

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 14 日現在

機関番号：10101

研究種目：基盤研究(B)

研究期間：2011～2014

課題番号：23300087

研究課題名(和文) 研究基盤としてのデジタルリポジトリのモデル化および実装

研究課題名(英文) Digital Repositories as Infrastructure of Research Community

研究代表者

行木 孝夫 (Namiki, Takao)

北海道大学・理学(系)研究科(研究院)・准教授

研究者番号：40271712

交付決定額(研究期間全体)：(直接経費) 11,000,000円

研究成果の概要(和文)：本研究課題においてはデジタルリポジトリコンテンツを研究者コミュニティが柔軟かつ高度に利用するためのモデルを提案し実装した。特に、リポジトリプラットフォームWEKOを拡張したSarabiWEKOにおいて、リポジトリ管理と外部付与メタデータ管理とを分離した上で統合的に扱うモデルを確立し、価値付与サービスとしてユースケースを提案した。同時にリポジトリへのアクセスログからアクセスキーワードを抽出し、外部付与メタデータとするモデルを提案した。これらはいずれも小規模な研究グループにとってデジタルリポジトリが研究基盤たりうることを示している。

研究成果の概要(英文)：Our research group has established a model of value-add service to use contents of digital repositories. Especially, we have implemented the service on SarabiWEKO, which is an extension of a repository software WEKO. As the feature of SarabiWEKO, the users can add external metadata by themselves and can integrate the basic metadata and external metadata. By the user model contents of digital repositories could be added disciplinary metadata by certain research community as some external metadata. We also proposed two use-cases of SarabiWEKO, as a showcase of subject based research community and as a showcase of a research laboratory. On the other hand, we also propose a metadata model in which one can add the search keywords from referrer fields of access logfile of digital repository as external metadata. These models show that digital repositories can be essential infrastructure for small-scale research communities.

研究分野：図書館情報学、複雑系

キーワード：デジタルリポジトリ 機関リポジトリ サブジェクトリポジトリ 価値付与サービス スモールサイエンス

1. 研究開始当初の背景

研究開始当初の2011年には、国立情報学研究所の主導したCyber Science Infrastructure (CSI) プロジェクト第一期(2005-2007)および第二期(2008-2009)による支援のもと、日本では国立大学法人を中心に機関リポジトリが広く普及し、第二期CSIプロジェクトの終了時には175機関の機関リポジトリが構築済みであり、874,587コンテンツが搭載された。現在は第三期(2010-2012)とJAIRO Cloudの提供開始を受けて更に281機関、1,526,978コンテンツに増加している(参考文献1,3,5)。また、経済学および理論物理学・数学を中心とする特定の研究分野においては既にRePEC, arXivを始めとするサブジェクトリポジトリが広く受け入れられていた。

これらサブジェクトリポジトリおよび機関リポジトリに代表されるデジタルリポジトリの特徴は、コンテンツの基本的なメタデータが整備されていること、そのメタデータを系統的にOAI-PMHによって提供することである。機関リポジトリおよびサブジェクトリポジトリに代表されるデジタルリポジトリの発展にも関わらず、これらの機能が活用されているとは言い難い状況にあった。

同時に、機関リポジトリへ研究者の理解を深めるためにも機関リポジトリコンテンツを高度に活用する何らかの手段を必要とした。

海外においては既にコンテンツの機械登録プロトコルとしてSWORDが提案され、アグリゲートの手段としてOAI-OREが提案されており、特にSWORDは広く活用されつつあった。

このように、デジタルリポジトリと単なるwebコンテンツを区別するものとしてリポジトリコンテンツを高度に活用する手法が必要である。リポジトリ構築側と研究者コミュニティの間は隔絶しており、研究者コミュニティがリポジトリコンテンツを柔軟に活用することは難しい。

研究データを機関リポジトリへ格納する方向性は研究開始当初から世界的な潮流ではあったものの、それらはコミュニティに帰するものとして扱い、機関リポジトリの対象としてはOpenRepositories2008の基調講演でも見られたスモールサイエンスを目指すものとした。小規模なコミュニティにおいて研究基盤として機関リポジトリおよびサブジェクトリポジトリを置くものである。

本課題ではこの問題に取り組み、リポジトリコンテンツを柔軟に活用し、コンテンツへの価値付与サービスのモデル構築に集中した。

2. 研究の目的

多様な研究分野、多様なユーザに適用できるデジタルリポジトリ相互運用の新たなモデルを提案することを目的とした。研究グル

ープの規模に応じて適応する研究基盤は異なると考えられるが、本研究では一定の規模スケールな研究グループを対象とし、メンバーが広く国内外の研究機関に分散しているケースを対象とする。このようなケースは決して少なくなく、多数派である。同時に研究サイクルにおける情報発信と情報流通に関して重点を置き、研究期間内に、特定の研究分野を例示しつつ必要な情報流通発信モデルを設定し、モデルに基づいたデジタルリポジトリを構築する。特に、コンテンツをリポジトリおよびwebサイト間において共有するモデルに重点を置く。実装においては、他の実装にも役立つよう、データモデルと操作体系を具体的に示す。本研究では単なるシステム構築に留まらず、ユーザモデルを提示し、既存のデジタルリポジトリとの違いを明確にする。提示するユーザモデルは、利用形態に即した具体的なデータモデルとその操作体系、およびリポジトリ間インターフェースであり、これを基に様々な実装も期待できる。

3. 研究の方法

研究代表者(行木)が本課題以前から構築を続けてきたリポジトリオーバーレイ型のサブジェクトポータルをモデルケースとし、分担者(山地)の開発するリポジトリプラットフォームWEKOを想定して価値付与サービスのあり方をクローズドなワークショップ形式で検討した。システム設計およびログ解析の一部には分担者(池田)が、ログ解析による価値付与サービスおよびアノテーション、多様なユースケースの展開に分担者(青山)が寄与した。

4. 研究成果

(1) SarabiWEKOと研究者コミュニティによる外部メタデータ付与、価値付与サービスとショウケースとしての機関リポジトリ。

機関リポジトリコンテンツに関わる価値付与サービスおよび研究者コミュニティとの連携について、主な発表論文3,5,6を中心に成果を述べる。

機関リポジトリ、サブジェクトリポジトリを問わず、リポジトリコンテンツには書誌情報をはじめとするメタデータを付与し公開される。このメタデータは機関リポジトリ運営者によって定められた属性に対応する属性値である。セルフアーカイビングによって機関リポジトリにコンテンツを登録するとともに、著者による解説や関連するファイルの提供、専門分野コードなどのコンテンツに対する有益の情報を公開したい場合でも、属性を自由に拡張することは難しいことを意味する。機関に所属するさまざまな分野の研究者の要求に応じた属性を追加することで、研究者によるコンテンツと価値のあるメタデータを公開することを支援できると考え

られる。そこで本研究では、機関リポジトリがコンテンツの書誌情報からなる基本的なメタデータを保持し、専門的あるいは補足的な情報を機関リポジトリ以外で付加し、公開するフレームワークを提案する。さらに、提案するフレームワークを実装したシステムとして SarabiWEKO を開発した。SarabiWEKO によってリポジトリシステム WEKO に対して、リポジトリ上のコンテンツのメタデータに付与する形でユーザが任意のメタデータを追加し、公開することが可能となる。

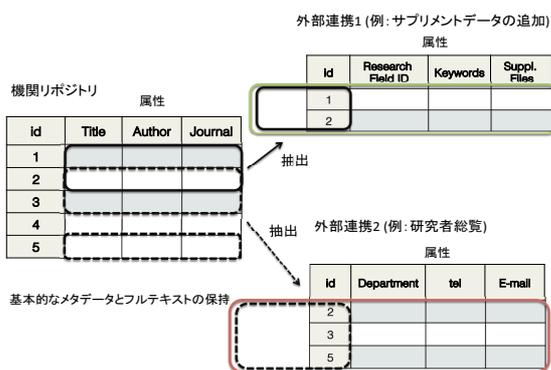


図 1 機関リポジトリのメタデータをリポジトリ外のメタデータにより拡張し、再利用する。

SarabiWEKO においては図 1 に示すようにリポジトリの管理者からコミュニティへ拡張メタデータの管理権限を委譲することができる。これによって、多様なユースケースへの対応が可能となった。提案したユースケースとして、学科および研究室等の単位における利用例とサブジェクトポータルにおける利用例を示した。

(2) 数学分野におけるリポジトリオーバーレイ、フルテキストとマイクロコンテンツ

研究代表者 (行木) が継続している DML-JP に関わる研究成果を主な発表論文 3, 4, 7, 8 と学会発表 8, 13 を中心に引用しつつ述べる。

機関リポジトリおよびデジタルリポジトリを中心とするコンテンツからサブジェクトベースのポータルを構築することを目的とし、数学分野におけるコンテンツ形成を主題として取り上げた。本事業で構築したポータルサイトを WDM (World Digital Mathematics Library) 計画のカウンターパートという意味で DML-JP と称する。

<http://sparc1.math.sci.hokudai.ac.jp/dmljp/>

WDM とは、2005 年に IMU (International Mathematics Union) の下部委員会である CEIC (Committee for Electronic Information and Communication) の提案した計画である。各国の数学コミュニティが保

持するコミュニティベースの学術誌を電子化、公開し、可能なすべての数学関連文献は電子的に可読とする計画である。

日本においては後述する紀要文化に応じて発展した数学系学術誌の多様性から、中央集約的な実現が困難であった。一部の先駆的なジャーナルの経験をもとに数誌が電子化を完了したのみである。それでも主要な数学系国際誌の電子化は国立情報学研究所の SPARC Japan 事業のもとで大きく前進した。一方で、機関リポジトリに大学紀要が搭載されることで、中小規模の紀要をもとにする数学系学術誌も多く電子化され、オーバーレイ型の DML を構築する基盤が整った。

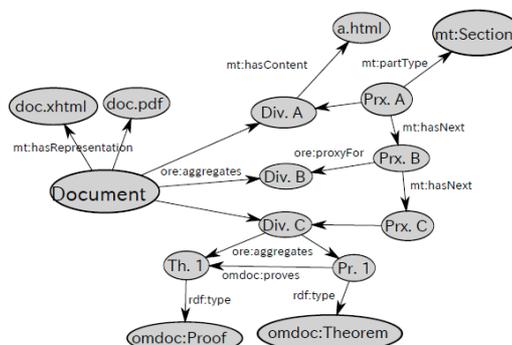
これらの背景のもとに、日本の数学系学術誌をオーバーレイするポータルとして DML-JP を構築している。

本成果では、OAI-PMH ベースのメタデータハーベスティングに関わる問題を解決し、SarabiWEKO とは異なるアプローチをとった。

まず、OAI-PMH のキャッシュサーバーを構築するツールを開発し、OAI-PMH の GetRecord, ListRecords によるアクセスの軽減を図った。これにより、小規模なコンテンツの価値付与サービスにおいては GetRecord もしくは ListRecords と Set 指定によるオンデマンドなメタデータ取得と外部メタデータのマッシュアップサービスを簡便に構築することができるようになった。

数学分野におけるフルテキストはマークアップ言語による表現に適しており、NLM-DTD+MathML および OpenMath による表現は数式の機能を込めた表現が可能である。

また、数学分野の論文における主要な成果は定理の形で提示される。これを適切にマークアップすることでマイクロコンテンツとして再利用可能な形に表現することができる。これらの関係は下図のように表現でき、ある定理が別の定理を参照して証明される場合、後者の証明を前者の証明の内部へ組み込むインターフェースを構築した。



ユーザ支援環境として wiki ベースの環境を構築し、数学分野のフルテキストとして必要な語彙を与えることで参照先のマイクロコンテンツを参照元に埋め込んでいる。具体的なマークアップ例を次に示す。

```

<statement content-type="theorem">
<title>Taylor's Theorem </title>
<p>Let<inline-formula><math
xmlns="http://www.w3.org/1998/Math/MathML"
display="inline"><mi>
* snip *
</math>
</p>
<uri xmlns:xlink="http://www.w3.org/1999/
xlink
"xlink:href="http://mw2011.matherial.org/
LaurentSeries/en/pr/TaylorTheorem"/>
</statement>

```

(3) ログ解析と価値付与サービス

研究分担者（青山）の関わるログ解析と価値付与サービスに関し、学会発表 7 を中心に記述する。

佐藤翔らの研究（参考文献 4）によれば、機関リポジトリへのアクセスの多くは検索エンジンを経由している。検索キーワードはメタデータに限らずフルテキスト中のテキスト情報へも誘導されており、検索キーワードがコンテンツへの有効な情報となることが期待される。

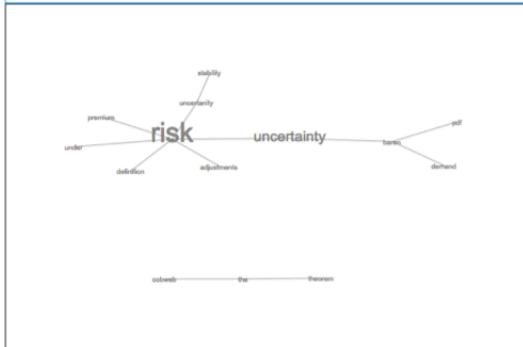
機関リポジトリへのアクセスは http サーバのログに記録されており、リポジトリ構築ソフトウェアによっては独自にアクセスを記録している。ここでは httpd 特に apache のログファイルに着目し、ログファイルの referrer 部分に記録される検索キーワードを抽出した上でコンテンツ自身のメタデータとして付与するシステムを構築した。

139

Risk premium and the stability of some adjustment processes under uncertainty

Uzawa, Masaru

risk,premium,stability,adjustment,uncertainty,331,経済学



対象とした機関リポジトリは小樽商科大学の機関リポジトリ Barrel および DML-JP である。

著者の指定したキーワードと referrer から取得した検索キーワードとを比較することで、何らかの知見を得ることを期待した。検索キーワードの 60%はメタデータと一致していない。メタデータへ検索キーワードを付与することで、コンテンツのユーザから見た有効なメタデータを付与することが可能になると考えられる。図は検索キーワードの共起情報から構成したタグクラウドである。

参考文献（研究成果論文を除く）

1. 国立情報学研究所「未来への飛躍 ～機関リポジトリの更なる発展を目指して～:学術機関リポジトリ構築連携支援事業 第3期報告書」2014年3月
2. 国立情報学研究所 CSI プロジェクト, 学術機関リポジトリ構築連携支援事業報告書, 領域2 数学ポータル構築, 2013年3月
3. 国立情報学研究所「変容する学術情報流通、進展する機関リポジトリ：学術機関リポジトリ構築連携支援事業 第2期報告書」2011年12月
4. 機関リポジトリへの登録が論文の被引用数と電子ジャーナルアクセス数に与える影響, 佐藤 翔ほか, 情報知識学会誌 Vol. 21 No. 3, pp. 383-402, 2011
5. 国立情報学研究所「学術コミュニケーションの新たな地平:学術機関リポジトリ構築連携支援事業第1期報告書」2008年12月
6. Takao Namiki and Hiraku Kuroda, "Digital Mathematics Library over digital repositories" (in Japanese), Digital Libraries, No. 37 pp. 18-23, (Nov. 2009)
7. Takao Namiki, "Current status of mathematical publications in Japan", In Proceedings of Towards Digital Mathematics Library 2008 pp. 97-102 (27 Jul 2008, Birmingham, UK)

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 8 件, 全て査読有)

- ① Takehiro Shiraishi, Toshihiro Aoyama, Kazutsuna Yamaji, Takao Namiki and Daisuke Ikeda: "Discovering Unpredictably Related Words from Logs of Scholarly Repositories for Grouping Similar Queries", Roger Y. Lee (Ed): Applied Computing and

- Information Technology, Studies Computational Intelligence Vol.553, DOI: 10.1007/978-3-319-05717-0_4, pp.48--60, May 2014. (査読有)
- ② 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」 情報知識学会誌 23 巻 3 号 pp.380-394 (2013), 情報知識学会 2014 年度論文賞受賞, open access, doi: 10.2964/jsik.23_380 (査読有)
- ③ SarabiWEKO: Metadata class management system for WEKO repository Toshihiro Aoyama, Kazutsuna Yamaji, Daisuke Ikeda, Takao Namiki Proceedings of 2nd International Conference on Academic Library, pp. 204-20 (Feb. 2013) (査読有)
- ④ Toshihiro AOYAMA, Kazutsuna YAMAJI, Daisuke IKEDA and Takao NAMIKI, “SarabiWEKO: A reuse system of WEKO contents and metadata”, Digital Libraries, No. 41, pp. 3-10 (Nov. 2011) (査読有)
- ⑤ Hiraku Kuroda and Takao Namiki, “Wiki Authoring and Semantics of Mathematiccal Document Structure”, In: Proceedings of the ITP 2011 Workshop on Mathematical Wikis pp.28-37. (MathWikis2011, Eds: Christoph Lange and Josef Urban, Nijmegen, Netherlands, August 27, 2011) (査読有)
- ⑥ Takao Namiki, Kazutsuna Yamaji, Toshiyuki Kataoka and Noboru Sonehara, “Time Stamping Preprint and Electronic Journal Server Environment”, In proceedings of DML2011 Towards a Digital Mathematics Library, pp 19-23, Eds: Petr Sojka and Thierry Bouche (July 20-21, 2011 Bertinoro, Italy) (査読有)
- [学会発表] (計 17 件)
- ① Tag Cloud of Search Queries for Repository System, Toshihiro Aoyama, Yuta Suzuki, Kazutsuna Yamaji, Open Repositories 2014, June 9-13 2014, Helsinki (Finland)
- ② Handshake ecosystem for Educational Contents between Institutional Repository and OER based Repository, Kazutsuna Yamaji, Hiroshi Kato, Toshihiro Aoyama, Tsuneo Yamada, Open Repositories 2014, June 9-13 2014, Helsinki (Finland)
- ③ Institutional Repository ecosystem in Japan, IRDB and JAIRO Cloud, Akira Maeda, Hiroshi Kato, Nanako Takahashi, Yukinae Yoshida, Kumi Ushirosako, Kazutsuna Yamaji, Open Repositories 2014, June 9-13 2014, Helsinki (Finland)
- ④ 行木孝夫, 青山俊弘, 池田大輔, 山地一禎 「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 第 22 回 (2014 年度) 年次大会 2014 年 5 月 25 日, 和歌山大学 (和歌山県和歌山市), 2014 年度情報知識学会論文賞受賞招待講演
- ⑤ A mash-up of a Japanese Open Repository and a Researcher CV Platform, Kazu Yamaji, Toshihiro Aoyama, Satoru Bannai, Noriko Arai, International Conference on Open Repositories 2013, July 8-12 2013, Charlottetown, PEI (Canada)
- ⑥ Automatic reproduce metadata from the log of HTTP server, Toshihiro Aoyama, Yuta Suzuki, Kazutsuna Yamaji, International Conference on Open Repositories 2013, July 8-12 2013, Charlottetown, PEI (Canada)
- ⑦ Client based interface and proxy server for content re-use framework based on OAI-PMH, Takao Namiki, International Conference on Open Repositories 2013, July 8-12 2013, Charlottetown, PEI (Canada)
- ⑧ A Handshake System for Japanese Academic Societies and Institutional Repositories, Kazutsuna Yamaji, Toshihiro Aoyama, Hideaki Takeda, The 7th International Conference on Open Repositories, July 9-13 2012, Edinburgh (United Kingdom)
- ⑨ JAIRO Cloud: National Infrastructure for Institutional Repositories in JAPAN, Shiozaki, R., Tanabe, M., Mori, I. and Yamaji, K., The 7th International Conference on Open Repositories, July 9-13 2012, Edinburgh (United Kingdom)
- ⑩ Takao Namiki, “Client based interfaces for conteten re-use framework based on OAI-PMH”, The 7th International Conference on Open Repositories, July 9-13 2012, Edinburgh (United Kingdom)
- ⑪ Takao Namiki and Daisuke Ikeda “Digital Mathematics Library in Japan as an overlay of digital repository systems” JCDL-DR 2011, June 2011, Ottawa (Canada)
- ⑫ Takao Namiki, Hiraku Kuroda, “Digital Mathematics Library in Japan: collaboration between disciplinary community and digital repository community”, Open Repositories 2011, June 6-11, 2011,

- Austin, Texas (USA)
- ⑬ Kazu Yamaji, Toshihiro Aoyama,
Daisuke Ikeda, Takao Namiki, “New
Function of the Institutional
Repository as Departmental Show
Cases”, Open Repositories 2011, June
6-11, 2011, Austin, Texas (USA)

[その他]

ホームページ等

DML-JP

sparc1.math.sci.hokudai.ac.jp/dmljp/

6. 研究組織

(1) 研究代表者

行木 孝夫 (NAMIKI, Takao)

北海道大学・大学院理学研究院・准教授

研究者番号：40271712

(2) 研究分担者

青山 俊弘 (AOYAMA, Toshihiro)

鈴鹿工業高等専門学校・電子情報工学科・

准教授

研究者番号：00294992

池田 大輔 (IKEDA, Daisuke)

九州大学・大学院システム情報科学研究

院・准教授

研究者番号：00373259

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