

学術資源リポジトリにおけるLightweight Information Describing Object(LIDO)の検討

著者	林 正治, 堀井 洋, 堀井 美里, 高田 良宏, 山地一禎, 上田 啓未, 古畑 徹
雑誌名	情報知識学会誌 = Journal of Japan Society of Information and Knowledge
巻	23
号	2
ページ	292-297
発行年	2013-01-01
URL	http://hdl.handle.net/2297/35259

doi: 10.2964/jsik.23_292

第21回年次大会予稿

学術資源リポジトリにおけるLightweight Information Describing Object(LIDO)の検討

A study of Lightweight Information Describing Object (LIDO) in Academic Resource Repository

林正治^{1*}, 堀井洋², 堀井美里², 高田良宏³, 山地一禎⁴, 上田啓未⁵, 古畑徹⁵
Masaharu HAYASHI^{1*}, Hiroshi HORII², Misato HORII², Yoshihiro TAKATA³,
Kazutsuna YAMAJI⁴, Hiromi UEDA⁵, Toru FURUHATA⁵

1* 一橋大学 情報化統括本部情報基盤センター

Center for Information and Communication Technology, Hitotsubashi University

2 合同会社AMANE

AMANE LLC.

3 金沢大学総合メディア基盤センター

Information Media Center, Kanazawa University

4 国立情報学研究所 学術ネットワーク研究開発センター

R & D Center for Academic Networks, National Institute of Informatics

5 金沢大学資料館

Kanazawa University Museum

文献資料およびモノ資料を含む学術資料の共有化を目的とした学術資源リポジトリ構想がある。これまでにDublin Core Metadata Element Setを拡張し、学術資源情報の記述を実現してきたが、ハーベスティング時の情報欠落に課題があった。一方、国外では博物館資料を対象とした国際的なハーベスティング・スキーマLIDOへの注目が高まっている。本稿では、学術資源リポジトリにおけるLIDOの可能性を検討する。

We are aiming to share bibliographic materials and non-bibliographic materials by using a qualified Dublin Core metadata on repository system named Academic Resource Repository. However, there is an information lack problem in OAI metadata harvesting via oai_dc. On the other hand, it is growing interest in LIDO (Lightweight Information Describing Object) that is an international harvesting schema for museum object recently. In this paper, we study about the potentiality of LIDO in Academic Resource Repository.

キーワード: 学術情報リポジトリ, OAI-PMH, LIDO, 非文献コンテンツ, Dublin Core

Academic resource repository, OAI-PMH, LIDO, non-Bibliographic contents, Dublin Core

1 はじめに

文献資料および文献資料以外の資料（文献資料と対比して非文献資料）を含む学術資料の共有化を目的とした学術資源リポジトリ構想がある[1-2]。大学等の高等教育機関が構築・運用する学術機関リポジトリを拡張し、非文献資料の収集、組織化、保存・発信のための仕組みの実現を目指している。

著者らのグループはこれまでにDublin Core Metadata Element Set (DCMES) [3]をベースにした共通メタデータ形式とリポジトリシステムWEKOへの実装を行ってきた [2]。共通メタデータ形式では、新たなメタデータ要素は追加せず、限定子を用いてDCMESを拡張し、非文献資料情報の記述を可能にした。Dumb-Down原則[4]により、限定子を取り除かれても、メタデータ要素と値との間で矛盾が生じることはなく、DCMESとの互換性も確保した。

共通メタデータ形式をDumb-Downすることで、学術機関リポジトリ連携を想定したoai_dcによるハーベスティング[5]にも対応できる。

しかしながら、共通メタデータ形式およびoai_dcによるハーベスティングには二つの問題がある。一つ目は共通メタデータ形式を用いた場合の表現の制約である。限定子によるDCMESの拡張であるため、DCMESの15要素の詳細化は可能としても、Dumb-Down原則に反するため、15要素を超えた表現はできない。二つ目は、ハーベスティング時はDumb-Downにより15要素によるメタデータ表現となる点である。これは学術機関リポジトリと連携する場合はメリットとなりうるが、非文献資料を対象とする学術資源リポジトリとしては大きな

制約となる。oai_dcを拡張し、共通メタデータ形式に対応させることも考えられるが、一つ目の問題、DCMES15要素以上の表現は困難であるという問題は残る。

他方、収蔵品の殆どが非文献資料となる博物館業界にはLIDO (Lightweight Information Describing Objects) [6]と呼ばれるハーベスティング・スキーマがある。本来は収蔵品管理システム上の情報収集を想定したハーベスティング・スキーマであるが、学術資源リポジトリにとってはoai_dcの代替手段となりうる。

本報告では学術資源リポジトリにおけるLIDOの適用可能性について検証する。

2 LIDO (Lightweight Information Describing Object)

LIDOは博物館資料情報のための国際的なハーベスティング・スキーマである。国際博物館会議 (ICOM) が組織するドキュメンテーション委員会 (CIDOC) のワーキンググループ「Data Harvesting and Interchange」により仕様が策定された。最新バージョンは2010年11月に公開されたバージョン1.0である [6]。

LIDOは博物館等の収蔵品管理システムに蓄積された情報を収集し、ポータルサイトなどで利用されることを想定している。LIDOの記述対象は幅広く、芸術・文化・技術・自然科学に関するあらゆる資料情報を表現できるとされる。また、CIDOC CRM (ISO 21127) に準拠したイベント指向の記述が可能という特徴をもつ。

LIDOは博物館等の収蔵品をLIDOレコードとよばれる記述単位で管理する。LIDOレコードは5領域・14情報グループに分類されたメタデータ要素で構成される。表1は

LIDOレコードの概要となる5領域・14情報グループの概要である。

表1の中に下線が引かれた項目が3つあるが、これらはLIDOレコードで必須となる情報グループである。博物館資料を識別するために最低限必要な情報を定義している。

LIDOレコードはXML文書として記述される。ルート要素はlidoWrapであり、その子要素lidoにLIDOレコードを記述する。LIDOレコードは構造化文書であるため、DCMESのように要素とコンテンツが1対1となることは少ない。階層化された要素とコンテンツが対応する形となる。

表 1 領域・情報グループ

分類情報領域 (Object Classifications)
<u>資料作品タイプ (Object / Work Type)</u>
資料分類 (Classification)
基本情報領域 (Object Identification)
<u>資料名称 (Title / Name)</u>
記述・記録 (Inscriptions)
所蔵機関・場所 (Repository / Location)
状態・版 (State / Edition)
解説 (Object Description)
測定 (Measurements)
イベント情報 (Event)
Event Set
関連情報領域 (Relations)
主題 (Subject Set)
関連作品 (Related Works)
管理メタデータ領域 (Administrative Metadata)
権利 (Rights)
<u>レコード情報 (Record)</u>
関連デジタル資料 (Resource)

3 LIDOへのマッピング

学術資源リポジトリへの適用可能性を検証するために共通メタデータ形式の要素とLIDOスキーマの要素とを対応付ける。各要素同士を対応付けることで、学術資源リポジトリにおけるLIDOスキーマの表現力を確かめることができる。また、既存の学術資源リポジトリのコンテンツをLIDO対応とする際にも役立てることができる。

共通メタデータ形式とLIDOスキーマとのマッピングは以下のルールで実施する。

1. 共通メタデータ形式とLIDOスキーマのマッピングは1対1マッピングを基本とする。
2. 共通メタデータ形式の要素を詳細化することでLIDOスキーマの要素とマッピング可能な場合はマッピング候補とする。
3. マッピングにより、共通メタデータ形式の要素がもつ意味が曖昧になる場合はそれを区別する。

3.1 title要素

共通メタデータ形式のtitle要素はDCMESで定義されているものと同じである。非文献資料のタイトル記述に用いる。

LIDOでは、資料名称グループで定義されているメタデータ要素でタイトルを記述する(表2参照)。これらはLIDOレコードを記述する上で必須要素である。要素titleはLIDOスキーマの要素と1対1で対応している。

表 2 titleのマッピング

共通メタデータ形式	LIDO
title	lidoWrap→lido→ descriptiveMetadata→ objectIdentificationWrap→ titleWrap→titleSet→ appellationValue

3.2 coverage.temporal要素

coverage.temporal要素はDCMI Metadata Terms[7]にて定義されている。時代についての情報を記述するためのメタデータ要素である。LIDOでは、このような資料に関する時間情報・場所情報をイベント情報として記述する(表3参照)。

coverage.temporal要素の場合、マッピ

ング方法は2通りある。表示用のdisplayEvent要素と索引用のevent要素による表現である。表示用のdisplayEvent要素は文字列そのままを内容として記述する。一方、event要素は場所・時間・人などイベントに関連する情報の詳細な記述が可能である。共通メタデータ形式の要素を詳細化することで、より詳細な表現が可能となる。

表 3 coverage.temporalのマッピング

共通メタデータ形式	LIDO
coverage.temporal	lidoWrap→lido→descriptiveMetadata→eventWrap→eventSet→displayEvent
	lidoWrap→lido→descriptiveMetadata→eventWrap→eventSet→event

3.2 description.history要素

description.history要素は共通メタデータ形式オリジナルのメタデータ要素である。資料の由来についての情報を記述する。この要素については1対1でマッピング可能な対象がない。意味づけが曖昧となるがdescriptiveNote要素とマッピングする。

表 4description.historyのマッピング

共通メタデータ形式	LIDO
description.history	lidoWrap→lido→descriptiveMetadata→objectIdentificationWrap→objectDescriptionWrap→objectDescriptionSet→descriptiveNote

3.3 その他の要素とのマッピング

共通メタデータ形式とLIDOスキーマとのマッピング結果を表5に示す。マッピング先の構造は一部省略している。表中の「@」で始まる文字は属性、「,」は同時に定義必要な属性、「/」はマッピング候補が

複数あることを表す。「△」はマッピングにより意味が曖昧となる場合、「○」は1対1マッピングが可能である場合、「◎」は共通メタデータ要素の詳細化により精度の高い情報を表現可能であることを示す。

4 考察

共通メタデータ形式のマッピング結果から、学術資源リポジトリによけるLIDOの適用可能性を考察する。

4.1 LIDOの表現力について

共通メタデータ形式とのマッピング過程から以下のことが明らかとなった。

1. 共通メタデータ形式で表現された非文献資料情報はLIDOでも表現できる。
2. イベント情報など一部の情報については、共通メタデータ形式より詳細な表現が可能である。
3. カテゴリなどの分類情報はLIDOでは並列で扱われるため、シソーラスなどによる語彙管理が必要である。

以上のことから、非文献資料に対するLIDOの表現力は十分であることがわかった。ただし、LIDOの構造は階層が深くなりやすいため、情報システムに搭載するにはユーザインタフェースをよく考える必要がある。

4.2 LIDOと共通メタデータ形式の互換性について

LIDOに備わっている表示用の要素displayXXXXにより、構造化されていない共通メタデータ形式との互換性を維持できることが確認できた。また、共通メタデータ形式とのマッピングにより、DCMESと

のマッピングも可能であることを示すことができた。これにより、文献資料用のハーベスティング・スキーマoai_dcとの共存にも期待がもてる。

5 おわりに

本稿では学術資源リポジトリにおけるLIDOスキーマの適用可能性を検討するために、学術資源リポジトリの共通メタデータ形式とのマッピングを実施した。マッピングについては、実際のデータと照らし合わせながらの検証が必要と考えるが、今回の検証結果から共通メタデータ形式との互換性とLIDOスキーマの表現力の高さを確認することができた。今後は本報告の成果をもとに、リポジトリシステムへの実装を進めていきたい。

謝辞

本研究は、科研費(基盤研究(B)24300310)の助成を受けたものである。関係各位に感謝する。

参考文献

- [1] 高田良宏, 笠原禎也, 西澤滋人, 森雅秀, 内島秀樹: 「非文献コンテンツのための可視性と保守性に優れた学術情報リポジトリの構築」, 情報知識学会, Vol. 19, No. 3, pp. 251-263, 2009.
- [2] 堀井洋, 堀井美里, 林正治, 塩瀬隆之, 高田良宏, 古畑徹: 「分野・組織横断的な

非文献資料リポジトリの実現を目指して」, 情報知識学会, Vol. 22, No. 2, pp. 91-96, 2012.

[3] Dublin Core® Metadata Initiative: "Dublin Core Metadata Element Set, Version 1.1",

<http://dublincore.org/documents/dces/> (accessed 10 Apr 2013)

[4] 杉本 重雄: 「メタデータについて: Dublin Core を中心として」, 情報知識学会誌, Vol. 10, No. 3, pp. 53-58(2000).

[5] Open Archives Initiative: "Open Archives Initiative Protocol for Metadata Harvesting",

<http://www.openarchives.org/pmh/> (accessed 10 Apr. 2013)

[6] ICOM-CIDOC Working Group Data Harvesting and Interchange: "LIDO - Lightweight Information Describing Object Version 1.0",

<http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf> (accessed 10 Apr. 2013)

[7] Dublin Core® Metadata Initiative: "DCMI Metadata Terms",

<http://dublincore.org/documents/dcmi-terms/> (accessed 10 Apr. 2013)

表 5 共通メタデータ形式とLIDOスキーマのマッピング結果

共通メタデータ要素名	共通メタデータ要素説明	LIDO	
title	タイトル	titleSet→appellationValue	○
title.alternative	タイトル(英文)	titleSet→appellationValue, @xml:lang=en	○
title.transcription	タイトル読み	titleSet→appellationValue	○
title.collection	統一資料名(大分類)	category→term	△

title.collection-transcription	統一資料名（大分類）読み	category→term	△
title.union	統一資料名（中分類）	category→term	△
title.union-transcription	統一資料名（中分類）読み	category→term	△
title.community	統一資料名（小分類）	category→term	△
title.community-transcription	統一資料名（小分類）読み	category→term	△
subject	トピック	subjectSet→displaySubject/subjectSet→subject→subjectConcept→term	◎
subject.application	用途	subjectSet→subject→subjectConcept→term	○
subject.domain	領域	subjectSet→subject→subjectConcept→term	○
language	言語	@xml:lang=ja, @pref=preferred	○
source	管理番号・請求番号	recordID	○
creator	電子化資料の作成団体	recordSource→legalBodyName→appellationValue	○
publisher	公開者	recordSource→legalBodyName→appellationValue	○
coverage	地空間範囲	eventSet→displayEvent	△
coverage.temporal	時代についての情報	eventDate→displayDate/eventDate→date	◎
coverage.spatial	場所についての情報	eventPlace→displayPlace/eventPlace→place	◎
description	資料解説	objectDescriptionSet→descriptiveNoteValue	○
description.author	作成者	eventActor→displayActorInRole/actor→nameActorSet→appellationValue	◎
description.author-transcription	作成者読み	actor→nameActorSet→appellationValue	○
description.author-alternative	作成者（英文）	actor→nameActorSet→appellationValue	○
description.author-description	作成者についての解説	objectDescriptionSet→descriptiveNoteValue	△
description.publisher	出版者	eventActor→displayActorInRole/actor→nameActorSet→appellationValue	◎
description.publisher-transcription	出版社読み	actor→nameActorSet→appellationValue	○
description.copyist	複写・模写作成者	eventActor→displayActorInRole/actor→nameActorSet→appellationValue	◎
description.copyist-transcription	複写・模写作成者読み	actor→nameActorSet→appellationValue	○
description.copyist-alternative	複写・模写作成者読み	actor→nameActorSet→appellationValue	○
description.copy-date	複写年代についての情報	eventDate→displayDate/eventDate→date	◎
description.collection	統一資料（大分類）解説	objectDescriptionSet→descriptiveNoteValue	△
description.union	統一資料（中分類）解説	objectDescriptionSet→descriptiveNoteValue	△
description.community	統一資料（小分類）解説	objectDescriptionSet→descriptiveNoteValue	△
description.note	注記	objectDescriptionSet→descriptiveNoteValue	○
description.source-format	形態	eventMaterialsTech→displayMaterialsTech/eventMaterialsTech→materialsTech	◎
description.history	由来	objectDescriptionSet→descriptiveNoteValue	△
description.physics	数量データ	objectMeasurementsSet→displayObjectMeasurements/objectMeasurementsSet→objectMeasurements	◎
contributor	寄与者	recordRights→rightsHolder→legalBody→appellationValue	○
contributor.author	電子化資料寄与者	recordRights→rightsHolder→legalBody→appellationValue	○
contributor.photographer	電子化資料撮影者	resourceSource→legalBody→appellationValue	○
format	記録形式	resourceType→term	○
format.mimetype	電子化資料のMIMEタイプ	resourceType→term	○
rights	権利管理	rightsWorkWrap→creditLine	○
rights.rightsholder	所蔵	repositorySet→repositoryName→legalBodyName	○
rights.licence	使用上の注意	rightsWorkWrap→creditLine	○
date	日付	recordInfoSet→recordMetadataDate	○
date.modified	電子化資料変更日	recordInfoSet→recordMetadataDate	△
date.created	電子化資料作成日	recordInfoSet→recordMetadataDate	△
date.available	電子化資料公開日	recordInfoSet→recordMetadataDate	△
date.issued	発行日	recordInfoSet→recordMetadataDate	△
identifier	資源識別子	lidoRecID	○
identifier.URI	URI	recordInfoSet→recordInfoLink	○
relation	関係	relatedWorkSet→relatedWork→displayObject	○