

# OER のファインダビリティ向上のための システムの設計と実装

武田俊之<sup>\*1</sup>, 重田勝介<sup>\*2</sup>, 金子大輔<sup>\*3</sup>, 八木秀文<sup>\*4</sup>, 永嶋知紘<sup>\*5</sup>

\*1 関西学院大学, \*2 北海道大学, \*3 北星学園大学, \*4 東北大学, \*5 カーネギーメロン大学

## Design and implementation of a system to improve the findability of OER

Toshiyuki Takeda<sup>\*1</sup>, Katsusuke Shigeta<sup>\*2</sup>, Daisuke Kaneko<sup>\*3</sup>, Hidefumi Yagi<sup>\*4</sup>, Tomohiro  
Nagashima<sup>\*5</sup>

\*1 Kwansai Gakuin University, \*2 Hokkaido University, \*3 Hokusei Gakuen University,  
\*4 Tohoku University, \*5 Carnegie Mellon University

オンライン授業の普及と教材のデジタル化に伴って無料で二次利用可能なオープン教材（OER）が増加している。OER は授業や自習に有用であるが、ファインダビリティ（検索可能性）が低いなどの理由で再利用につながらないという課題がある。この報告では、まず OER のライフサイクル、ファインダビリティ、再利用、メタデータ標準について検討した上で、収集、前処理、メタデータ等のタグ付与、データベース化をおこなうシステムの設計と実装法について述べる。

キーワード: Open Educational Resources, OER, メタデータ, ファインダビリティ, 教材制作

### 1. はじめに

ユネスコ総会で採択された「オープン教育資源（OER）に関する勧告」[1]において、オープン教育資源（またはオープン教材、Open Educational Resources: OER）は以下のように定義されている。

1 オープン教育資源（OER）とは、パブリック・ドメインとなった、又はオープンライセンスの下で公開されている著作権のあるあらゆる形式及び媒体の学習、教育及び研究の資料であって、他の者による無料のアクセス、再使用、別の目的のための再利用、改訂及び再配布を認めるものをいう。

2 オープンライセンスとは、著作権者の知的所有権を尊重し、並びに公衆に教材を再使用し、別の目的のために再利用し、改訂し、及び再配布し、並びに教材にアクセスする権利を付与する許可を与え

るライセンスをいう。

さらに、OER は ICT を通じて「個々の学習者のニーズを満たし、及びジェンダー平等を効果的に促進することを助長し、並びに革新的な教育学的、教訓的、及び方法論的な取組を奨励し得る。」とされている。

MIT の OpenCourseWare（OCW）が活性化させた教材のオープン化は、OER を利用したオープン教育運動へと発展しており[2]、対象や社会的なインパクトを広げている。OCW だけではなく、MOOC（Massive Open Online Course）その他でのコンテンツ公開だけではなく、大学教員以外による YouTube 等での教材公開も増加している。さらに COVID-19 により授業のオンライン化を余儀なくされた中で、利用機会が増えた OER の価値が高まっているといつてよいであろう。

一方で OER の利用においては課題もある。その一

つがファインダビリティ（検索可能性）である[3, 4, 5]。OER は年々増え続けている。一方、OER はコースを担当する教員によって作成されたものが多く、コンテンツの内容、難易度、構成、構造、メディアなどがさまざまであり、自分の授業で OER を取り入れる場合の選択が困難である。海外では Merlot、ARIADNE、OER Commons などの OER（教材そのものやメタデータ）集積と検索を提供するサイトがあるが、現在の日本では集積サイトも検索サービスも存在しない。

この報告では、日本語の OER のファインダビリティを高めるようなシステムの設計と試作について述べる。設計にあたっては、OER のライフサイクル、ファインダビリティ、再利用、メタデータ標準について検討した上で、収集、前処理、メタデータ等のタグ付与、データベース化について検討をおこなう。なお、教材を指す用語として、e ラーニングの文脈では LO（Learning Object）、LOM（Learning Object Metadata）の用語が用いられてきたが、本稿ではより広義の教材を指す Educational Resource(s)を主に用いることにする。

## 2. システムの概要

OER のファインダビリティは OER 普及にあたって常に大きな課題となっている。図書の場合は分類、キーワードの付与のプロセスが既に確立していて、各図書ごとの利用数も大きい。OER 等の教材の場合は、教員等が自分自身のために制作した教材が多く、検索の手がかりとなるメタデータの付与も本人に依存しがちである。それゆえにメタデータ、キーワード等の質は安定しない。特に配布ライセンスや教材の品質が不明であることは利用の障壁である。OCW や Merlot のように大学や大学連携組織は一定の役割を果たしているが十分とはいえない。たとえば、Merlot の検索画面（図 1）は複雑で、教材の形式に精通している人向けに作られている。

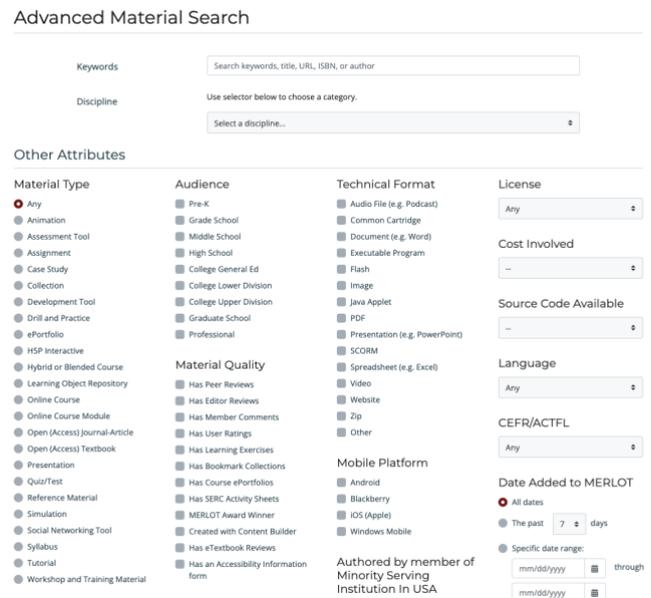


図 1 Merlot の検索画面<sup>1</sup>

さらに、日本には OER を集積、メタデータ付与、カタログ化する実用サイトがない。教材の整備としては、教育情報ナショナルセンター（NICER）[6]や、放送大学が開発した日本オープンコースウェアコンソーシアム（Japan OpenCourseWare Consortium: JOCW）の横断検索システム[7]があるが、いずれも運用を終了している。システムを構築するにしても運用コストが低くなければ持続可能ではないと思われる。

このように持続可能な OER のファインダビリティ向上策は容易ではない。一方でオンライン授業の定着により、質の高い OER のニーズも、作成される教材の量も増加が見込まれる。本研究では以下の背景から以下の手順でシステムの設計と実装をおこなう。

- (1) 教材制作（含 OER）のライフサイクルの検討とシステムのユースケース作成
- (2) OER 情報の収集とカタログ化
- (3) OER に標準に準拠したメタデータの付与
- (4) OER 間の関連視覚化に向けた検討

このシステムでは、教材そのものを蓄積するのではなく、メタデータを保存する。しかし、OER の利点を

<sup>1</sup><https://www.merlot.org/merlot/advSearchMaterials.htm>

活かして、教材をコピーしてキャッシュする。キャッシュされた教材がテキスト分析されて、新たなメタデータが生成される。

以上のシステムは可能なかぎり自動化をおこない運営者の負担がないようにする。一方で、利用者自身が OER の追加、改善やメタデータの改善に関与するオープンソース・ソフトウェア的な仕組みを追加で検討する。

### 3. OER のライフサイクル

一般に教材制作のフローは以下の通りである。ここで教材は教育で用いられる資料（コンテンツ、materials）のことであり、単一のファイルだけではなく、複数の階層構造を持つファイル群や、データベースに保存されている場合もある。

- (1) 教材制作の目的を定める。
- (2) 開始: 新規で教材を制作開始する、または、他の資料をコピーして教材の初期状態とする。この時点で教材として利用可能な場合もある。
- (3) 編集: 教育実践で利用可能（意味のある）状態になるまで教材を改変する。
- (4) 保存: 教材をファイルシステムやデータベース等に保存する。
- (5) 公開 (publish) : 教材をアクセスできるようにする。
- (6) 教材を廃棄する。

一方、教材の利用は以下のように分類される。

- (1) 利用 (Use) : 教材を利用して講義する; 自習で教材を利用する、など。
- (2) 改変 (Revise) : 既存の教材を修正する。
- (3) 複製 (Copy) : 他の教材の一部または全部を複製する。
- (4) リミックス (Remix) : 複数の教材の一部を組み合わせる新しいコンテンツを作成する。
- (5) 配布 (distribute) : 教材を配布する。

作成者以外が教材を利用する場合には、著作権の制

限がかかる。Wiley は OER 利用許諾のパターンを検討して、Reuse、Revise、Retain、Remix、Redistribute の 5R からなるフレームワークを提案している<sup>2</sup>[8, 3]。

オープンソース・ソフトウェアの開発において、複製（発生）、編集、保存、公開（配布）で GitHub などのバージョン管理システムが普及している。教材の制作と利用においてもバージョン管理システムは有用であり、特に OER の配布と利用において利便性が高いと思われる[9]。

### 4. OER のメタデータ

このシステムでは、主なメタデータとして、“OER Metadata Rosetta Stone”のコア項目を採用する。これは米国の SPARC が組織する“OER Discovery Working Group”が提案した OER のメタデータ項目案で、MARC21、DublinCore、Schema.org などのメタデータ標準の語彙を、OER へのアクセスに必要なニーズによって整理したものである。コア項目は、Required（必須）、Recommended（推奨）、Optional（任意）のカテゴリーで分類される。これらに加えて、NICER LOM (IEEE LOM) の項目を必要に応じて追加する。

### 5. OER 探索システムの設計と実装

OER 探索システムの機能、動作手順は以下の通りである。

#### (1) 初期データソースの設定

OER 収集においては、OCW や「数理・データサイエンス・AI 教育強化拠点コンソーシアム」のように組織的に OER を提供するサイトが重要なデータソースである。これらについては、定期的にクローリングをおこない、OER そのものも取得して、メタデータを付与する。それぞれのデータソースについて、OER のライセンスやクローリングの制限についてのデータを記述する。

#### (2) クローリング

---

<sup>2</sup>元の教材と異なる意図、目的でリミックスを許諾する Reboot も必要かもしれない。

クローリングをおこない、Web コンテンツ（多くの場合 HTML）を取得する。

### (3) コンテンツとメタデータの取得

取得した Web コンテンツをパース（スクレイピング）して、コンテンツおよびメタデータを取得する。

### (4) メタデータの保存

取得したメタデータをメタデータ用データベースに保存する。ポータルやコースページなどのように複数のコンテンツが配置されている場合はそれぞれを新しくクローリングする対象として保存する。

### (5) コンテンツの保存

OER を保存する。可能であれば OER のテキストを抽出する。

以上の機能を持つシステムを Python で記述した。データベースには PostgreSQL とテキスト（csv）を併用している。

クローリングは robots.txt などアクセス制限をおこなうサイトがあることから、自動と半自動（一部人が操作）によっておこないアクセス制限を回避している。また、JavaScript による動的レンダリングをおこなうサイトが多いため、レンダリングサーバー Splash をプロキシサーバーとして用いている。スクレイピングは Python で記述している。

検索には Elasticsearch を用いている。取得したメタデータを保存して検索対象としている。

## 6. まとめと今後の課題

教材制作の生産性を向上させるために OER の利用は有効である。本稿では OER のファインダビリティを向上させるためのシステムの設計と実装法について説明した。

開発したプロトタイプの評価については、利用者の満足度のような主観的指標ではなく、ある教材に関連する OER 検索結果の適合度のような客観的な指標の開発を検討している。現在はメタデータを検索することができるだけであるが、メタデータを RDF (Resource Description Framework) で記述して、Linked Open Data (LOD)化することによって検索の

精度も向上すると思われる。

OER のファインダビリティは、単にメタデータの付与やその検索システムによって向上するものではない。教材は、対象分野の知識マップやオントロジーと、教育実践の間につくられる。教材の制作プロセスにおいて、利用可能なオープンなコンテンツを見出すことがシステムに必要である。したがって、教材やコンテンツに表現される知識に基づいた OER 間の関連の表現が今後のファインダビリティ向上において重要な技術的課題であろう。

## 謝辞

本研究は JSPS 科研費 20H01713 の助成を受けたものです。

## 参考文献

- (1) オープン教育資源 (OER) に関する勧告 (仮訳) 2019 年 11 月 25 日 第 40 回ユネスコ総会採択  
[https://www.mext.go.jp/unesco/009/1411026\\_00001.htm](https://www.mext.go.jp/unesco/009/1411026_00001.htm)
- (2) 竹村治雄: “OCW から OER へ: その転換は可能か?”, 工学教育研究講演会講演論文集, 第 61 回年次大会 (平成 25 年度) (2013)
- (3) 重田勝介: “アフターコロナにおけるオープンエデュケーションの展望”, 知能と情報, 33 巻, 4 号, p. 111-116 (2021)
- (4) C. K. Pereira, S. W. M. Siqueira, B. P. Nunes and S. Dietze: "Linked Data in Education: A Survey and a Synthesis of Actual Research and Future Challenges," in IEEE Transactions on Learning Technologies, vol. 11, no. 3, pp. 400-412, 1 July-Sept (2018)
- (5) Herrera-Cubides, J. F., Gaona-García, P. A., Montenegro-Marín, C. E., & Sánchez-Alonso, S.: "Improving OER descriptions to enhance their availability, reuse, and enrichment", Education and Information Technologies, 27(2), 1811-1839. (2022)
- (6) 山田恒夫“放送大学オープンコースウェアの現状と課題”, 工学教育研究講演会講演論文集 25(61) 270-271
- (7) 清水康敬 "高等教育における e-Learning の支援と教育コンテンツの共有" メディア教育研究 1.1: 1-10

- (2004)
- (8) Wiley, D., & Hilton III, J.: “Defining OER-enabled pedagogy”, IRRODL - The International Review of Research in Open and Distributed Learning, 19(4). (2018)
- (9) 重田勝介 , 武田俊之 , 金子大輔 , 八木秀文 , 永嶋知紘: “OER の構造化とバージョン管理を行う Moodle プラグインの開発”, 研究報告教育学習支援情報システム (CLE) ,2021-CLE-35(11),1-5 (2021-11-27) , 2188-8620 (2021)
- (10) OER Discovery Working Group. OER Metadata Rosetta Stone  
<https://docs.google.com/document/d/14tYwNEzr1EKMJAEDHVx2BXeyXFCCSCLGe1lCuySkZE/>

表 1 OER Metadata Rosetta Stone の項目と内容

**Required**

Label	Usage
Title	The title of the resource.
Author/ Creator	The individual credited with the creation of the resource.
Subject	Topic related to the content of the resource. Can include geographic or temporal dimensions. Should not duplicate the discipline of study captured in oer:discipline
Description	brief description of the content of the resource, not including a table of contents
Language	The primary language of the resource. Recommended practice is to use a non-literal value representing a language from a controlled vocabulary such as ISO 639-2 or ISO 639-3.
Date	The date on which the resource was created. Recommended practice is to use the ISO 8601-1.
Material type	The predominant type or kind characterizing the learning resource. For example, 'full course', 'lecture', or 'ancillary material'
Media Format	nature or genre of the resource; newspaper, textbook, video presentation, . not strictly related to the digital object, but the content of the object
Rights Holder	Agent, personal or corporate, that holds the rights to the resource
License Description	Additional information for non-standard rights description
License Title	The proper name of the license or rights statement
License URL	The URL where the owner specifies permissions for using the resource.

**Recommended**

Audience	A class of entity for whom the resource is intended or useful.
Contributor	any agent, personal or corporate, contributing to the creation of the resource, but not

	considered the primary creator(s)
Editor	Any agent having managerial responsibility related to the writing, compilation, and revision of content for a publication
Table of Contents	table of contents of a resource
File Type	Media type typically expressed using a MIME format (see IANA site and MDN reference)
File Size	A measure of how much data a computer file contains or, alternately, how much storage it consumes.
Duration	Time duration of audio/video resource
Identifier	unique identifier for the resource (Examples: ISBN, ISSN, Call Number, File name, Manuscript number, Institutional Repository #, doi)
Peer Review	Is the item peer reviewed? (boolean value) Using exact note “Refereed/Peer-Reviewed” may create display icon in some ILS’s
Course Title	Course name associated with the intended course
Course Identifier	Course code associated with the intended course or course mapping (e.g. CHEM 200)
Education Level	Defines a level or stage within an education system for which the resource being described is intended or useful.As an example, mimic the oercommons Education Level values: <a href="https://www.oercommons.org">https://www.oercommons.org</a>

### Optional

Alternate Title	An alias for the item.
Edition Statement	Example/instance/realization/derivation of the concept of this work. eg. The paperback edition, first edition, or eBook.
Page Count	Number of pages that a resource has
Publisher	The organization credited with publishing the resource.
Provider	Source institution providing metadata and collection
Place	Geographic location of creation
Provenance	Purpose of creation, including grant sponsorship
Relationship	Resources with a relationship to the resource described (Examples: translations, revisions, other formats, etc.)
Is Ancillary	What is the resource an ancillary to?
Has Ancillaries	Does the item have ancillary resources? Are the resource not standalone (Example: test bank)