

令和元年6月19日現在

機関番号：17102

研究種目：基盤研究(B) (一般)

研究期間：2015～2018

課題番号：15H02787

研究課題名(和文) 拡張性と簡便性を両立するデータ流通基盤の汎用的フレームワークと複数分野での実証

研究課題名(英文) An extensible, easy-to-use data framework in multi-disciplines

研究代表者

池田 大輔 (Ikeda, Daisuke)

九州大学・システム情報科学研究所・准教授

研究者番号：00294992

交付決定額(研究期間全体)：(直接経費) 10,500,000円

研究成果の概要(和文)：様々なデータを分野外の人でも簡単に使えることを目指し、データ流通基盤の概念的なフレームワークを提案し、複数分野に適用し評価する。複数分野の研究者や産業界を対象に調査を行い、ニーズの存在を明らかにした上で、概念的なモデルを構築した。既存の地理情報収集システムを改良し、複数のデータリポジトリを構築し、試験的に運用した(論文誌、国際会議)。インターフェイスについては複数分野の研究者向けと、教育など一般向けのデータ利用に関するプロトタイプを構築、実運用の上、国際学会等で発表を行った。データを拡充する手段として、論文から自動的に関連するデータセットを特定するアルゴリズムを構築した(国際会議)。

研究成果の学術的意義や社会的意義

研究期間内に、実際に検索を用いてメタデータを管理するフレームワークを、すでにあるクラウドサービス上で仮想的に実装し、個別化できることを示し、概念の検証を行った。また、非専門家の利用を含む、複数の分野での利用を前提にしたデータリポジトリの利用者インターフェイスも構築した。これらにより、異分野のデータを発見することが容易となり、また、データの利用方法の可視化にもつながり、研究の再現性の確保にもつながる。一方で、専門家が深く利用するには、まだ多くの課題があることも分かった。

研究成果の概要(英文)：Aiming to make variable data easily accessible from non-expert users, we propose a conceptual framework for data circulation and then evaluate by applying it to multi-disciplines. Through surveys we conducted to people in both academia and industry, we revealed the existence of needs for open data, and the proposed a conceptual framework. Modifying an existing GIS-like database, we implemented the framework and made test operations in multi-disciplines (presentations at an international conference). We also implemented interface parts for both experts in multi-disciplines and non-experts, and made test operations (journal papers and presentations at international conferences). To increase data on repositories, we developed an algorithm to extract dataset names from scholarly papers (presentation at an international conference).

研究分野：情報学

キーワード：データリポジトリ ベクトル空間モデル クイックルック 研究データ

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

研究の透明性や再現性、データ取得にかかるコストの回収などを目指してデータリポジトリの必要性は強く認識されているが、まだ一般的とは言い難い状況であった。実際、高度だが専門的な分野に特化したものが、一般向けだが機能が限られたものしか存在しない。文書を対象にした機関リポジトリでは、分野によらず対象とするデータは文書(論文)であり、形式も一定の書式が存在するため、検索も利用も容易である。一方で、データの扱いは分野に依存するため、特定の分野に特化したサイトはあるものの、データリポジトリは学術情報の流通基盤として成立していない状況にあった。

2. 研究の目的

様々なデータを分野外の人でも簡単に使えることを目指し、本研究では、データ流通基盤の概念的なフレームワークを提案し、複数分野に適用した実装により評価する。検索という簡便な手法でフレームワークを個別化し、分野別のリポジトリが得られる(図1参照)点が独創的で、フレームワークは分野を越えた学術情報流通基盤となることが期待できる。

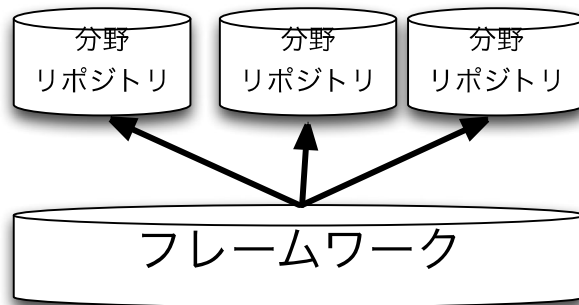


図1: 提案のフレームワーク

3. 研究の方法

これを達成するために、学術情報流通の基本的な仮説を設定し、これを検証する仮説検証型で研究を進める。仮説にあたる(1)フレームワーク、これを検証するプロトタイプを(2)データと(3)インターフェイスに分け、計3つの大項目を設定する。

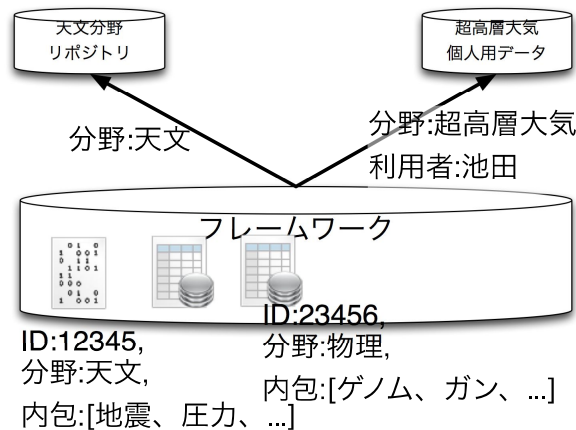


図2: 提案のフレームワーク

(1)では、代表者の研究シーズと複数分野から抽出したニーズとあわせて、概念的なモデル提案(図2参照)を目指す。研究分野に加えて、産業界からもニーズの調査を行い、データリポジトリの必要な形を明確にする。プロトタイプは、検証に加え、ニーズ抽出の時にも用い、現実的な議論の工夫とする。(2)は、データ自体に加え、その表現となるメタデータの拡充も含む。(2)および(3)で、データの確認や検証が簡単に可能になるようにする。互いの項目は関連しているが、仮のデータやモックアップ等を用いて、可能な限り独立に進め、外注も利用して効率よく研究を進める。

4. 研究成果

(1)のモデル化の一環として、これまでのリポジトリ機能の調査し、歴史的な位置付けを行った。また、複数分野の研究者や産業界からのニーズについて調査を行い、ニーズの存在を明らかに

した。その上で、概念的なモデルを、クラウド上の実サービスを用いて仮想的に実装し、概念検証を行った(国際会議発表)。複数の分野での検証は行ったものの、小規模なデータに留まるため、大規模な利用時における課題等の検証は今後の課題である。

(2)について、機関リポジトリ等で公開されている論文から、自動的に関連するデータセットを特定するアルゴリズムを構築した(国際会議発表)。

(3)については、複数分野の研究者向けと、教育など一般向けのデータ利用に関するインターフェースのプロトタイプを構築の上、国際学会等で発表を行った。

5. 主な発表論文等

[雑誌論文](計6件)

1. Development of food culture database system using Kataribe Cloud to support open science revise, Go Hirakawa and Daisuke Ikeda, Proceedings of the 8th International Conference on Data Science and Institutional Research, 2019. (in print) 査読有
2. Toward Automatic Identification of Dataset Names in Scholarly Articles, Daisuke Ikeda and Yuta Taniguchi, Proceedings of the 8th International Conference on Data Science and Institutional Research, 2019. (in print) 査読有
3. 地域科学技術指標 2018、荒木寛幸、野澤博一、科学技術・学術政策研究所 調査資料, RM278, 2018. 査読無
4. 地域イノベーションシステムに関する意識調査(2016)の要因分析, 荒木寛幸, 科学技術・学術政策研究所 DISCUSSION PAPER 165, 2018. 査読無
5. 再現と共有を可能とするデータ可視化方法について、今井 弘二, 村山 泰啓, 海老沢 研, 池田 大輔, 瀬口 大介, 宇宙科学情報解析論文誌, Vol.7, pp.1-7, 2018. 査読有
6. 地域イノベーションシステムに関する意識調査報告, 荒木寛幸 犬塚隆志, 文部科学省科学技術・学術政策研究所 調査資料 260, 2017. 査読無

[学会発表](計15件)

1. (invited)Text Mining with Variety and Frequency, Daisuke Ikeda, 第28回MRS日本年次大会国際ワークショップ「先端プラズマ技術が拓くナノマテリアルズフロンティア」(国際学会), 2018
2. (invited)Toward Improving Research Processes using Big Data, Daisuke Ikeda, 2nd Asia-Pacific Conference on Plasma Physics(国際学会), 2018
3. Historical Transition of Repository Functions and Latest Trends of Repository System Architecture, Masaharu Hayashi, Kazutsuna Yamaji, Yohei Hayashi, Kosuke Tanabe, Toshihiro Aoyama, Takao Namiki, Daisuke Ikeda, The 13th International Conference on Open Repositories(国際学会), 2018.
4. Data visualization service for the earth science, 今井 弘二、村山 泰啓、海老沢 研、池田 大輔, Japan Geoscience Union Meeting(国際学会), 2018.
5. Reproducible and Shareable Data Visualization Method, 今井 弘二、村山 泰啓、池田 大輔, Japan Geoscience Union Meeting(国際学会), 2018.
6. Consideration of the metadata schema extension for upper atmospheric research field, Morio YAMAUCHI, Yukinobu KOYAMA, Japan Geoscience Union Meeting(国際学会), 2018.
7. Web service for reproducible multidisciplinary data visualization, Koji IMAI, Yasuhiro MURAYAMA, Ken EBISAWA, Daisuke IKEDA, Daisuke KITAO, International Workshop on Sharing, Citation and Publication of Scientific Data across Disciplines(国際学会), 2017.
8. Web service for reproducible multidisciplinary data visualization, Koji IMAI, Yasuhiro MURAYAMA, Ken EBISAWA, Daisuke IKEDA, Daisuke KITAO, WDS Asia-Oceania

Conference 2017(国際学会), 2017.

9. Web service for reproducible multidisciplinary data visualization, Koji IMAI, Yasuhiro MURAYAMA, Ken EBISAWA, Daisuke IKEDA, Daisuke KITAO, 日本地球惑星科学連合 連合大会 2017 年大会(国際学会), 2017
10. 考古地磁気学と考古学の交流のための発掘情報共有ツールの設計, 北原 優、平川 剛、池田 大輔, 日本地球惑星科学連合 連合大会 2017 年大会, 2017
11. 地域イノベーションシステムに関する意識調査における考察 地域の自己認識の考察 , 荒木 寛幸, 第 32 回研究・イノベーション学会, 2017
12. CGI クエリでコントロールするウェブシステム開発の試み, 今井 弘二、村山 泰啓、海老沢 研、池田 大輔、瀬口 大介, 宇宙科学情報解析シンポジウム, 2017.
13. 豊富な検索語で検索可能なデータリポジトリの構築に向けて、池田 大輔, 第 3 回オープンサイエンスデータ推進ワークショップ, 2016.
14. Web service for reproducible multidisciplinary data visualization, Koji IMAI, Yasuhiro MURAYAMA, Ken EBISAWA, Daisuke IKEDA, Daisuke KITAO, SciDataCon(国際学会), 2016.
15. 地球惑星科学(地学)の新しいウェブサービス, 今井 弘二、松下 幸市朗、海老沢 研、池田 大輔, 日本地球惑星科学連合, 2016.

〔図書〕(計 件)

〔産業財産権〕
出願状況(計 件)

名称：
発明者：
権利者：
種類：
番号：
出願年：
国内外の別：

取得状況(計 件)

名称：
発明者：
権利者：
種類：
番号：
取得年：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1)研究分担者
研究分担者氏名：小山 幸伸
ローマ字氏名：Koyama, Yukinobu
所属研究機関名：大分工業高等専門学校
部局名：情報工学科
職名：講師
研究者番号(8桁)：50598513

研究分担者氏名：今井 弘二
ローマ字氏名：Imai, Koji

所属研究機関名：国立研究開発法人情報通信研究機構
部局名：戦略的プログラムオフィス
職名：研究員
研究者番号（8桁）：50711230

研究分担者氏名：荒木 寛幸
ローマ字氏名：Araki, Yukihiro
所属研究機関名：文部科学省科学技術・学術政策研究所
部局名：その他部局等
職名：上席研究官
研究者番号（8桁）：80542026

(2)研究協力者

研究協力者氏名：比良松 道一
ローマ字氏名：Hiramatsu, Michikazu

研究協力者氏名：平川 剛
ローマ字氏名：Hirakawa, Go

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、研究者個人に帰属されます。