interoperability of metadata, this research proposes a novel method that enables the simple description, sharing, and reuse of application profiles used to validate and document descriptive metadata items. A platform based on this method was developed, along with a set of tools that were made publicly available. Specifically, we proposed a method for a formal structural description of a metadata vocabulary. This method verifies the graph representation of Linked Data datasets based on the RDF shape representation schema and facilitates easy documentation and tracking of modifying metadata items. We have demonstrated the feasibility of this method by applying it to a number of practical Linked Data datasets.

研究分野：図書館情報学

キーワード：情報組織化 メタデータ アプリケーションプロファイル Linked Data

## 1. 研究開始当初の背景

ウェブ上で流通するコンテンツや資料群に対する整理組織化手法として、データベース化、デジタルアーカイブの構築が進んできた。国は『知的財産基本計画 2019』の中で、デジタルアーカイブの構築が文化の保存と継承、多様なコンテンツの融合の場となることを期待するとともに、このようなデジタルアーカイブの構築と利活用、連携の推進のための取り組みをうたっている。また、複数の機関から提供されているデジタルアーカイブのコンテンツを検索するためのポータルサイト「ジャパンサーチ」では絵画や古文書等の文化資源や科学技術情報、図書情報、放送コンテンツなどさまざまな分野のデータベースと連携し、2024年5月現在、約3千万点以上のコンテンツが検索可能となっている。デジタルアーカイブ構築は、今後とも多くの領域で進み、相互に連携しながら利活用されることが期待されている

このような多様な情報資源を整理組織化する際には、それぞれのデジタルアーカイブにおけるそれぞれの用途や領域の知識構造にあわせた異なるメタデータ語彙が使われるため、これを利用したり、他のデジタルアーカイブ等と連携して活用したりするには、メタデータ項目とその値、データ構造、その他の外部情報源へのリンク状態などの制約をまとめた情報として提供してもらい、その情報を共有しながら利用していかなければならない。このような情報資源の組織化を文書化し、連携共有するための枠組みとして、Dublin Core Metadata Initiativeは、2008年に「シンガポール・フレームワーク」と呼ぶ勧告を発表し、各コレクション単位でのメタデータの構造などを含む総合的な情報を「アプリケーションプロファイル (Application Profile)」としてまとめる手法が提案された。その後、アプリケーションプロファイルは、大規模なプロジェクトでメタデータモデルを共有する手法として一定程度確立しつつある。例えば、ヨーロッパ全域のデジタルアーカイブの横断検索ポータル Europeana は Europeana Data Model (EDM) と呼ばれるメタデータモデルを採用しており、提携参加機関に対してその情報をアプリケーションプロファイルとして文書化し、公開している。同様に、米国の DPLA (Digital Public Library of America) や日本の国立国会図書館サーチなどでもアプリケーションプロファイルの公開が行われてきた。

こうしたアプリケーションプロファイルの構築やアプリケーションプロファイルの一部を改変して別のコレクション群に流用・再利用したりするような方法論、機械可読で再利用しやすい形で記述する方法論の提案・実践はともに少なく、個別事例においてウェブサイトでの HTML や PDF 形式でのアプリケーションプロファイルが公開されるにとどまっている。本研究では、このような状況を抜本的に解決するために、アプリケーションプロファイルの記述と共有、活用をまとめて解決できる手法の開発を目指す

## 2. 研究の目的

本研究では、アプリケーションプロファイルをさらに活用してデジタルアーカイブの公開を促進することを目標として、アプリケーションプロファイルの設計、記述、提供、利用にあたって、1) 新しいコレクション群のデジタル化を行う際に適切な品質のアプリケーションプロファイルを設計するにはどのようにすればよいか、2) このようなアプリケーションプロファイルを簡便に記述するにはどのようにすればよいか、3) デジタルコンテンツの整理組織化の状況にあわせてメタデータとアプリケーションプロファイルの変更を随時に追跡するためにはどのような手段があるか、という3つの研究課題を立て、これらの問いを解決する提案手法とツール群の研究開発を進める。

### 3. 研究の方法

アプリケーションプロファイルにおいて記述される要素には以下のものがあり、これらを精細に記述し、共有する必要がある。0) データセットが対象とする事物のもつ性質全般とその利用要件、1) データセット内で規定されるリソース群の種類やクラス、2) データセット内で用いられる URI 識別子とリソース種別の関係、3) 使われるメタデータ語彙の由来と性質、4) データセット内に用いられるデータ項目値の書式や出現回数等の値制約と構造制約。これらの記述すべき要素のうち、前項 1) -4) の要素を機械可読で再利用可能な形で記述することにより、他のデータセットの開発者にとっても有用な情報となる。また、機械可読のアプリケーションプロファイルを用いて、入力されるメタデータインスタンスの値や構造が正しい状態となっているかを検証したり、場合によっては入力するための画面設計やユーザインタフェースそのものを生成したりすることにも活用できる。近年ではこのようなデータセットの形式記述として Shape Expression に基づく記述フォーマットが普及しつつある。このような記述フォーマットを、アプリケーションプロファイル共有のための枠組みと組み合わせ、アプリケーションプロファイル共有を実現する要件を検討し、整理した形で方法論として提案する。

提案した方法論において、アプリケーションプロファイルの記述内容や、記述する場合のコストと便益などは、具体的なデータセットごとに、求められるアプリケーションプロファイルの様態は異なるものと思われる。したがって、複数の具体的なデータセットを取り上げて、これらのアプリケーションプロファイルの記述を試行しながら、アプリケーションプロファイルの記述手法の適用可能性を検証した。

### 4. 研究成果

本研究課題の研究期間を通じた研究成果を以下に挙げる。

アプリケーションプロファイルの記述手法として、シェイプ記述標準 Shapes Constraint Language (SHACL) を基にした記述法を提案した。さらに、提案手法の簡便性を高めるために、エクセル形式から SHACL 形式を簡易に生成できる方法も提案した(図 1)。両者の方法を有機的に組み合わせたツールのワークフローを提案するとともに、実践的に利用可能なツール `ttl2html`、`xlsx2shape` ツールとして実装し、公開提供した。

研究代表者を中心とする研究チームにおいて開発保守を進めてきた Linked Open Data データセット、教科書 LOD、単元 LOD、学習指導要領 LOD の 3 つのデータセットを対象として、提案手法の記述法、提供公開法などを試用し、ユーザビリティ、使い勝手、対応可能な記述要素の範囲、さらなる問題点の洗い出しなどを行い、必要に応じて、提案手法や記述法の拡張を行い、研究を進展させた。各 LOD データセットにおいて必要となったメタデータ項目、要素種別や拡張は以下の通りである：

	A	B	C	D	E	F	G	H	I	J
1	<a href="https://w3id.org/jp-textbook/">https://w3id.org/jp-textbook/</a>	sh:path	sh:name@ja	sh:name@en	sh:description@ja	sh:description@en	skos:exam	sh:maxCo	sh:minCo	sh:class
2	sh:targetClass	<a href="https://w3id.org/jp-textbook/Textbook">https://w3id.org/jp-textbook/Textbook</a>					NEW CRO	1	1	
3	sh:property	<a href="http://schema.org/name">http://schema.org/name</a>	書名	Title						
4	sh:property	<a href="http://schema.org/editor">http://schema.org/editor</a>	編者名	Editor(s)			両佐田 ほか	1	1	
5	sh:property	<a href="http://schema.org/publisher">http://schema.org/publisher</a>	出版者名	Publisher	<code>textbook:ps:Link to the publisher</code>					
6	sh:property	<a href="http://schema.org/bookEdition">http://schema.org/bookEdition</a>	版情報	Edition						
7	sh:property	<a href="http://schema.org/isbn">http://schema.org/isbn</a>	ISBN	ISBN						
8	sh:property	<a href="http://id.loc.gov/ontologies/bibframe/extent">http://id.loc.gov/ontologies/bibframe/extent</a>	ページ数	The total pages			120	1		
9	sh:property	<a href="http://id.loc.gov/ontologies/bibframe/dimensions">http://id.loc.gov/ontologies/bibframe/dimensions</a>	大きさ	Dimensions			94	1		
10	sh:property	<a href="https://w3id.org/jp-textbook/grade">https://w3id.org/jp-textbook/grade</a>	対象学年	Target grade			1, 2, 3	6		
11	sh:property	<a href="https://w3id.org/jp-textbook/authorizedYear">https://w3id.org/jp-textbook/authorizedYear</a>	検定年	Authorized year			2001			
12	sh:property	<a href="https://w3id.org/jp-textbook/catalogue">https://w3id.org/jp-textbook/catalogue</a>	掲載教科書目録	Textbook cat	<code>textbook:cc:Link to the catalogue</code>					
13	sh:property	<a href="https://w3id.org/jp-textbook/usageYear">https://w3id.org/jp-textbook/usageYear</a>	使用年	Usage year			2005	1		
14	sh:property	<a href="https://w3id.org/jp-textbook/usageYearRange">https://w3id.org/jp-textbook/usageYearRange</a>	使用年の区間	Range of Usa "yyyy-yyyy"	もし L < Either of "yyyy-yyyy"2005-2006			1	1	
15	sh:property	<a href="https://w3id.org/jp-textbook/textbookSymbol">https://w3id.org/jp-textbook/textbookSymbol</a>	教科書記号	Textbook symbol			英語	1	1	
16	sh:property	<a href="https://w3id.org/jp-textbook/textbookNumber">https://w3id.org/jp-textbook/textbookNumber</a>	教科書番号	Textbook number			904	1	1	
17	sh:property	<a href="https://w3id.org/jp-textbook/subject">https://w3id.org/jp-textbook/subject</a>	科目 (科目)	Subject	<code>textbook:cc:Link to the subject</code>					
18	sh:property	<a href="https://w3id.org/jp-textbook/subjectArea">https://w3id.org/jp-textbook/subjectArea</a>	教科	Subject area	<code>textbook:cc:Link to the subject</code>					
19	sh:property	<a href="https://w3id.org/jp-textbook/curriculum">https://w3id.org/jp-textbook/curriculum</a>	対応する学習	Curriculum G	<code>textbook:cc:Link to the curriculum</code>					
20	sh:property	<a href="https://w3id.org/jp-textbook/school">https://w3id.org/jp-textbook/school</a>	学校種別	Type of school	対応する学校種別を Link to the type rest textbooks					
21	sh:property	<a href="https://w3id.org/jp-textbook/item">https://w3id.org/jp-textbook/item</a>	所属情報	Holding item	<code>nier:record Link to the blank node resource(s)</code>, which					
22	sh:property	<a href="http://www.w3.org/2000/01/rdf-schema#seeAlso">http://www.w3.org/2000/01/rdf-schema#seeAlso</a>	関連リンク	Related link	(繰り返し繰り返しあり。 <a Repeatable. Links (http://id.ndl.go.jp/jpno/99034331					
23	sh:property	<a href="http://id.loc.gov/ontologies/bibframe/note">http://id.loc.gov/ontologies/bibframe/note</a>	注記	Notes			2009年度より出版社を「日本文芸出版株式会社			
24	sh:property	<a href="http://purl.org/dc/terms/bibliographicCitation">http://purl.org/dc/terms/bibliographicCitation</a>	書誌情報	Bibliographic information			「あたりし	1	1	

図 1 エクセル形式を用いたシェイプ記述形式の例

- 教科書 LOD(図2): 14 種類のメタデータクラス、38 回のリリースバージョン履歴情報、空白ノード対応、複数選択から成る切替構造項目 (sh:or) SPARQL エンドポイント
- 単元 LOD: 5 種類のメタデータクラス、5 回のリリースバージョン履歴情報、空白ノード対応、SPARQL エンドポイント
- 学習指導要領 LOD: 14 種類のメタデータクラス、6 回のリリースバージョン履歴情報、空白ノード対応、SPARQL エンドポイント

各リソースの説明  
教科書

対象クラス: <https://w3id.org/jp-textbook/Textbook>

教科書を選択するクラス。以下のようなURIにより参照されます。  
<https://w3id.org/jp-textbook/> (学級単位の) (教科書) (条件書番号)

※教科書を各リソースのURIにアクセスすると、HTML形式で教科書についての基本的な情報を返します。URIの後ろに出をついてアクセスするとturtle形式のRDFデータを返します。

教科書リリースが持つプロパティを下表に示します

プロパティ名	説明	プロパティ値の例	必須・重複の有無	備考
<code>schema:name</code>	書名	NEW CROWN ENGLISH SERIES 3	必須 繰り返しの有無	
<code>schema:editor</code>	編者名	森山 隆 (もりやま たか)	必須 繰り返しの有無	
<code>schema:publisher</code>	出版社名	textbook:publisher/2003/三栄堂	必須 繰り返しの有無	<code>textbook:publisher</code> (対応関係) (条件書番号) (条件書番号) (条件書番号) として示される発行元をあらわすリソースと関連付ける
<code>schema:bookid:form</code>	版情報	改訂版	省略可能 繰り返しの有無	バージョン2.0.1までのデータ内には版情報が入っていないレコードが多い (古い教科書の備忘のため念のため用意している)。

図2 教科書 LOD におけるメタデータスキーマ記述の例 (抜粋)

本研究課題からの今後の展望として、以下の側面が考えられる。

- (1) 研究成果をさらに活かしてもらうためには、国内外での広報、啓発的な活動が欠かせないと思われる。他のデータセットへの適用などをさらに進めて、手法としての完成度を上げるとともに、ツール群のマニュアル化や英語化をさらにすすめて、「使えるツール」としての価値を高めることが必要となってくる。
- (2) 提案した手法によりアプリケーションプロファイル記述が簡便になったことにより、アプリケーションプロファイルの再利用可能性を高め、共有する手法の重要性がさらに高まると思われる。こうした共有のための方法論は、国内外を問わず、いまだ十分に深められていない研究領域であり、さらなる検討と研究開発が求められるものと思われる。
- (3) 上記(2)で述べた共有手法を深めるためにも、提案手法の評価の際に使用した LOD データセット3点は相互に関連付けられリンクしたデータセットであり、これらのリンク関係を適切に管理するような手法を検討することが、単なるメタデータスキーマの共有に留まらない、データセットの発見や推薦にもつながる可能性を持っていると考えられる。

## 5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 6件 / うち国際共著 0件 / うちオープンアクセス 5件）

1. 著者名 叢艶, 高久雅生	4. 巻 32
2. 論文標題 漢詩の構造化に関する研究: 唐詩作品を中心に	5. 発行年 2022年
3. 雑誌名 情報知識学会誌	6. 最初と最後の頁 15 ~ 38
掲載論文のDOI (デジタルオブジェクト識別子) 10.2964/jsik_2022_002	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 Kikkawa Jiro, Takaku Masao, Yoshikane Fuyuki	4. 巻 9
2. 論文標題 Dataset of first appearances of the scholarly bibliographic references on Wikipedia articles	5. 発行年 2022年
3. 雑誌名 Scientific Data	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1038/s41597-022-01190-z	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 高久雅生	4. 巻 2020
2. 論文標題 討議報告 『分類を構成する要素の解釈の明確化 Linked Data化を想定して 』	5. 発行年 2020年
3. 雑誌名 TP&Dフォーラムシリーズ 整理技術・情報管理等研究論集	6. 最初と最後の頁 27-29
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 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 コミュニティの取り組み	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 佐藤いつみ, 高久雅生	4. 巻 2020
2. 論文標題 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	5. 発行年 2020年
3. 雑誌名 人文科学とコンピュータシンポジウム2020	6. 最初と最後の頁 107-114
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 有
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 -

1. 著者名 Sato Itsumi, Takaku Masao	4. 巻 12645
2. 論文標題 Aggregation and Utilization of Metadata for Intangible Folk Cultural Properties Using Linked Open Data	5. 発行年 2021年
3. 雑誌名 Lecture Notes in Computer Science (LNCS)	6. 最初と最後の頁 154 ~ 164
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-71292-1_14	査読の有無 有
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 -

〔学会発表〕 計11件 (うち招待講演 1件 / うち国際学会 1件)

1. 発表者名 榎本 聡、大井 将生、高久 雅生、阿兒 雄之、有山 裕美子、江草 由佳
2. 発表標題 学習指導要領のLinked Open Data化による学習への利活用に向けた検討
3. 学会等名 日本教育工学会研究会
4. 発表年 2022年

1. 発表者名 小林恭輔, 高久雅生
2. 発表標題 楽曲探索を支援するための類似楽曲提示手法
3. 学会等名 第30回情報知識学会年次大会
4. 発表年 2022年

1. 発表者名 藤岡雄大, 高久雅生
2. 発表標題 日本のテレビ放送における番組のシリーズを同定する手法の開発
3. 学会等名 第30回情報知識学会年次大会
4. 発表年 2022年

1. 発表者名 阿児雄之, 有山裕美子, 江草由佳, 榎本聡, 大井将生, 高久雅生
2. 発表標題 学習指導要領LODの公開と活用の検討
3. 学会等名 2022年度アート・ドキュメンテーション学会年次大会
4. 発表年 2022年

1. 発表者名 島村翔, 高久雅生
2. 発表標題 ビデオゲームMOD配布プラットフォームにおける説明文のコンテンツ分類
3. 学会等名 情報処理学会 第147回 情報基礎とアクセス技術研究会 (SIG-IFAT) 研究会
4. 発表年 2022年

1. 発表者名 榎本聡, 大井将生, 高久雅生, 阿児雄之, 有山裕美子, 江草由佳
2. 発表標題 学習指導要領コードで紐づくデジタルコンテンツ
3. 学会等名 第48回全日本教育工学研究協議会全国大会
4. 発表年 2022年

1. 発表者名 叢艶, 高久雅生
2. 発表標題 漢詩作品の自動生成とその評価システムの試み
3. 学会等名 第27回情報知識学フォーラム
4. 発表年 2022年



1. 発表者名 Jiro Kikkawa, Masao Takaku, Fuyuki Yoshikane
2. 発表標題 Time Lag Analysis of Adding Scholarly References to English Wikipedia: How Rapidly Are They Added to and How Fresh Are They?
3. 学会等名 iConference 2023 (国際学会)
4. 発表年 2023年

1. 発表者名 高久雅生, 大井将生, 榎本聡, 江草由佳, 有山裕美子, 阿児雄之
2. 発表標題 学習指導要領LOD
3. 学会等名 LODチャレンジ2021オンライン授賞式シンポジウム (招待講演)
4. 発表年 2022年

1. 発表者名 高久雅生, 江草由佳
2. 発表標題 SHACL (Shapes Constraint Language) によるアプリケーションプロファイル記述の試み
3. 学会等名 第55回セマンティックウェブとオントロジー研究会
4. 発表年 2021年

1. 発表者名 阿児雄之, 有山裕美子, 江草由佳, 榎本聡, 大井将生, 高久雅生
2. 発表標題 教科書編修趣意書を介した学習内容とデジタルコンテンツ接続可能性の検討
3. 学会等名 アート・ドキュメンテーション学会2021年度年次大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

開発したツールはGitHubにて公開: <https://github.com/masao/ttl2html/>

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
研究分担者	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)	

	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
研究協力者	叢 艶  (Cong Yan)	筑波大学・大学院図書館情報メディア研究科・博士後期課程学生    (12102)	

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
---------	---------