

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 20 日現在

機関番号：62615

研究種目：基盤研究(C) (一般)

研究期間：2015～2017

課題番号：15K00446

研究課題名(和文) 研究データリポジトリにおける時間軸を意識した版管理モデルの開発と実装

研究課題名(英文) Development of Metadata Version Control Model in Research Data Repository

研究代表者

林 正治 (Hayashi, Masaharu)

国立情報学研究所・サイバーセキュリティ研究開発センター・特任助教

研究者番号：90552084

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)：本研究では、研究データリポジトリにおけるメタデータの版管理手法およびバージョン管理された研究データ引用手法を提案し、プロトタイプシステムによる機能検証を実施した。版管理手法では、コンテンツおよびメタデータの双方の版を管理することとした。永続識別子はランディングページ毎に付与し、Memento Frameworkによる版管理情報の提供を実施した。研究データ引用手法では、Citation Style Languageによる引用情報の提供手法についての検討を実施した。

研究成果の概要(英文)：In this research, we propose version control and citation method of research data in research data repository and develop proto-type system using these methods. Version control method for research data targets content file and its metadata. A persistent identifier is given for each landing page. version control method for research data repository is developed based on Memento framework and provides versioning information of research data to machines. In the research data citation method, we examined the method of providing citation information by Citation Style Language.

研究分野：図書館情報学

キーワード：機関リポジトリ 版管理 メタデータ

1. 研究開始当初の背景

近年、論文とその根拠となる研究データを一緒に公開することで、研究過程を明らかにし、論文の再現性と再利用性を高めることを目的とした試み、オープンサイエンス[1]に注目が集まっている。これまでも研究データを共有する試みはあったが、オープンサイエンスの特徴はウェブという共通基盤を前提としている点にある[2]。たとえば、DataCite[3]は、データへのアクセス性の確保、データの再利用と検証の促進を目的とした、一意で持続的なデジタルオブジェクト識別子(DOI)のデータへの付与を行っている。figshare[4]および Dryad[5]は、研究データのためのデジタルリポジトリ(研究データリポジトリ)機能を提供し、それまで論文の一部でしかなかった図表やデータを研究成果として公開する仕組みを提供している。しかしながら、これらの仕組みは、研究成果としての論文の流通基盤であるデジタルリポジトリに係わる技術を転用した仕組みであり、研究成果としてのデータの流通基盤、すなわち真の研究データリポジトリを実現するには、データ固有の問題に取り組む必要がある。とくに、論文の場合と比較してデータは動的であり、研究データリポジトリにおけるデータおよびメタデータ(データに関する説明)の版管理は重要な課題の一つである[6]。一般的に論文の公開後、論文の内容や論文のメタデータが修正されることは殆どないが、研究データについてはその可能性を想定しなければならない。たとえば、観測機器を用いて取得したデータについては、機器の特質や観測時の状況誤差を補正するための校正作業により、データは大きく異なるものとなる。研究過程で校正作業の不備が発覚したり、新たな校正アルゴリズムが発見されたりする可能性もあり、データの修正を想定しておく必要がある。また、メタデータが公開後に修正される可能性もある。たとえば、博物館資料や古文書資料等に関するメタデータについては、対象物が未知であることから、研究過程で修正されたり、補充されたりする可能性がある。すなわち、メタデータ自体が研究過程の産物であり、公開後のメタデータ修正を想定しておかなければならない。

以上のように、研究データリポジトリが対象とする研究データおよびメタデータについては、公開後の修正についての想定が必要で、公開から引用、修正までの時間軸を意識した版管理の実現および版管理を意識したデータ引用モデルの確立が重要である。また、版管理においては、過去のデータ内容を保証し、引用元にデータを適切に提供する仕組みも必要である。申請者らは、国内におけるデータリポジトリの先駆者であり、これまでの研究およびデータリポジトリに関する研究会「学術資源リポジトリ協議会」での議論から、研究データリポジトリにおける時間軸を意識した版管理モデルの確立が必要である

との結論に至った。

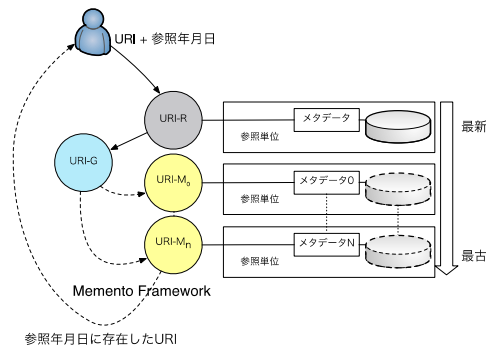


図1 時間軸を意識した版管理モデル

2. 研究の目的

本研究では研究データリポジトリにおける時間軸を意識した版管理モデルの開発、およびそれを実現する基盤技術の開発を目的とする。具体的には次の2つの課題に取り組む。

(1) 研究データリポジトリにおける時間軸を意識した版管理モデルの開発

本課題では研究データの公開から引用、および修正までの各時間軸における研究データの版管理モデルの確立および実現のための基盤技術の開発を目的とする。また、引用時のデータを保証し、論文公開後も引用時のデータを適切に提供する仕組みの確立も重要課題の一つである。本研究においては、時間軸を意識したHTTPアクセス手法の一つであるMemento Framework[7]と電子署名等による改ざん検知機能を備えたデジタルリポジトリにおける時間軸を意識した版管理モデルの実現に取り組む(図1参照)。Memento Frameworkは、時間軸の概念を導入したHTTPプロトコルであり、HTTPアクセス時に指定された時間軸に対応したウェブのアーカイブを返すために考案された仕組みである。本研究では、この仕組みを応用することで、ある時点で引用されたデータに対するアクセスに対して、適切なデータおよびメタデータの版を返すことが可能となる。また、研究データ公開時の時間と内容に対しての電子署名を行うことで、引用時のデータから意図しない改ざんが行われたか否かの検知が可能となり、引用時のデータおよびメタデータの内容を保証する仕組みを提供することが可能となる。

(2) 研究データリポジトリにおけるデータ引用モデルの確立

本課題では「時間軸を意識した版管理モデル」におけるデータ引用モデルの確立を目的とする。国内外の存在する研究データリポジトリの状況および研究データリポジトリを利用する研究者を調査し、データの引用がどのように行われるのか、データの種類により引用方法がどのように変化するのか、複数の版が存在するデータに対してどのように対

俊弘, 池田 大輔, 行木 孝夫, 山地 一禎: 次世代リポジトリプラットフォームに求められる機能像 ~ オープンソースリポジトリソフトウェアの比較調査を通して ~ , 情報知識学会誌 , 査読無 , 27 , pp. 366-369 , 2017 .

堀井 洋, 堀井 美里, 上田 啓未, 林 正治, 高田 良宏, 山地 一禎: 学術資料を対象にしたサブジェクトリポジトリの構築 - 科学実験機器資料および教育掛図資料を事例として - , 情報知識学会誌 , 査読無 , 27 , pp. 357-361 , 2017 .

根本 しおみ, 高田 良宏, 堀井 洋, 堀井 美里, 飯野 孝浩, 林 正治: DOI を用いた天文学研究資料の横断的な整理・公開 - 岡山天体物理観測所を一例として - , 情報知識学会誌 , 査読無 , 27 , pp. 347-352 , 2017 .

高田 良宏, 古畑 徹, 林 正治, 堀井 洋, 堀井 美里, 上田 啓未: 金沢大学における研究データ公開用リポジトリの構築の試み, 情報知識学会誌 , 査読無 , 27 , pp. 337-342 , 2017 .

高田 良宏, 古畑 徹, 林 正治, 堀井 洋, 堀井 美里, 上田 啓未: 金沢大学資料館ヴァーチャル・ミュージアムの 開発思想と構築の歩み, 情報知識学会誌 , 査読無 , 27 , pp. 343-346 , 2017 .

堀井 洋: 科学実験機器資料サブジェクトリポジトリプロジェクトの現状と課題, 大学の物理教育 , 査読有 , 23 , pp. 27-29 , 2017 .

林 正治, 夏目 琢史, 松田 訓典, 山本和明, 赤木 完爾: ウェブアノテーションを用いた仮想コレクション定義の試み-二つの「幸田文庫」を事例に-, 人文科学とコンピュータシンポジウム論文集 , 査読有 , 2017 , pp. 81-86 , 2016 .

古畑 徹, 高田 良宏, 堀井 洋, 林 正治, 堀井 美里, 上田 啓未: 金沢大学資料館におけるヴァーチャル・ミュージアム構築の歩み, 大学 ICT 推進協議会 2016 年度年次大会 (AXIES2016) 論文集 , 査読無 , WD36 , 2016 .

堀井 洋, 林 正治, 堀井 美里, 上田 啓未, 山地 一禎, 高田 良宏: 博物資料情報に対する DOI 付与の意義と展望, 情報知識学会誌 , 査読無 , 26 pp. 217-220 , 2016 .

林 正治, 田中 克明, 長谷 海平, 高見澤 秀幸, 松村 芳樹, 中島 康, 万代勝信: 学術的動画の評価傾向とその要因, 人文科学とコンピュータシンポジウム論文集 , 査読有 , 2015 , pp. 159-164 , 2015 .

高田 良宏, 林 正治, 堀井 洋, 堀井 美里, 山地 一禎, 山下 俊介, 古畑 徹: 非文献資料のための学術資源群によるサブジェクトリポジトリの構築 (構想と進捗状況), 大学 ICT 推進協議会 2015 年度年次大会論文集 , 査読無 , 2B1-8 ,

pp. 1-4 , 2015 .

[学会発表] (計 18 件)

林 正治, Re*poN からみた JAIRO Cloud の可能性, 第 3 回 CODH セミナー, 2017 .

根本 しおみ, 天文学資料リポジトリの構築 ~ 岡山天体物理観測所における現代天文学黎明期研究資料のリポジトリ構築 ~ , 平成 28 年度宇宙科学情報解析シンポジウム, 2017 .

堀井 洋, 学術資料情報の生成と活用に向けた取り組みについて, アート・ドキュメンテーション学会 研究会・見学会 現代アートとアーカイブの問題」 in 金沢 21 世紀美術館, 2017 .

高田 良宏, 研究資源の蓄積と利活用を目指した学術資源リポジトリについて ~ ヴァーチャル・ミュージアムからサブジェクトリポジトリ、そして今後 ~ , 人文科学とコンピュータシンポジウム, 2016 .

仲山 悠也, STP 観測データ公開・解析用データリポジトリと次世代認証技術の研究, 第 332 回生存圏シンポジウム「宇宙プラズマ波動研究会」, 2016 .

林 正治, 歴史資料の情報化とその可能性, 2016 年度加賀藩研究ネットワーク大会, 2016 .

林 正治, 研究資料の蓄積と共有を目指した学術資源リポジトリについて ~ ヴァーチャル・ミュージアムやサブジェクトリポジトリの構築・運用経験から ~ , 金沢大学国際文化資源学研究中心 第 8 回文化資源学フォーラム, 2016 .

林 正治, 古典籍画像のための実験的ウェブアノテーションツール, 第 2 回日本語の歴史的典籍国際研究集会, 2016 .

高田 良宏, オープンサイエンス時代に向けた非文献資料のための学術資源群によるサブジェクトリポジトリの構築, 第 11 回博物科学会, 2016 .

高田 良宏, 科学研究費プロジェクトの報告と展望 ~ デジタル化のためのリポジトリから利活用のためのリポジトリへ ~ , 学術資源リポジトリ協議会 2016 総会, 2016 .

林 正治, 学術資源リポジトリ協議会における JaLC DOI 付与の試み, 国立情報学研究所学術情報基盤オープンフォーラム 2016, 2016 .

山下 俊介, 地域資料のアーカイブとフィールドワーク, 平成 27 年度公開セミナー「民具が語る白山麓の自然と生活」, 2016 .

山下 俊介, 映像ステーションの果たした役割と新しい映像ステーションへの期待, 第 3 回京都大学研究資源アーカイブ研究会: 「さようなら映像ステーション, こんにちは映像ステーション: 稲盛財団記念館から総合博物館へ」, 2016 .

Shunsuke YAMASHITA, Applications of Video in Museums , 2015 University Museums Symposium on Museum in Everyday Life , 2015 .

Hiroshi Horii, An implication for categorizing and organizing the astronomical instruments from the museological Viewpoint , International Symposium on the NAOJ Museum , 2015 .

馬淵 嵩大, 科学データ解析・公開に最適なデータ公開システムの開発, 平成27年度電気関係学会北陸支部連合大会, 2015 .

堀井 洋, 高精細パノラマ撮影技術を利用した博物館展示空間記録・公開の試み, 2015 .

堀井 洋, 学術資源を活用した地域観光の実現に向けた試み, 第12回観光情報学会全国大会, 2015 .

6. 研究組織

(1) 研究代表者

林 正治 (HAYASHI, Masaharu)

国立情報学研究所・サイバーセキュリティ
研究開発センター・特任助教

研究者番号：90552084

(2) 研究分担者

高田 良宏 (TAKATA, Yoshihiro)

金沢大学・総合メディア基盤センター・准
教授

研究者番号：30251911

堀井 洋 (HORII, Hiroshi)

合同会社AMANE・調査研究ユニット
・代表社員

研究者番号：40372495

山下 俊介 (YAMASHITA, Shunsuke)

北海道大学・総合博物館・助教

研究者番号：50444451