

研究支援部署間の相互理解を目的とした 研究事業の実施記録のメタデータ管理手法の検討

清重 周太郎¹⁾, 三上 絢子¹⁾, 込山 悠介^{2),3)}, 長谷川 晃⁴⁾

1) 北海道大学 附属図書館 研究開発室

2) 国立情報学研究所 コンテンツ科学研究系

3) 国立情報学研究所 オープンサイエンス基盤研究センター

4) 北海道大学 附属図書館

kiyoshige@lib.hokudai.ac.jp

Study of Metadata Management Method of Research-Strategy Activity Logs

for Mutual Understanding among Research Supporting Sections

Shutaro Kiyoshige¹⁾, Ayako Mikami¹⁾, Yusuke Komiyama^{2),3)}, Ko Hasegawa⁴⁾

1) Research and Development Department, Hokkaido University Library

2) Digital Content and Media Sciences Research Division, National Institute of Informatics

3) Research Center for Open Science and Data Platform, National Institute of Informatics

4) Hokkaido University Library

概要

中期目標・中期計画に基づいて大学が取り組む研究力強化の取り組みにおいては、技術部・URA 局等の支援機関や研究推進部等の事務部署の連携が重要であり、連携を長期的に継続するためには相互理解が不可欠である。しかしながら、現状では情報共有のためのコストが大きく、必ずしも互いの得手不得手を把握できていない。研究力強化の取り組みの実施記録は年報や評価点検報告書等の形で存在するため、これらを対象として図書館情報学における情報組織化の技法を応用し、同一規格のメタデータを付与することで蓄積・検索・再利用を可能とする体系化の手法を検討した。

1 はじめに

1.1 背景

近年の学術研究機関は財源・人員・時間に制約があり、これは2018年の科学技術・学術政策研究所(National Institute of Science and Technology Policy: NISTEP)による『科学技術の状況に係る総合的意識調査(NISTEP 定点調査)』の中でも報告されている[1]。同時に、組織内では研究の国際化・学際化の推進に伴って学内外の連携を強化する動きがある。北海道大学における研究力強化の取り組み(以下、「研究事業」という)の例として、研究推進部、University Research Administrator (URA)、技術本部、附属図書館等の研究支援を担う部署間で設備・人員の共有が進められてきた[2][3]。部署間での連携を長期的に継続するためには各部署の相互理解が不可欠であり、事業内容、背景や課題意識などの情報共有が重要である。

しかしながら、北海道大学においては部署間の情報共有の手法は伝統的なメール・打ち合わせが中心で、個人間のやり取りによるところが大きく、また効率的ではない面がある。言い換えると「連携によって事業の効率化・高度化を図るための取り組み」を行うために必要な情報共有することに労力がかかる」というジレンマである。2019年に林豊らは、この問題を『コミュニケーション、タスク管理、ナレッジ蓄積をよりオープンかつスピーディに行い、業務の効率化及び活性化、引いては図書館サービスの向上を実現できないか』とし、大学図書館業務の効率化の観点から提唱している[4]。

一方で、図書館情報学では、20世紀初頭にRanganathanが『利用者の時間を節約せよ』[5]と提言した事を起点に、情報(媒体としては図書・雑誌等の資料)の概要と所在を含むメタデータを、グローバルな統一規格で蓄積し、これを検索可能とする仕組みを整備してきた歴史があり、これを「情報組織化(information

organization)」という[6]. さらにインターネットの普及とデータの電子化が進んだ現在、情報の発見性・アクセス性・再利用性は高度化しており、情報探索者の時間的コストの縮減に繋がっている[7].

また、業務のノウハウ蓄積とその継承や再利用は企業経営においても重要な課題であり、2015年に深尾らは『過去のプロジェクトの豊富な現場実績による実現知の再利用によるスピード・コスト・品質のさらなる向上』と表現し、『提案書を作成する現場 SE の探索行動を分析し、探索に有効なメタデータ』の生成技法を報告している[8]. 他にも、早矢仕らが 2016 年に提唱した組織を超えたデータの共有技法としてメタデータを用いる事例もある[9]. ただし、メタデータ “Data Jacket” が人対人のデータ価値の相互理解を目的として設計されているように、目的によってメタデータの設計が異なる点は注意が必要である. 加えて、情報共有による事務業務改善の視点では、九州大学附属図書館が Backlog 試験導入を受けて『重要なポイントは「ドキュメント化」と「オープン化」の 2 点』、『多くの人が低負荷で習慣的にドキュメント化を続けられるような職場環境』、『部署間の垣根を超えて遠慮なく口と手を出していけるようなオープンなカルチャー』と解決の方向性を示している[3].

本研究では、研究支援部署間の情報共有コストの縮減を目標として、図書館情報学における情報組織化の技法を転用し、研究事業の実施記録の網羅的な蓄積および検索を実現するべく、“研究事業実施記録の組織化”(以下、「体系化」という)のフレームワークの構築手法について述べる.

2 事業実施記録メタデータの設計

2.1 設計の方向性

本研究では、体系化の対象を次のように定義する. 第一に、大学の「中期目標・中期計画・年度計画」、第二に大学が計画に基づき実施した「研究事業の活動報告」である. これらに対して、各活動の内容や相互の関連を表現可能であるメタデータの設計を行った. このメタデータを「事業実施記録メタデータ」という. メタデータ入力は研究支援部署の事務担当者が行うことを想定しており、作業負担軽減のため、自由記述を極力減らし、語彙群からの選択式や ID 入力を中心とした. 人物の表現には ORCID iD[10], または researchmap のハンドル[11]を、学術研究機関の表現には ROR ID[12], または『NISTEP 大学・公的機関名辞書』[13]の ID を使用する.

2.2 基底とするメタデータ規格

事業実施記録メタデータの設計には、学術研究分野において既に規格化されている JPCOAR スキーマと IEEE 1484.12.1 を基底とする.

JPCOAR スキーマは学術成果物である論文や研究データ(論文のエビデンスデータ)への付与を目的として、2017 年にオープンアクセスリポジトリ推進協会 (JPCOAR : Japan Consortium for Open Access Repositories) によって策定された[14]. 大園らが述べるように、オープンサイエンスや国際的な相互運用性に対応した規格であり、事業実施記録メタデータの記述においても有用な特徴をもつ.

IEEE 1484.12.1 Draft Standard for Learning Object Metadata は e-Learning コンテンツを対象として、2002 年に IEEE によって策定された国際規格である[15].

3 サンプルデータの作成

著者らは、北海道大学が 2012 年度から 2017 年度に実施した事業計画・活動について、ウェブ公開情報を情報源に収集し、整備過程で判明した情報のうち、JPCOAR スキーマに対応する項目があるものをメタデータとして記述した.

3.1 中期目標・中期計画・年度計画データ

北海道大学の『中期目標・中期計画・年度計画及び実績』(2012 年度から 2017 年度)[16]の記載内容から、項番に基づきリスト化を行い、独自に ID を付与した. メタデータ項 “Relation” には、項番が上位であるデータの ID を記述し、関連リンクを整備した.

3.2 研究事業の活動報告データ

ウェブ上で公開されている PDF ドキュメントおよびウェブページからテキストを収集した[17]. 日付を起点に研究事業活動記録をリスト化し、独自に ID を付与した. ウェブ上で報告が見られない基盤的な研究支援活動については、北海道大学内でヒアリングを行いリスト化した. 加えて、『業務の実績に関する報告書』(平成 24 年度から 29 年度)[18]の記載情報に基づき、研究事業活動に対応する年度計画データの ID を “Relation” に記述し、関連リンクを整備した.

4 メタデータ項目の追加定義

4.1 『情報の 6W2H』への拡張

2015 年に深尾らによって『情報の探索者が求める情報の 6W2H : What, Who, When, Where, Why, for Whom, How, How much』[8]という分析が提唱された. 作成したサンプルデータにおいてどのメタデータ項が 6W2H に該当するかを分析したところ, “Why”, “for Whom”, “How much”

の表現が欠けていると考えられた。したがって“Why”の表現として“Objective”，“for Whom”を“Target”，“How much”を“Input Cost”として独自に定義した。

4.2 “Relation”における述語の追加

JPCOAR スキーマにおいては“Relation”の性質を表現する述語が用意されている。研究事業の内容を表現するために、「共催」を意味する“Collaborated with”，「広報（PR：Public Relations）」を記述する“PR by”を独自に定義した。

4.3 メタデータ項のグループ化

JPCOAR スキーマのメタデータ項の順序は図書館情報学における「情報組織化」に基づいているため、IEEE 1484.12.1 を参考に事業実施記録メタデータ項のグループ化を行い、研究支援部署の担当者が理解しやすい順序とした。

4.4 “Subject”，“Activity Type”語彙群の作成

サンプルデータにおいて，“Subject”および“Activity Type”項の記述が多様化したため整理した。収集テキストにおける語彙の出現頻度をカウントし、高頻度またはユニークである語彙

から固有名詞(国・組織名称等)を差し引いた後、クラスタリングを行い、残った語彙群を使用してメタデータを修正した。

5 結果

本研究において設計した「事業実施記録メタデータ」の項目一覧を表 1 に示す。JPCOAR スキーマ項番の列において“**”と表記した項は本研究独自のメタデータである。研究事業活動の内容を表現するために，“Subject”および“Activity Type”項に用いた語彙群は表 2 のとおりである。

また、サンプルデータとして、中期目標・中期計画・年度計画 1,366 件、研究事業の活動記録 1,693 件、計 3,059 件のメタデータを作成した。一部を表 3 に示す。各データの関連性(リンク)を視覚化するために、グラフ可視化ツール Graphviz のレイアウトコマンド sfdp を使用してナレッジグラフとして描画した(図 1)[19]。

表 1 事業実施記録メタデータの項目一覧

グループ	項目名（英）	項目名（日）	JPCOAR スキーマ項番	入力形式	本研究における追加定義
1. General	ID	ID	17	英数字	-
	Title	タイトル	1	文字列	What に該当
	Alternative Title	その他のタイトル	2	文字列	-
	Subject	主題	9	選択式	What に該当
	Objective	目的	**	選択式	Why に該当
	Target	対象	**	選択式	for Whom に該当
	Description	内容記述	10	文字列	What に該当
	Language	言語	13	選択式	-
	Activity Type	手段	14	選択式	How に該当
	Temporal	時間的範囲	20	英数字	When に該当
	Geo Location	位置情報	21	21.1 を参照	Where に該当
	Point Longitude	経度	21.1.1	英数字	-
	Point Latitude	緯度	21.1.2	英数字	-
2. Rights and Permission	Access Rights	アクセス権	5	選択式	部署内、学内、公開
	Rights Holder	権利者情報	8	8.1 を参照	-
	Rights Holder Name Identifier	権利者識別子	8.1	選択式	ORCID または NISTEP 大学・公的機関名辞書等

3. Organizer	Creator	作成者	3	3. 1 を参照	Who に該当
	Creator Name Identifier	作成者識別子	3. 1	選択式	ORCID または NISTEP 大学・公的機関名辞書等
	Affiliation	作成者所属	3. 6	3. 6. 1 を参照	-
	Affiliation Name Identifier	所属機関識別子	3. 6. 1	選択式	ORCID または NISTEP 大学・公的機関名辞書等
4. Funding	Funding Reference	助成情報	22	22. 1 を参照	-
	Funder Identifier	助成機関識別子	22. 1	英数字	-
	Award Number	研究課題番号	22. 3	英数字	-
	Award Title	研究課題名	22. 4	22. 3 を参照	-
	Input Cost	投入コスト	**	英数字	How much に該当
5. Refarences	Relation	関連情報	19	19. 1 を参照	-
	Related Identifier	関連識別子	19. 1	英数字	-
6. Result and Output	File	ファイル情報	35	文字列	-
	URI	本文 URL	35. 1	英数字	-
	Date	日付	35. 4	英数字	-
7. Citations	Relation	関連情報	19	19. 1 を参照	-
	isReferencedBy	～で参照されている	19 の述語	いずれかを選択	5. Refarences においても同様の述語を用いる
	references	～を参照している			
	isFormatOf	～の別の記録形式である			
	hasFormat	～は以前から存在していた別の記録形式である			
	isIdenticalTo	～と同一である			
	isDerivedFrom	～に由来している			
	isSourceOf	～の由来になっている			
	Collaborated with	当該 ID を持つ機関と関連がある（共催等）	**		
	PR by	～の媒体で広報を行った	**		
	Related Identifier	関連識別子	19. 1	英数字	-

表 2 研究事業活動の内容を表現する語彙 (Subject, Activity Type)

大学運営	情報基盤	成果発信	社会連携	自然科学	セミナー企画
大学教育	安全管理・倫理 コンプライアンス	成果公開	他機関連携	社会科学	コンテンツ化
中期目標	点検評価・情報公開	産学連携	オープンファシリティ	情報収集・意見交換	研修受講
資源基盤	事業評価	研究戦略	連携協定	研究環境	スキルアップ
基盤整備	資産管理	病院戦略	異分野連携	助成・顕彰	キャリアアップ
事業管理	知財管理	制度設計	国際力強化	技術提供	

表3 北海道大学における研究事業活動のメタデータ（サンプルデータからの一部抜粋）

ID	タイトル	作成者	所属	主題	言語	手段	関連情報 ID	本文 URL
NEN24_02_01_08	理・工・農系分野の女性教員の 新規採用及びワークライフバ ランスに配慮した雇用・育成を 継続的に推進するとともに、研 究キャリアの継続及び向上に 必要と考えられる支援策のさ らなる展開と次世代女性研究 者の育成を行う。	-	北海道 大学	大学運営	JP	事業管理	CHM2_02_01	https://www.hokudai. ac.jp/introduction/p lan/folder888/folder 1/
ACT24_F0 HRED_149	Hi-System, 研究者管理 システム	上級人材育 成ステーシ ョン	人材育 成本部	キャリア アップ	JP	制度設計	ACT24_ResDev	http://www2.synfoste r.hokudai.ac.jp/
ACT24_F0 HRED_191	北大理系女子コミュニ ティ “RingS”	女性研究者 支援室			JP	事務執行	NEN24_02_01_08	https://freshu.ist.h okudai.ac.jp/
ACT24_CR IS_107	分析受託	グローバル ファシリテ ィセンター	創成研 究機構	オープン ファシリ ィティ	JP	技術提供	NEN24_01_02_02_05	https://www.gfc.hoku dai.ac.jp/
ACT30_CR IS_126	第5回 オープンファ シリィシンポジウム				JP	セミナー 企画	NEN29_01_02_02_01_02_01	

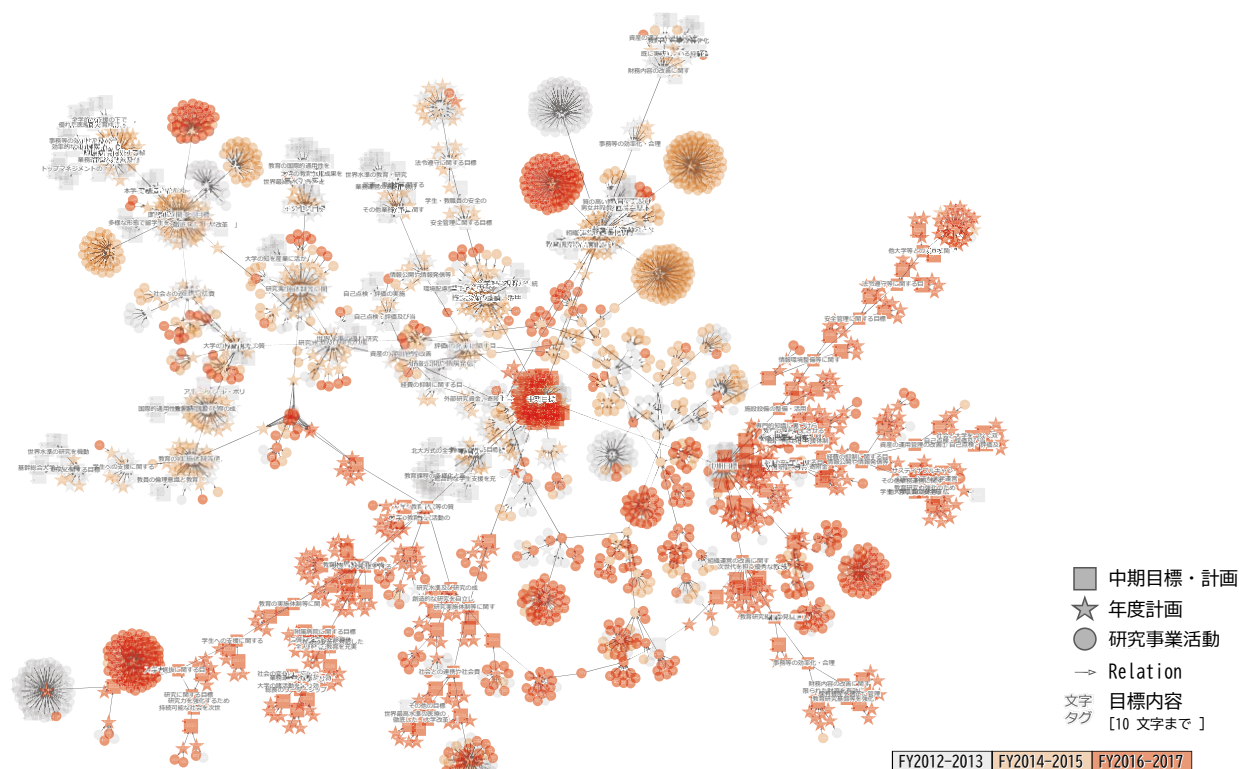


図1 北海道大学における中期目標・中期計画・年度計画と研究事業活動の
関係性リンクに基づくナレッジグラフ

4 まとめと今後の展望

本研究では、事業実施記録メタデータの定義を行い、体系化のフレームワークの構築について検討した。また、北海道大学がウェブ上で公開している事業実施記録を対象に、図書館情報学における情報組織化の技法を転用し、サンプルデータとして 3,059 件のメタデータを作成した。今後の課題として、実務フローへの導入および利活用の展望について述べる。

本研究のサンプルデータはウェブ上の公開情報を収集したが、運用段階では研究支援部署の事務担当者が実務の中で直接記録を行うため、実用的なツールが求められる。国立情報学研究所による「GakuNin RDM」は、研究データの収集・蓄積を担うほか、機関リポジトリへのスムーズな移行・公開の機能を有する持続可能な研究データ管理サービスである[20]。また、アドオン開発が可能であるため、事業実施記録メタデータを JPCOAR スキーマへと読み替える翻訳機能を含む、直感的に使いやすいインターフェースを開発することで、研究プロジェクトを管

理するように研究事業のプロジェクト管理が可能と考えられる(図 2)。このアドオンを「活動情報メタデータ管理インターフェース：ACtivity Metadata Management Interface (ACMMI)」と仮称し、開発に取り組む。

また、蓄積され検索可能となった事業実施記録の利活用は、利用者の目的に応じて多岐にわたると考えられる。北海道大学において考えられる例としては、研究支援部署間における業務改善のための相互理解や、大学と企業間における産学連携のための相互理解への利用がある。一方で大学運営の視点からは、研究資金配分機関および国や社会への説明責任としての情報公開のための利用、または大学 IR(Institutional Research)への拡張が考えられる。

深尾らと早矢仕らによる先行研究[8][9]が示すように、利用目的によってメタデータの各項目の重要性は変動する。今後は大学関係者の様々な立場や視点からフィードバックを得ながら、あらゆる研究事業の表現に対応し、利活用の可能性を広げるメタデータの設計を進め、ACMMI の実現に繋げたい。

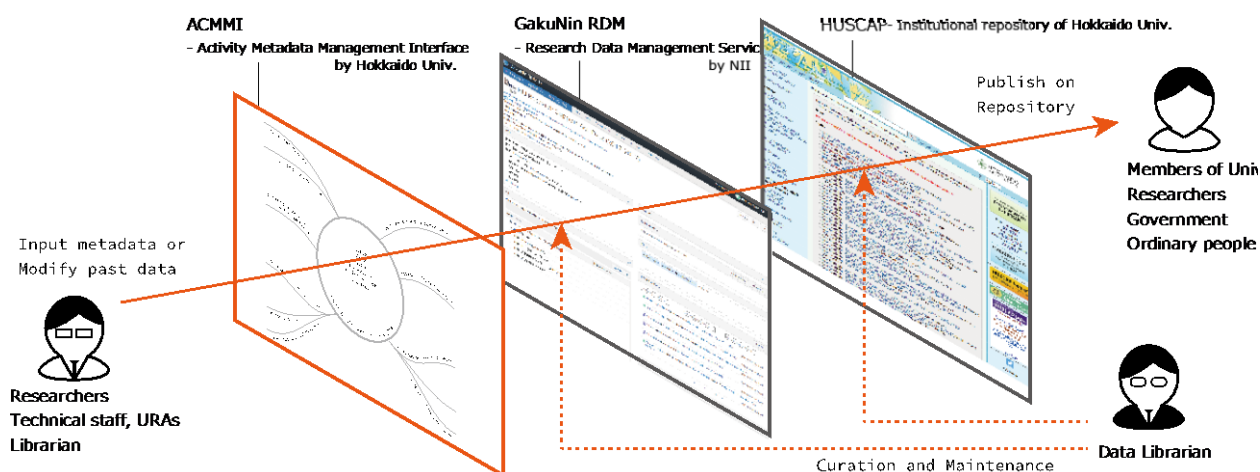


図 2 活動情報メタデータ管理インターフェース (ACMMI) と研究データ管理サービスの連携

謝辞

本研究は、2019 年度国立情報学研究所公募型共同研究（19S0303）の助成を受けている。また、サンプルデータの収集においては、北海道大学附属図書館事業 DB 作業部会のメンバーに多大なる助力をいただいた。この場を借りて感謝申し上げる。

注・参考文献

- [1] 科学技術・学術政策研究所，“科学技術の状況に係る総合的意識調査(NISTEP 定点調査2018)報告書”，
<http://doi.org/10.15108/nr179> (参 照 2019-9-10)
- [2] 技術支援本部発足(2013):北海道大学技術支援本部，“技術支援本部について”，
<https://tech.general.hokudai.ac.jp/organization/outline/>(参照 2019-9-10)
- [3] グローバルファシリティセンター発足(2016):北海道大学グローバルファシリティセンター，“GFC パンフレット(2019 年度版)”，
https://www.gfc.hokudai.ac.jp/wp-content/uploads/2019/07/gfc_Pamphlet_201907.pdf(参 照 2019-9-10)
- [4] 林豊，大村武史，堀優子，“九州大学附属図書館におけるプロジェクト管理ツールの導入”，九州大学附属図書館研究開発室年報，2018/2019，p.14-20，2019.
- [5] 原文 “Save the time of the reader.” : Shiyali Ramamrita Ranganathan , “ Five Laws of Library Science ” , 1931.
- [6] 高山正也，植松貞夫監修，“情報資源組織論”，p.2，樹書房，2011.
- [7] 原聡子，片岡真，“ディスカバリーサービスとこれからの図書館”，図書館雑誌，108(3)，p185-187，2014.
- [8] 深尾浩ほか，“情報探索者の探索行動分析に基づく情報探索効率化のためのメタデータ体系化技法の開発”，人工知能学会知識・技術・技能の伝承支援研究会(SIG-KST-026-03)，2015.
- [9] 早矢仕晃章，大澤幸生，“Data Jacket Store: データ利活用知識構造化と検索システム”，人工知能学会論文誌，31(5)，A-G15_1-9，2016.
- [10] ORCID, “Format of the ORCID iD”，
<https://support.orcid.org/hc/en-us/articles/360006897674>(参照 2019-9-10)
- [11] 仕様では“permalink”:国立研究開発法人 科学技術振興機構 (JST), “WebAPI について”，
<https://researchmap.jp/public/about/rmap2/function/WebAPI/>(参照 2019-9-10)
- [12] Research Organization Registry : “ ROR announces the first Org ID prototype ” ,
<https://www.crossref.org/blog/ror-announces-the-first-org-id-prototype/> (参照 2019-9-10)
- [13] 科学技術・学術政策研究所，“NISTEP 大学・公的機関名辞書”，
<https://www.nistep.go.jp/research/scisip/randdon-university/>(参照 2019-9-10)
- [14] 大園隼彦ほか，“JPCOAR スキーマの策定：日本の学術成果の円滑な国際的流通を目指して”，情報管理，60(10)，p.719-729，2018.
- [15] Marilyn McClelland, “Metadata Standards for Educational Resources”，Computer，36(11)，p.107-109，2003.
- [16] 北海道大学，“中期目標・中期計画・年度計画及び実績”，
<https://www.hokudai.ac.jp/introduction/plan/folder888/>(参照 2019-9-10)
- [17] テキスト収集に用いた PDF ドキュメントとウェブページの一覧を『附録』に示す.
- [18] 北海道大学，“業務の実績に関する報告書”，
<https://www.hokudai.ac.jp/pr/tenken/hyouka/houjin/2nd/>(参照 2019-9-10)
- [19] Graphviz - Graph Visualization Software :
<http://www.graphviz.org/documentation/>(参 照 2019-9-10)
- [20] 2020 年 9 月末まで実証実験中である :
<https://rdm.nii.ac.jp>

[illegible]

序 号	姓 名	学 号	考 试 成 绩	备 注
-----	-----	-----	---------	-----

--	--	--	--	--