

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 16 日現在

機関番号：62615

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500137

研究課題名(和文)論文を構成するサプリメンタルデータを共有・再利用するための基盤技術開発

研究課題名(英文)Development of sharing and reuse technology on supplemental data of the research paper

研究代表者

山地 一禎(YAMAJI, Kazutsuna)

国立情報学研究所・コンテンツ科学研究系・准教授

研究者番号：50373379

交付決定額(研究期間全体)：(直接経費) 4,200,000円

研究成果の概要(和文)：本研究では、論文のサプリメンタルデータを共有するためのコンテンツ流通・保護システムの構築を目的とした。まず、論文と紐づく成果公開プラットフォームの開発を実施した。従来開発したリポジトリシステム「WEKO」をベースとし、サプリメンタルデータを登録するリポジトリを構築できるシステムへと拡張した。次に、登録されたコンテンツに軽量な方法で、商用タイムスタンプを取得する方法を実現した。また、文献リポジトリとサプリメンタルリポジトリにSAML認証を導入し、成果物のセキュアな公開環境を実現した。最後に、統合環境による実験を行い、所望の要件を満たすことを実証した。

研究成果の概要(英文)：This study proposed a secured scholarly communication system in order to facilitate sharing and reuse the supplemental data of the research paper. The publication platform of supplemental data was developed by enhancing functionalities of the previously developed repository system named WEKO. The existence proof and tamper resistant of the registered supplemental data can be secured by applying the newly developed light weight time stamping method. The method obtains timestamp token for the hashed-value of the supplemental data using the time stamping authority managed by a trusted third party. The system allow users to single sign on to both paper and supplemental data repository by means of the SAML authentication mechanism. The validity of the integrated system was confirmed through the experiment.

研究分野：学術コンテンツ流通

キーワード：オープンサイエンス データリポジトリ サプリメンタルデータ 学術コンテンツ流通 機関リポジトリ シングルサインオン SAML

1. 研究開始当初の背景

研究成果の発信を出版社のみに依存するのではなく、各大学図書館などでリポジトリを構築し、電子化された学術情報をアーカイブする動きが世界的に活発化している。国内でも約 130 のリポジトリが運営されており、コンテンツの総数は 100 万件を超える[1]、現在、国内の機関リポジトリに登録されているコンテンツの殆どが雑誌論文や紀要論文であり、ソフトウェアやデータ類、あるいは、教材といった文献以外のコンテンツの割合は、全体の 2%程度である[2]。これに対し、欧米のリポジトリでは、研究の過程を捉えるコンテンツの蓄積ならびに再利用に焦点があてられている[3]。実際に、SWORD[4]や OAI-ORE[5]といった新しいプロトコルの提案や、利用実験が活発に進められており、世界のリポジトリは、大学の業績を羅列するリポジトリから、研究や教育を推進する上で不可欠なリポジトリへの進化を遂げつつある。

現在、研究成果の主たる公開方法は論文であり、最もわかりやすい研究業績の表現形態となっている。したがって、論文以外の様々なコンテンツを共有・再利用する社会を構築する場合でも、論文を軸とした研究資源のネットワークを構成することにより、エンドユーザへの到達性が向上すると考えられる。また、論文というパブリックな研究成果と紐づくことにより、研究資源としての有用性が理解されやすくなると同時に、そのコンテンツの価値に繋がるものと期待される。

2. 研究の目的

研究代表者はこれまで、汎用的なリポジトリシステム WEKO を開発し[6]、情報系学会の電子ジャーナルプラットフォームとしての利用を進めてきた[7]。また、研究開発の発展として、より最新の研究情報発信源としてのプレプリントに注目し、プレプリント発行時の研究成果の先取権を確保するための、電子的な長期署名技術の開発を行ってきた[8]。ただし、これらはともに、論文 PDF による研究成果の発信を対象としたものである。世界の動向を加味した更なる課題としては、文献以外のコンテンツを、リポジトリ上でいかに扱うかを検討し、そのコンテンツ流通モデルを確立することが挙げられる。こうした新しい研究成果の発信モデルを検討する上では、論文のように研究者に受け入れやすい形で知の蓄積化を実現し、かつ、研究者が研究資源を提供するインセンティブを満足するフレームワークを、システム面と運用面から構築する必要がある。しかしながら、それに対する適切なモデルやケーススタディは世界的にも提案・報告されておらず、今後の展開が期待されている。

こうした課題への解決策として、論文のサプリメンタルデータは、様々な学術コンテンツが利用できる仕組みを構築する上で、利用者や提供者にとって最適な形態であると考

えられる。このとき、コンテンツの著作権管理にも配慮し、研究者が自分のコンテンツを安心して提供できる技術を導入することも重要な要素となる。そこで本研究では、研究代表者がこれまで構築してきた文献を主体とした研究成果発信システムを活用し、研究成果の新たな公開メディアとして、論文の資源であるサプリメンタルデータを安心・安全に公開できる、学術流通システムの基礎基盤を構築することを目的とする。

3. 研究の方法

本研究では、サプリメンタルデータと呼ばれる論文の補遺となる研究資源を共有するための、次世代の学術コンテンツ流通・保護システムを構築するために、その基礎基盤技術として、以下の3つの課題に大別して研究・開発・実証を行う。

(1) 課題1: 論文と紐づく成果公開プラットフォームの開発を実施する。本課題では、サプリメンタルデータを登録するリポジトリ(サプリメンタルリポジトリ)を WEKO の対象コンテンツを拡張することにより実現する。このとき、データの登録は、一般的には、研究者によるセルフアーカイビングを想定する。これは、コンテンツホルダーによる著作権管理をより明確にする上でも有効である。本課題においては、以下のポイントに焦点をあてた研究開発を実施する。

セルフアーカイブを容易にする HTML5 を採用した UI の仕様検討
サプリメンタルデータをインスタントに利用するためのフラッシュ対応
複数のサービスをシングルサインオンできる SAML 対応認証仕様の検討

(2) 課題2: 成果物を安心して公開できるセキュアな環境整備を実施する。具体的な実現方法としては、研究代表者が従来プレプリントサーバで実現してきた、研究成果へのタイムスタンプ付与技術を軽量化した方法を提案し、実装する。成果物の電子ファイルから計算したハッシュ値、あるいは、エンコード値そのものに対して署名を付与することにより、従来より軽量で、かつ、研究成果の先取性を保障できる技術を確立する。新たに構築するタイムスタンプ付与システムは、複数のリポジトリからのリクエストを受けることを可能とするために、タイムスタンプ取得・検証 API を実装する。

(3) 課題3: 提案するシステムの実証実験を実施する。研究代表者が学会向けに提供している情報系学会の電子ジャーナルプラットフォームの試験サイトに開発した機能を適用し、その実用性を検証する。

上記で具体的な方法を述べた、本研究で研究・開発する基盤と各課題との関係を図1に示す。

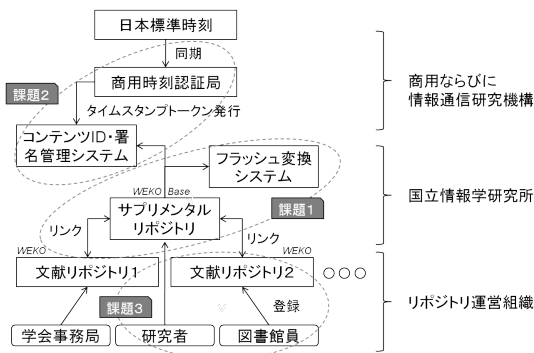


図 1 開発基盤の全体像と各課題との関係

4. 研究成果

(1) 課題 1 では、論文と紐づく成果公開プラットフォームの開発を実施した。従来の研究で、研究代表者は文献の登録をベースとしたリポジトリシステム「WEKO」の構築に取り組んできた。この WEKO をベースとし、サブリメンタルデータを登録するリポジトリ（サブリメンタルリポジトリ）を構築することができるシステムへと拡張した。

具体的には、まず、文献リポジトリとサブリメンタルリポジトリを連動させ、文献に関するサブリメンタルデータを研究者自身がセルフアーカイブできるフレームワークへと拡張した。次に、サブリメンタルデータの登録を容易にするために、HTML5 を採用した登録 UI の改良を行った。サブリメンタルデータの登録では、その操作性を高める目的で、複数のファイルをドラッグ&ドロップにて簡便に登録できるインターフェースを装備した。また、コンテンツのハンドリングを最適化することを目的に、HTML5 とフラッシュを融合するインターフェースを構築した。サブリメンタルリポジトリのバックエンドとして、フラッシュ変換サーバを用意することで、登録したコンテンツをインスタントに描画できるようにした。これにより、エンドユーザにコンテンツを直接ダウンロードさせたくない利用形態を満足できることになる。また、学会事務局のような組織が各論文のサブリメンタルデータを代理登録する場合も想定し、SWORD プロトコルを用いたサブリメンタルデータの一括登録機能も実装した。さらに、各リポジトリシステムを、SAML 認証に対応させた。この認証機構の導入により、文献リポジトリとサブリメンタルリポジトリをシングルサインオンによりシームレスに利用できる環境へと改良した。

(2) 課題 1 で拡張・改良したリポジトリシステムは、外部 ID サーバと連携して永続識別子としてのパーマネント URL を取得する機能を有する。この ID サーバの機能を拡張し、登録されたコンテンツのハッシュ値に対してタイムスタンプを取得するシステムを構築した。

具体的には、ID サーバから登録コンテンツのハッシュ値を受け取り、RFC3161 準拠のタイムスタンプサービスを利用してタイムスタンプトークンを取得する「タイムスタンプ取得 API」と、取得したタイムスタンプトークンとハッシュ値により改ざん等を検証した結果を返す機能を提供する「タイムスタンプ検証 API」を開発した。これらの機能は、ID サーバ内で運用するものとし、JSON 形式でのデータの通信により実現した。コンテンツの利用者は ID サーバの存在を意識することなく、リポジトリへのアクセスのみで検証結果を取得することができるインターフェースも用意した。コンテンツを登録した時刻は、商用の時刻認証局によって提供されるタイムスタンプトークンにより保証されるため、学術成果の先取性を確保するという目的にも威力を発揮する。

(3) 課題 3 では、開発したシステムを図 1 に示すような統合環境として実現し、各 API の実証実験と全体的なワークフローの確認を実施した。登録の方法としては、論文の著者が、自分の論文に対してサブリメンタルデータをセルフアーカイブする方法と、学会事務局のような代理登録者が一括して登録する場合を想定し、文献リポジトリとサブリメンタルリポジトリの連携機能、SAML によるシングルサインオン、SWORD による一括登録システムなど、それぞれに必要な機能が適切に動作することを確認した。また、タイムスタンプ機能に関しても、複数のリポジトリからのリクエストに API が適切に動作することを確認した。

本研究では、論文と紐づくサブリメンタルデータをセルフアーカイブするとともに、安心して公開するための基盤の構築を実施した。こうしたボトムアップ的なアプローチによるオープンサイエンスの実現に加えて、最近では、研究資金助成機関による、研究データのオープン化のための動きが活発化している。研究データ公開の義務化は、今では世界的な潮流であり、我が国の助成機関においても、近い将来において同様のポリシーをもつことが予想される。本研究の成果は、そうしたオープンサイエンス時代の学術情報流通において、研究者が自らのデータを安心・安全・簡便に公開するための基盤技術を提供するものであり、今後、より具体的な実社会での利用が期待される。また、新たなイノベーションを創出する目的で検討が進められているオープンサイエンスと並行して、現在、研究成果の改ざん、剽窃、不正利用などが社会的な問題となっている。機関による研究データの長期的な保全を検討する場合にも、研究データの真正性や先取性を担保できる本研究の成果は多面的な応用が期待される。今後も、開発した基盤技術に更に改善を加え、オープンサイエンスの礎となる研究開発に取り組みたいと考えている。

<引用文献>

- [1] 国内の機関リポジトリ一覧 :
<http://www.nii.ac.jp/irp/list/>
(accessed 2011-10-18)
- [2] 国内の機関リポジトリ統計情報 :
<http://irdb.nii.ac.jp/analysis/>
(accessed 2011-10-18)
- [3] The 5th International Conference on Open Repositories :
<http://or2010.fecyt.es/> (accessed 2011-10-18)
- [4] Allinson, J. et al.: "SWORD: Simple Web-service Offering Repository Deposit", Ariadne, Issue 54, 2008. Online at
<http://www.ariadne.ac.uk/issue54/>
- [5] OAI-ORE specification :
<http://www.openarchives.org/ore/>
(accessed 2011-10-18)
- [6] 山地一禎, 青山俊弘, 武田英明 : "学術資源共有基盤WEKOの開発", デジタル図書館, Vol.36, pp51-61, 2009.
- [7] 情報学広場プロジェクト :
<http://www.ixsq.nii.ac.jp/>
- [8] 山地一禎, 片岡俊幸, 行木孝夫, 曽根原登 : "プレプリントへの長期署名付とおよび検証システムの構築", 情報知識学会誌, vol.18, no.3, pp.240-248, 2008.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計1件)

青山俊弘, 山地一禎, 池田大輔, 行木孝夫: 機関リポジトリコンテンツの多面的な学内利用フレームワークの提案と実装, 情報知識学会, 査読有, Vol.23, No.3, pp. 380-394 (2013).

[学会発表](計8件)

Yamaji, K., Aoyama, T.: Shared Supplemental Data Repository for Japanese Academic Societies in Information Sciences, 10th International Digital Curation Conference (2015).

Maeda, A., Kato, H., Takahashi, N., Yoshida, Y., Ushirosako, K., Yamaji, K.: Institutional Repository ecosystem in Japan, IRDB and JAIRO Cloud, 9th International Conference on Open Repositories (2014).

Yamaji, K., Kato, H., Aoyama, T., Yamada, T.: Handshake ecosystem for Educational Contents between Institutional Repository and OER

based Repository, 9th International Conference on Open Repositories (2014).

Aoyama, T., Suzuki, Y., Yamaji, K.: Tag Cloud of Search Queries for Repository System, 9th International Conference on Open Repositories (2014).

Shiraishi, T., Aoyama, T., Yamaji, K., Namiki, T., Ikeda, D.: Preliminary Results for Discovering Related Words from Logs of Scholarly Repositories, International Conference on Advanced Information Technologies (2013).

Aoyama, T., Suzuki, Y., Yamaji, K.: Automatic reproduce metadata from the log of HTTP server, 8th International Conference on Open Repositories (2013).

Yamaji, K., Aoyama, T., Bannai, S., Arai, N.: A mash-up of a Japanese Open Repository and a Researcher CV Platform, 8th International Conference on Open Repositories (2013).

Yamaji, K., Aoyama, T. and Takeda, H.: A Handshake System for Japanese Academic Societies and Institutional Repositories, The 7th International Conference on Open Repositories (2012).

[その他]

ホームページ等

<http://weko.at.nii.ac.jp/>

<http://forge.at.nii.ac.jp/svn/nc/repository/>

6. 研究組織

(1)研究代表者

山地 一禎 (YAMAJI, Kazutsuna)

国立情報学研究所・コンテンツ科学研究系・准教授

研究者番号 : 50373379

(2)研究協力者

行木 孝夫 (NAMIKI, Takao)

池田 大輔 (IKEDA, Daisuke)

青山 俊弘 (AOYAMA, Toshihiro)