

Dublin Core — 最近の話題から

杉本重雄
図書館情報大学
sugimoto@ulis.ac.jp

概要

本稿では、Dublin Core Metadata Initiative(DCMI)における最近の話題について簡単に述べる。はじめにDCMIの現在の構成とその再構成について述べ、DCMIのワーキンググループによる最近の成果を紹介する。その後、Resource Description Framework (RDF)形式に基づくDublin Coreメタデータの記述形式に関する現在の推奨方法を紹介し、それに関する検討を示している。

キーワード

Dublin Core Metadata Element Set (DCMES), Dublin Core Metadata Initiative (DCMI), エレメント, 限定子 (qualifier), Resource Description Framework (RDF) による Dublin Core 記述

Dublin Core - Recent Activities

Shigeo Sugimoto
University of Library and Information Science
sugimoto@ulis.ac.jp

abstract

This paper briefly shows the current activities of the Dublin Core Metadata Initiative (DCMI). Firstly, this paper describes the organization structure of DCMI and its re-organization which is being considered and, then recent outcomes from DCMI working groups. Secondly, this paper shows RDF/XML description of both Simple and Qualified Dublin Core based on the proposed recommendation documents issued in the summer 2001. Discussion on the proposed description is given in this paper.

Keywords

Dublin Core Metadata Element Set (DCMES), Dublin Core Metadata Initiative (DCMI), elements, qualifiers, Expressing Dublin Core in Resource Description Framework (RDF)

1. はじめに

Dublin Core の利用が多様な分野で進められている。Dublin Core Metadata Element Set (DCMES) の開発は、Dublin Core Metadata Initiative (DCMI)[1] として、OCLC の Stuart Weibel を中心に、多くの草の根参加者によって進められてきたものである。基本 15 エレメントの定義を与える Simple Dublin Core に関しては 1997 年秋に実質的に確定し、その後、qualifier (限定子) の議論やいろいろな分野での利用に関する議論が進められてきた。2000 年 7 月には最初の推奨限定子セットをアナウンスした。また、エレメントや限定子の拡張に関するプロセスを確立させてきている。その間、2000 年にヨーロッパの CEN/ISSS での標準化 (CWA 13874) に続き、2001 年 10 月に ANSI での承認 (Z39.85) がなされた。

2000 年にカナダ国立図書館 (オタワ) で開催された第 8 回ワークショップ (DC-8) 以降、それまで OCLC の大きな補助とボランティアである草の根参加者によって実質運営されてきた DCMI の再構成に関する議論が行われてきた。その意味では、現在 Dublin Core は開発のフェーズから、維持管理と継続的な開発のフェーズへと移りつつあるといえる。一方、第 9 回ワークショップとして計画された国立情報学研究所 (東京) での会議は、Dublin Core ワークショップの歴史の中では初めて研究や開発に関する論文発表を含む国際会議 International Conference on Dublin Core and Metadata Applications 2001 (DC-2001) として 2001 年 10 月 22 日から 26 日に開催されることになった。

以下、本稿では、筆者の 2000 年 9 月の報告 [2] 以降の話題について簡単に述べる。本稿で述べる話題を理解するのに必要な限定子や Dumb-down 原則などについては、昨年の報告等を参照していただきたい。

2. 最近の話題から

2.1 DCMI の組織再構成

現在の DCMI は全体の運営を行う DCMI Directorate に Executive Director と Managing Director をおき、運営に関して Directorate への助言、フィードバックを行う Executive Committee, ワーキンググループのリーダー他からなる Advisory Committee, 特定分野に関して議論するワーキンググループ, エレメントや限定子に関して議論する Usage Board からできている。これらは毎年行われてきているワークショップで設立、継続あるいは廃止を決めている。

Dublin Core の開発の進展とともに、様々な応用分野への適用が進み、多様な分野の要求に基づくエレメントや限定子に対する要求が出てきている。一方、メタデータの表現と解釈の基礎を与えるデータモデルの定義や DCMES の維持管理のためのレジストリ開発等も求められている。このように、DCMI に対していろいろな活動を進める要求が寄せられる一方、OCLC 以外には大きなスポンサーはないこと、また、DCMI は基本的にボランティアによる組織であることなどのため、DCMES の維持管理や継続的な開発を進める上での限界が見えてきていた。そのため、DC-8 以降 DCMI を再構成することが議論されてきている。現時点では会員組織 (membership organization) 化することや、地域毎に組織をつくり、地域での要求に応じた新しいエレメントや限定子の開発、言語 (たとえば日本語) での標準のメンテナンスなどを行うことが検討されている。

2.2 ワーキンググループ

現在あるワーキンググループとその目的を表 1 にまとめて示す。以下にワーキンググループの活動成果のいくつかをごく簡単に紹介する。

Education ワーキンググループからは、2000 年 10 月に出された working draft で

- (1) 資料の利用者を表す dc-ed:audience エlementとその限定子
- (2) 教育やトレーニングに関する標準を表す dc-ed:standard エlementとその限定子
- (3) relation エlementの詳細化限定子として、教育に関する標準との関連を表すための dc-ed:conforms 限定子
- (4) IEEE LOM の namespace 内に定義されるデータElementの承認が提案されている。

Library グループからはタイトルを必須とすること、資源の作成に関わった者に関する記述を条件付きで必須とすること、資料の利用者や所蔵者に関するElementなどを定義した application profile が提案されている。

Architecture グループからは Simple Dublin Core と Qualified Dublin Core の RDF による記述方法が Proposed Recommendation として出されている。レジストリワーキンググループではレジストリに関する議論とレジストリの試作を進めてきている。これについては後節で述べる。

表1 DCMIのワーキンググループ(2000-2001)

グループ名	設置目的
Administrative Metadata	Administrative Metadataに関するElementセットの提案
Agents	Creator, Contributor, Publisherに関わる、いわゆるagentに関する表現のための方法とElementセットに関して提案する。
Architecture	Dublin Coreを実際的な環境で利用していくためのデータモデルや方法等を提案する。
Citation	雑誌や雑誌記事などに関する書誌引用情報のDublin Coreに適合した表現方法を提案する。
Collection Description	コレクションのためのDublin Coreに基づく表現方法(application profile)を提案する。
Education	教育情報資源のためのElementを提案する。
Government	政府・行政に関する情報資源の利用を進めるためのメタデータに関して議論し、政府・行政情報資源のためのElementを提案する。
Libraries	図書館でのDublin Coreの利用を進めること、またそのためのコミュニティ作りなどを行う。
Registry	DCMIの活動を支えるためのメタデータレジストリの開発を進める。
Standards	Dublin Coreの標準化に関して議論するとともに、国家あるいは国際標準化機関でのDublin Coreの標準化に寄与する。
Tools	Dublin Coreに関連する様々なツール作りを進めるための活動を行う。
Type	Type Elementに関する議論を行い、RFC 2413の修正とType Elementの限定子に関する議論を進める。
User Guide	Dublin CoreのUser Guideの作成を進める。

3. Resource Description Framework (RDF) による記述

Architecture ワーキンググループからは Simple DC (DCS) と Qualified DC (DCQ) の RDF による記述形式を定義した文書が Proposed Recommendation として、それぞれ本年 9 月と 8 月に出されている [3][4]。以下にそこから記述例をいくつか紹介したい。なお、ここで参照している文書は recommendation (推奨, 勧告) と認められたものではなく、その前段階であるのでこれからの変更も有得る。

3.1 Simple DC の場合

例： DCMI のホームページのメタデータ

```

<?xml version="1.0"?>
<!DOCTYPE rdf:RDF PUBLIC "-//DUBLIN CORE//DCMES DTD 2001 09 20//EN"
"http://dublincore.org/documents/2001/09/20/dcmes-xml/dcmes-xml-dtd.dtd">
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://dublincore.org/">
    <dc:title>Dublin Core Metadata Initiative - Home Page</dc:title>
    <dc:description>The Dublin Core Metadata Initiative Web site.
      </dc:description>
    <dc:date>1998-10-10</dc:date>
    <dc:format>text/html</dc:format>
    <dc:language>en</dc:language>
    <dc:contributor>The Dublin Core Metadata Initiative</dc:contributor>
    <!-- guesses for the translation of the above titles -->
    <dc:title xml:lang="fr">L'Initiative de metadonnees du Dublin
      Core</dc:title>
    <dc:title xml:lang="de">der Dublin-Core Metadata-Diskussionen
      </dc:title>
  </rdf:Description>
</rdf:RDF>

```

DCS の場合、rdf:Description タグで記述対象の資源を指示するとともに、エレメントを囲んでいる。各エレメントを表すタグは Dublin Core のエレメント名の前に namespace 名として"dc" が付いたタグになっている。また、この例の下の方にはフランス語とドイツ語によるタイトルエレメントの記述がある。ここでは記述言語をあらわすために rdf:lang 属性を用いている。

3.2 Qualified DC の記述

DCQ はエレメント詳細化 (element refinement) とコード化スキーム (encoding scheme) の 2 種類の限定子をどう表すかの定義をきちんとしなければならない。これらの限定子の意味を RDF のモデル上で定義することが求められるため、Proposed Recommendation 文書もデータモデルの説明にかなりの場所を費やしている。

例 1： 詳細化 (element refinement)

内容記述 (description) エレメントの詳細化限定子である abstract を用いた例である。RDF Schema によって abstract が description の sub-property として定義される。(この例では RDF Schema 記述が含まれている。) abstract が description の限定子であることは一意に定まることもあり、dc:description はこの例には現れず、直接 dcq:abstract で表している。

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns:dcq="http://purl.org/dc/terms/">
  <rdf:Description>

```

```

    <dcq:abstract>The paper resolves the issues of the data model
      draft.  </dcq:abstract>
  </rdf:Description>
  <rdf:Description about="http://purl.org/dc/terms/abstract">
    <rdfs:subPropertyOf
      rdf:resource="http://purl.org/dc/elements/1.1/description"/>
  </rdf:Description>
</rdf:RDF>

```

例 2: コード化スキーム (encoding scheme)

subject エレメントに MESH の値を入れている例である。dcq:MESH タグを dc:subject の中に入れ、その中に rdf:value と rdf:label の二つのタグを入れている。rdf:value は主たる値を表すために用いられるもので、rdf:label は主たる値に付随する名前や説明を書くために用いられている。

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcq="http://purl.org/dc/terms/">
  <rdf:Description>
    <dc:subject>
      <dcq:MESH>
        <rdf:value>D08.586.682.075.400</rdf:value>
        <rdfs:label>Formate Dehydrogenase</rdfs:label>
      </dcq:MESH>
    </dc:subject>
  </rdf:Description>
</rdf:RDF>

```

例 3: 言語属性

DCS ではメタデータの記述に用いられる言語をタグ内に rdf:lang 属性を用いて記述している。一方、DCQ に対する RDF 記述の文書では、記述言語をタグとして表す方法を提案している。下の二つの記述のうち、上は xml:lang 属性を用いたもの、下は xml:lang をタグとして用いたものであり、この提案では後者を勧めている。

・ xml:lang 属性

```

<?xml version="1.0"?>
<rdf:RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description>
    <dc:creator>Karl Mustermann</dc:creator>
    <dc:title>Algebra</dc:title>
    <dc:subject xml:lang="en">mathematics</dc:subject>
    <dc:subject xml:lang="de">Mathematik</dc:subject>
  </rdf:Description>
</rdf:RDF>

```

```
</rdf:Description>
</rdf:RDF>
```

・ xml:lang タグ

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:dcq="http://purl.org/dc/terms/">
  <rdf:Description>
    <dc:creator>Karl Mustermann</dc:creator>
    <dc:title>Algebra</dc:title>
    <dc:subject rdf:parseType="Resource">
      <rdf:value>mathematics</rdf:value>
      <dc:language>
        <dcq:RFC1766>
          <rdf:value>EN</rdf:value>
          <rdfs:label>English</rdfs:label>
          <rdfs:isDefinedBy
            rdf:resource="http://www.ietf.org/rfc/rfc1766.txt"/>
        </dcq:RFC1766>
      </dc:language>
    </dc:subject>
    <dc:subject rdf:parseType="Resource">
      <rdf:value>mathematik</rdf:value>
      <dc:language>
        <dcq:RFC1766>
          <rdf:value>DE</rdf:value>
          <rdfs:label>German</rdfs:label>
          <rdfs:isDefinedBy
            rdf:resource="http://www.ietf.org/rfc/rfc1766.txt"/>
        </dcq:RFC1766>
      </dc:language>
    </dc:subject>
  </rdf:Description>
</rdf:RDF>
```

Dumb-down に関する検討 :

DCQ の記述は Dumb-down 原則に基づき限定子を取り除くことができなければならない。上の記述の場合、ごく単純には、詳細化限定子の場合、その限定子が定義されたエレメントに置き換えればよい。また、コード化スキーム限定子の場合には rdf:vaue で示された値を取り出し、その限定子が定義されたエレメントの値とするという方法で Dumb-down を行うことができる。

3.3 検討

上の記述形式に関して筆者の観点から検討を加えてみたい。Simple DC に関する RDF 記述形式は単純であり、これまでの HTML の Meta タグを使った例と基本的に同じである。一方、Qualified DC は限定子という新たな概念が入ってきているため、複雑になっている。ここでは Qualified DC の RDF 記述に関して簡単に検討してみたい。

(1) 詳細化限定子 (element refinement qualifier) について

現在までに承認されている詳細化限定子は、それが属するエレメントが一意に定まるという関係で与えられている。(コード化スキーム限定子の場合は異なるエレメントに対して同じ限定子が対応する場合がある。付録参照。) 筆者には、ある詳細化限定子に対するエレメントが一意に決まらなければならないという関係はかなり制約としては厳しく感じられる。今後、もしこの関係が崩れた場合には対応できないことになる。

(2) コード化スキーム限定子について

従来の SGML の記述方法に基づくと、コード化スキーム限定子のように値の属性を定義するものはタグそのものではなく、タグの属性として定義するのが好ましいと思われる。しかしながら、RDF は、Resource, Property, Value の 3 つ組みでの表現を原則とするため、属性による表現が行いづらい。そのため、ここではコード化スキーム限定子をタグとして記述するとともに、主たる値を表すための `rdf:value` と `rdf:label` を用いて主たる値とそれに付随する情報の表現を行っている。また、構造を持つ値に関しては、推奨方式 (DCSV) [5] を利用して文字列として表している。

RDF の制約上仕方ない点はあると思うが、構造が複雑になること、詳細化とコード化スキームの別を構文から知ることができないことなどの問題があるように思う。また、同様にメタデータの記述言語を属性からはずしてタグとして記述していることも構造を複雑にしていると感じられる。RDF の記述は複雑になることから、適切なソフトウェアツールの整備が必要であることが良く理解できる。

4. Dublin Core レジストリ

DCMI では Dublin Core のエレメントの定義をネットワーク上で登録、提供するレジストリサービスが重要であると認め、レジストリワーキンググループを作ってソフトウェアの開発を進めてきている。筆者等も基本 15 エレメントの参照記述とその翻訳を RDF Schema 記述し、蓄積した多言語メタデータレジストリを開発してきた [6]。DCMI では Eric Miller を中心として EOR と呼ぶレジストリのためのソフトウェアツールの開発を進めた。メタデータ規則を収集し提供することを目的としてヨーロッパで進められた Schemas プロジェクト [7] では EOR を用いてレジストリを開発してきている。現在、DCMI では新たにレジストリを開発を進めている。

今後、レジストリはメタデータ規則 (エレメント, 限定子, application profile の定義や参照記述) の登録、蓄積、提供のために用いられ、メタデータ規則の維持管理と継続的な開発や、メタデータを利用するコミュニティ間でのメタデータの相互利用等のために重要な役割を持つと期待されている。

5. おわりに

Dublin Core は現在既に様々な分野での利用がひろがってきている。今後もこうした利用は広がりつづけると期待される。一方、WWW コンソーシアムによる Semantic Web activity のように、より内容的な理解を計算機に行わせることを目的とした研究開発が盛んに行われている。こうした動きによって、語彙の定義などより意味的内容を形式的に定義することに関心がもたれている。Dublin Core の場合、応用側で

は、こうした形式的取り組みとはあまり直接的なつながりは出てこないかもしれないが、アーキテクチャやレジストリなど基本的なデータモデルを扱う分野にはごく近い取り組みであると思われる。

参考文献

- [1] Dublin Core Metadata Initiative. <http://dublincore.org/>
- [2] 杉本重雄, Dublin Core について－最近の動向, 特に qualifier について, デジタル図書館, No.18, 2000.9, <http://www.dl.ujis.ac.jp/DLjournal/No.18/4-sugimoto/4-sugimoto.html>
- [3] Dave Beckett, Dan Brickley, Eric Miller, Expressing Simple Dublin Core in RDF/XML, <http://www.dublincore.org/documents/2001/09/20/dcmes-xml/>, 2001.9
- [4] Stefen Kokkeliink, Roland Schwaenzl, Expressing Qualified Dublin Core in RDF / XML, <http://www.dublincore.org/documents/2001/08/29/dcq-rdf-xml/>, 2001.8
- [5] SimonCox, Renato Ianela, DCMI DCSV: A syntax for writing a list of labelled values in a text string, <http://dublincore.org/documents/2000/07/28/dcmi-dcsv/>, 2000.7
- [6] Nagamori Mitsuharu, et al. A Multilingual Metadata Schema Registry Based on RDF Schema, Proceedings of DC-2001, 2001.10
- [7] Forum for Metadata Schema Implementers, <http://www.schemas-forum.org/>

付録 基本 15 エlementおよび限定子 (qualifier) 一覧

Dublin Coreの基本15Element

要素名	identifier	定義および説明
タイトル	Title	情報資源に与えられた名前。
作成者	Creator	情報資源の内容の作成に主たる責任を持つ実体。
主題およびキーワード	Subject	情報資源の内容のトピック。
内容記述	Description	情報資源の内容の記述。
公開者(出版者)	Publisher	情報資源を利用可能にすることに対して責任を持つ実体。
寄与者	Contributor	情報資源の内容への寄与に対して責任を持つ実体。
日付	Date	情報資源のライフサイクルにおける何らかの事象に対して関連付けられた日付。
資源タイプ	Type	情報資源の内容の性質もしくはジャンル。
形式	Format	物理的表現形式ないしデジタル形式での表現形式。
資源識別子	Identifier	与えられた環境において一意に定まる情報資源に対する参照。
情報源(出処)	Source	現在の情報資源が作り出される源になった情報資源への参照。
言語	Language	当該情報資源の内容の言語。
関係	Relation	関連情報資源への参照。
対象範囲(空間的・時間的)	Coverage	情報資源の内容が表す範囲あるいは領域。
権利管理	Rights	情報資源に含まれる、ないしは関わる権利に関する情報。

DCMI Qualifierの一覧

DCMES Element	Element Refinement(s)	Element Encoding Scheme(s)
Title	Alternative	-
Creator	-	-
Subject	-	LCSH MeSH DDC LCC UDC
Description	Table Of Contents Abstract	-
Publisher	-	-
Contributor	-	-
Date	Created Valid Available Issued Modified	DCMI Period W3C DTF
Type	-	DCMI Type Vocabulary
Format	Extent Medium	- IMT
Identifier	-	URI
Source	-	URI
Language	-	ISO 639-2 RFC 1766
Relation	Is Version Of Has Version Is Replaced By Replaces Is Required By Requires Is Part Of Has Part Is Referenced By References Is Format Of Has Format	URI
Coverage	Spatial	DCMI Point ISO 3166 DCMI Box TGN
	Temporal	DCMI Period W3C-DTF
Rights	-	-