

DCMIメタデータスキーマ レジストリとその応用

筑波大学図書館情報メディア研究科

永森光晴

nagamori@slis.tsukuba.ac.jp

DCMI Metadata Schema Registry

- DCMI (Dublin Core Metadata Initiative) で定義するメタデータスキーマを提供するための authoritative source を作成する
- 人間だけではなくソフトウェアからも利用可能にする
- 様々な言語に対応できるようにする

メタデータ

- セマンティックウェブやデジタルライブラリを中心としてメタデータが注目されている
- 情報資源の組織化と発見を目的としている
- *Metadata is the information and documentation which makes data understandable and shareable for users over time. (ISO/IEC 11179)*

2004/10/29

3

様々な種類のメタデータ

- Dublin Core Metadata Element Set
- IEEE LOM
- MODS (Metadata Object Description Schema)
- METS (Metadata Encoding and Transmission Standard)

2004/10/29

4

メタデータスキーマ

- メタデータの記述規則の定義
- 識別子、ラベル、記述内容、作成日、言語などを記述
- DCMIでは以下のメタデータスキーマを定義
 - Dublin Core Metadata Element Set (DCMES)
 - Element Refinements
 - Encoding Schemes
 - DCMI Type Vocabulary

2004/10/29

5

DCMES の title 要素の定義

URI	http://purl.org/dc/elements/1.1/title
Label	Title
Definition	A name given to the resource.
Comment	Typically, a Title will be a name by which the resource is formally known.
Type of Term	element
Status	recommended
Date Issued	1999-07-02

2004/10/29

6

RDFによるDCMESの定義(一部)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE rdf:RDF (View Source for full doctype...)>
- <rdf:RDF xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="http://purl.org/dc/elements/1.1/">
    <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/title">
      <rdfs:label xml:lang="en-US">Title</rdfs:label>
      <rdfs:comment xml:lang="en-US">A name given to the resource.</rdfs:comment>
      <dc:description xml:lang="en-US">Typically, a Title will be a name by which the
        resource is formally known.</dc:description>
      <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
      <dcterms:issued>1999-07-02</dcterms:issued>
      <dcterms:modified>2002-10-04</dcterms:modified>
      <dc:type
        rdf:resource="http://dublincore.org/usage/terms/types/#dcterms-Title" />
      <dcterms:hasVersion
        rdf:resource="http://dublincore.org/usage/terms/types/#dcterms-Title-1.1" />
    </rdf:Property>
  - <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator">
    <rdfs:label xml:lang="en-US">Creator</rdfs:label>
```

titleの定義

識別子, ラベル, 定義内容, 作成日, 更新日, バージョンなどを記述

メタデータスキーマを利用するのは...

- メタデータ作成者
 - リファレンスとして参照する
- メタデータスキーマ作成者
 - 新たなメタデータスキーマ作成の参考
 - メタデータスキーマの関連を参照
- ソフトウェア
 - レジストリに登録されている情報をソフトウェアの中でデータとして利用

利用が進むにつれ

- 多言語
 - ひとつのメタデータスキーマが様々な言語に翻訳される (DCMESは24言語以上に翻訳されている)
- バージョン
 - メタデータスキーマの改版にともない、新たな要素 (element) が追加されたり、定義が修正されることがある

2004/10/29

9

メタデータスキーマの管理が重要

- 様々な種類、言語、バージョン
- 長期に渡ってメタデータを利用したい



メタデータスキーマを管理するための仕組み
(= **メタデータスキーマレジストリ**) が必要

2004/10/29

10

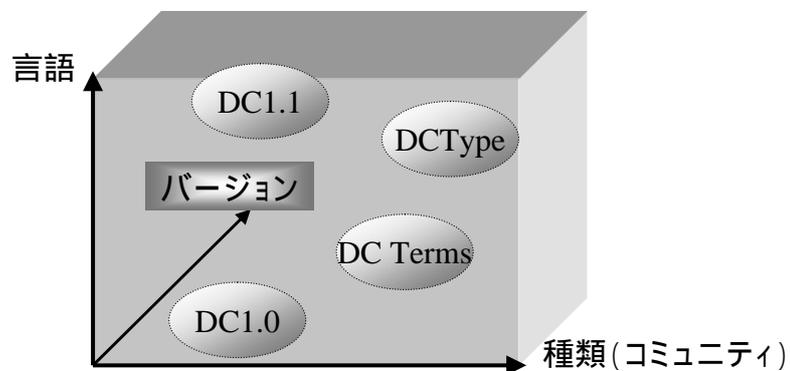
DCMI Metadata Schema Registry

- DCMI (Dublin Core Metadata Initiative) で定義するメタデータスキーマを提供するための authoritative source を作成する
- 人間だけではなくソフトウェアからも利用可能にする
- 様々な言語に対応できるようにする

2004/10/29

11

メタデータスキーマレジストリの管理するもの



2004/10/29

12

Registry WG の活動

- DCMIでは Registry Working Group においてメタデータスキーマレジストリに関する議論と実装をおこなっている
- 1999年フランクフルト(独)で開催された会議(DC-7)が始まり

2004/10/29

13

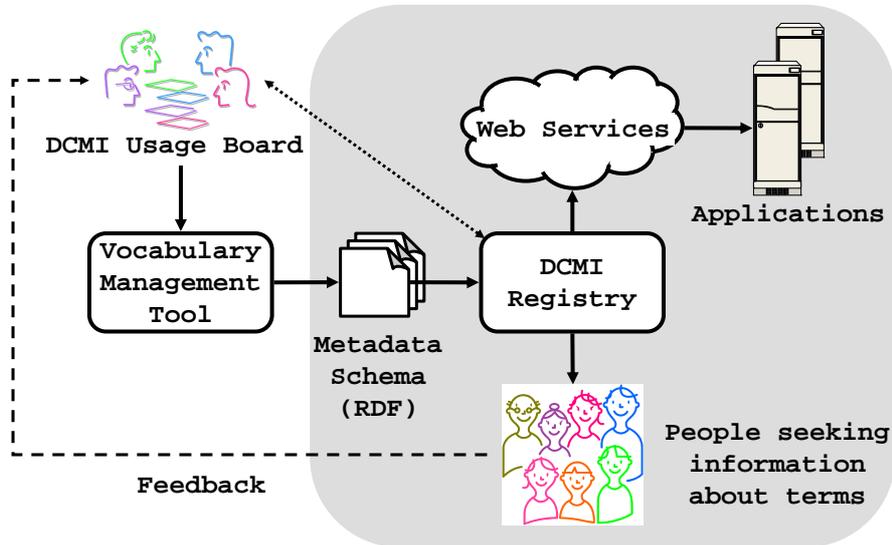
DCMI Metadata Schema Registryの 機能

- Web インタフェース
 - Webブラウザを利用したスキーマの検索と閲覧
- Application インタフェース
 - Webサービスを利用したアプリケーションからのレジストリの利用
- スキーマ管理機能
 - Webブラウザを利用したスキーマの登録と修正

2004/10/29

14

Registry Model



2004/10/29

15

実装のポリシー

- メタデータスキーマの記述形式として標準的に利用されているRDFを基礎とする
- 各国の利用者を考え、登録するデータだけではなくレジストリのインターフェースも多言語に対応させる

2004/10/29

16

DCMI Registry (トップページ)



インタフェースの切り替え



検索インターフェース



検索結果(一覧表示)



検索結果(詳細表示)

The screenshot shows a search result for the term 'alternative' in the Open Metadata Registry. The search criteria are set to 'exact match' (語句の完全一致). The results table lists the following details:

項目	値
定義	対象情報資源の正式なタイトルの代わりもしくは別タイトルとして利用される タイトル。
内容記述	この qualifier はタイトルの翻訳と同等省略形を表すのに用いてもかまわない。
定義元(is defined by)	http://purl.org/dc/terms/
RDFタイプ	Property
資源タイプ	element-refinement
版として持つ(関係)	alternative-002
発行された(日付)	2000-07-11
修正・更新された(日付)	2002-06-15
下位属性(sub-property)	title

カテゴリからの検索

The screenshot shows the search interface of The Open Metadata Registry. A dropdown menu is open, listing various metadata schema categories. The categories listed are:

- コード化形式限定子(encoding scheme)
- 要約
- スキーマ
- クラス
- 属性
- エレメント
- エレメント詳細化限定子(element refinement)
- 統制語彙の用語(controlled vocabulary terms)
- コード化形式限定子(encoding scheme)

A text box below the dropdown menu states: **メタデータスキーマの種類から検索することも可能** (It is also possible to search by metadata schema type).

メタデータスキーマの管理

ひとつのメタデータスキーマを3階層に分けて管理している

- Collection (DCMESやDCQなど)
- Term (title, creator, dateCreated, etc...)
- Property (各エレメントの定義)

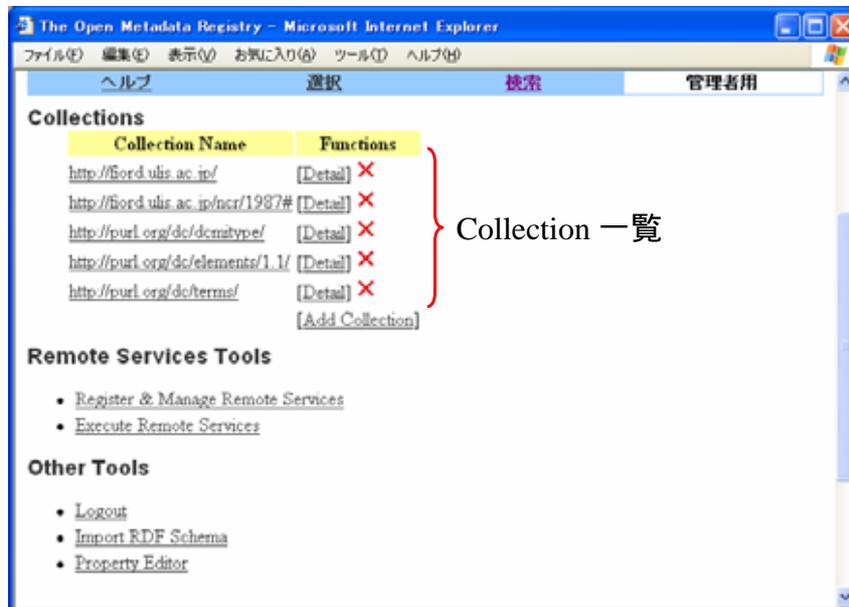
2004/10/29

23

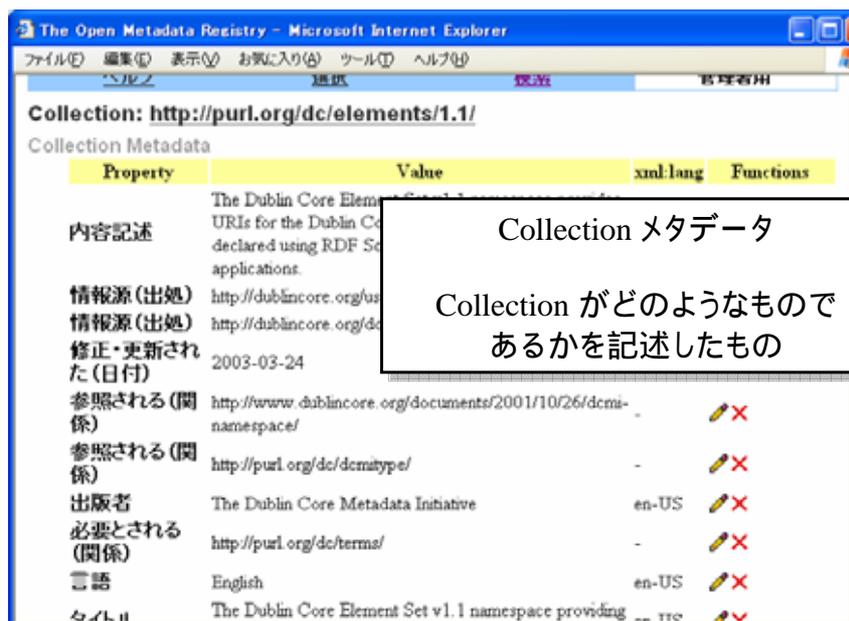
RDFによるDCMESの定義(一部)

```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE rdf:RDF (View Source for full doctype...)>
- <rdf:RDF xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description rdf:about="http://purl.org/dc/elements/1.1/">
  - <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/title">
    <rdfs:label xml:lang="en-US">Title</rdfs:label>
    <rdfs:comment xml:lang="en-US">A name given to the resource.</rdfs:comment>
    <dc:description xml:lang="en-US">Typically, a Title will be a name by which the
      resource is formally known.</dc:description>
    <rdfs:isDefinedBy rdf:resource="http://purl.org/dc/elements/1.1/" />
    <dcterms:issued>1999-07-02</dcterms:issued>
    <dcterms:modified>2002-10-04</dcterms:modified>
    <dc:type
      rdf:resource="http://dublincore.org/usage/documents/principles/#element" />
    <dcterms:hasVersion
      rdf:resource="http://dublincore.org/usage/terms/history/#title-004" />
  </rdf:Property>
  - <rdf:Property rdf:about="http://purl.org/dc/elements/1.1/creator">
    <rdfs:label xml:lang="en-US">Creator</rdfs:label>
```

メタデータスキーマ管理機能



Collection (1)



Collection(2)

Defined Terms/Schemes in the Collection

Name	Functions
http://purl.org/dc/elements/1.1/contributor	[Detail] ✖
http://purl.org/dc/elements/1.1/coverage	[Detail] ✖
http://purl.org/dc/elements/1.1/creator	[Detail] ✖
http://purl.org/dc/elements/1.1/date	[Detail] ✖
http://purl.org/dc/elements/1.1/description	[Detail] ✖
http://purl.org/dc/elements/1.1/format	[Detail] ✖
http://purl.org/dc/elements/1.1/identifier	[Detail] ✖
http://purl.org/dc/elements/1.1/language	[Detail] ✖
http://purl.org/dc/elements/1.1/publisher	[Detail] ✖
http://purl.org/dc/elements/1.1/relation	[Detail] ✖
http://purl.org/dc/elements/1.1/rights	[Detail] ✖
http://purl.org/dc/elements/1.1/source	[Detail] ✖
http://purl.org/dc/elements/1.1/subject	[Detail] ✖
http://purl.org/dc/elements/1.1/title	[Detail] ✖
http://purl.org/dc/elements/1.1/type	[Detail] ✖

そのCollectionで定義しているtermの一覧

Termの管理

http://purl.org/dc/elements/1.1/title

Metadata about the Term

Property	Value
RDFタイプ	http://www.w3.org/1999/02/22-rdf-syntax-ns#Property
資源タイプ	http://dublincore.org/usage/documents/principles/#element
内容記述	一般而言, 这一名称指的是资源对象的正式公开的名称。
修正・更新された(日付)	2002-10-04
定義	当該情報資源に与えられた名前。一般には作者もしくは公開者によって与えられる。
ラベル(label)	표제
定義元(is defined by)	http://purl.org/dc/elements/1.1/
版として持つ(関係)	
発行された(日付)	

定義しているプロパティとその値の一覧を表示

複数の言語で定義している場合はメニューによって切り替える

Application インタフェース

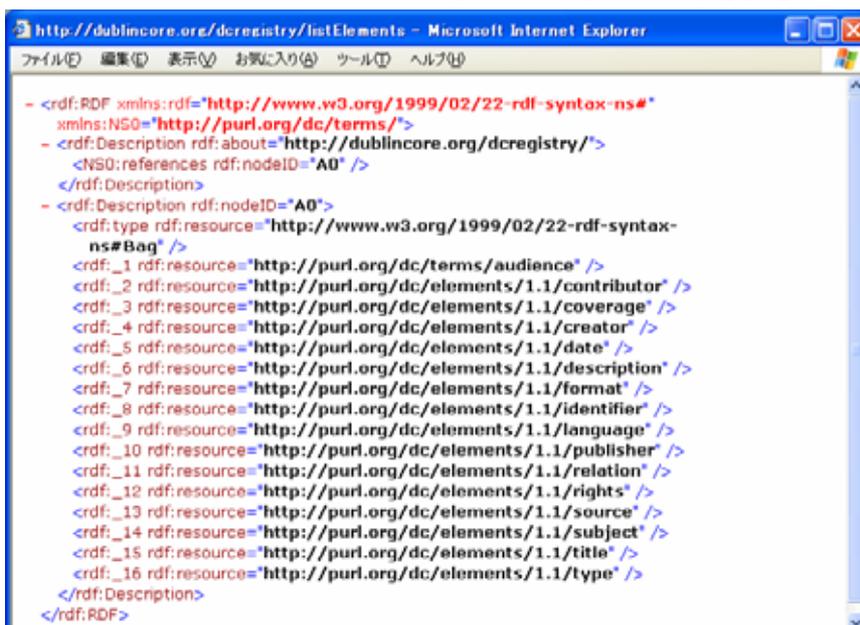
- GET・POST、SOAPの2種類
- RDF形式のデータを返す
- メタデータスキーマをやり取りするための基本的な機能を準備している
 - listItemSummary()
 - listItemDetail()
 - listElements()
 - listItemRefinements()
 - listVocabularyTerms()
 - listEncodingSchemes()
 - listSupportedLanguages()
 - getUpdates()

<http://dublincore.org/dcregistry/helpServlet>

2004/10/29

29

listElements を呼び出した結果



```
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:NS0="http://purl.org/dc/terms/"
  - <rdf:Description rdf:about="http://dublincore.org/dcregistry/"
    <NS0:references rdf:nodeID="A0" />
  </rdf:Description>
  - <rdf:Description rdf:nodeID="A0">
    <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns:#Bag" />
    <rdf:_1 rdf:resource="http://purl.org/dc/terms/audience" />
    <rdf:_2 rdf:resource="http://purl.org/dc/elements/1.1/contributor" />
    <rdf:_3 rdf:resource="http://purl.org/dc/elements/1.1/coverage" />
    <rdf:_4 rdf:resource="http://purl.org/dc/elements/1.1/creator" />
    <rdf:_5 rdf:resource="http://purl.org/dc/elements/1.1/date" />
    <rdf:_6 rdf:resource="http://purl.org/dc/elements/1.1/description" />
    <rdf:_7 rdf:resource="http://purl.org/dc/elements/1.1/format" />
    <rdf:_8 rdf:resource="http://purl.org/dc/elements/1.1/identifier" />
    <rdf:_9 rdf:resource="http://purl.org/dc/elements/1.1/language" />
    <rdf:_10 rdf:resource="http://purl.org/dc/elements/1.1/publisher" />
    <rdf:_11 rdf:resource="http://purl.org/dc/elements/1.1/relation" />
    <rdf:_12 rdf:resource="http://purl.org/dc/elements/1.1/rights" />
    <rdf:_13 rdf:resource="http://purl.org/dc/elements/1.1/source" />
    <rdf:_14 rdf:resource="http://purl.org/dc/elements/1.1/subject" />
    <rdf:_15 rdf:resource="http://purl.org/dc/elements/1.1/title" />
    <rdf:_16 rdf:resource="http://purl.org/dc/elements/1.1/type" />
  </rdf:Description>
</rdf:RDF>
```

DCMIレジストリのまとめ

- メタデータスキーマの検索、閲覧、管理機能
- Dublin Core だけではなく、RDF形式で記述されたものであれば登録できる
- UTF-8を利用した様々な言語への対応
- インタフェースを24言語に切り替えられる
- Web Serviceを利用したレジストリへの問い合わせ機能

2004/10/29

31

開発環境

- Java 1.4
- Apache 1.3
- Jakarta Tomcat 4
- Jena 1.6 (RDF API, <http://www.hpl.hp.com/semweb/>)
- PostgreSQL
- Apache Axis (SOAP)
- XSLT (Apache Xerces, Xalan)

2004/10/29

32

レジストリの設置状況

- 現在5ヶ所に設置
 - DCMI(オハイオ州ダブリン)
 - 筑波大学
 - UKOLN(英国) DC以外のスキーマが登録されている
 - Goettingen大学(独)
 - 中国科学院文献情報中心(Library of Chinese Academy of Science, 中国)
- オーストラリア、FAO(Food and Agriculture Organization of the UN)で準備を進めている

2004/10/29

33

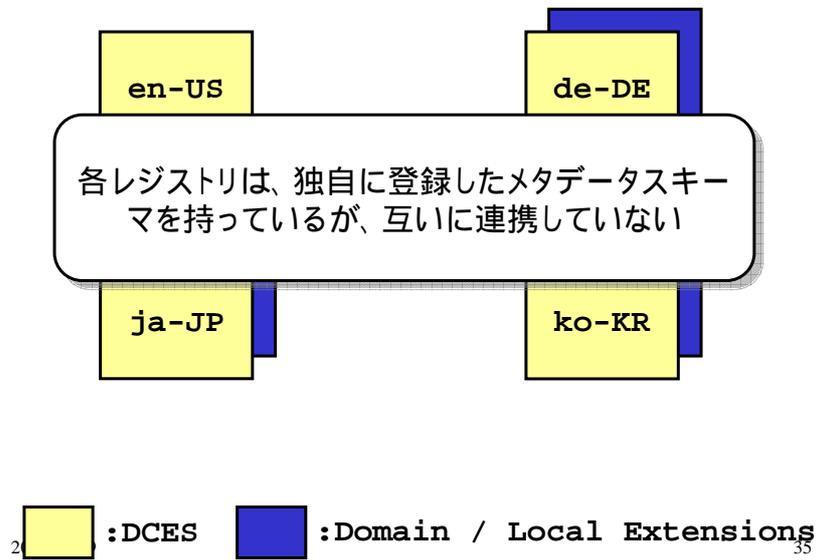
今後の課題

- 分散型メタデータスキーマレジストリ
- RDF形式以外のドキュメントやApplication Profileをどのようにサポートするか
- メタデータスキーマのバージョン管理
 - メタデータスキーマだけではなく、その中で定義している個々のtermのバージョン管理も考える必要がある
- レジストリを利用したソフトウェアの開発
 - 横断検索、メタデータの変換
 - ボキャブラリ作成支援
 - メタデータを利用したソフトウェアの構築支援

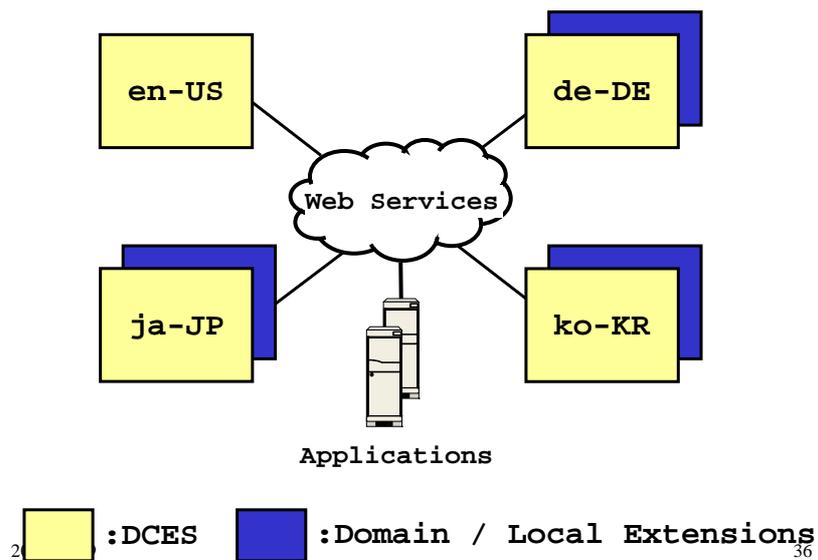
2004/10/29

34

現在のレジストリのモデル



Distributed Registry Model



今後の課題

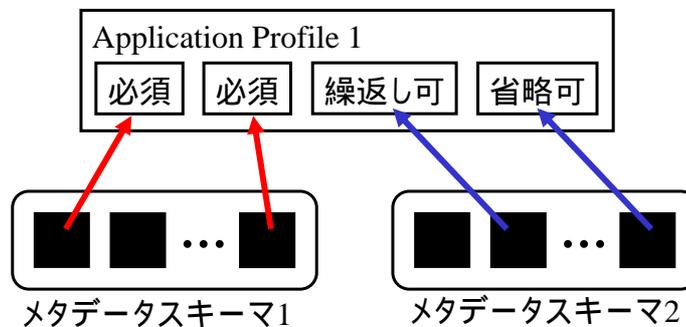
- 分散型メタデータスキーマレジストリ
- RDF形式以外のドキュメントやApplication Profileをどのようにサポートするか
- メタデータスキーマのバージョン管理
 - メタデータスキーマだけではなく、その中で定義している個々のtermのバージョン管理も考える必要がある
- レジストリを利用したソフトウェアの開発
 - 横断検索、メタデータの変換
 - ボキャブラリ作成支援
 - メタデータを利用したソフトウェアの構築支援

2004/10/29

37

Application Profile

- 応用ごとに決める構造的・構文的制約
- 必須・省略可能、繰り返し回数...



2004/10/29

38

今後の課題

- 分散型メタデータスキーマレジストリ
- RDF形式以外のドキュメントやApplication Profileをどのようにサポートするか
- メタデータスキーマのバージョン管理
 - メタデータスキーマだけではなく、その中で定義している個々のtermのバージョン管理も考える必要がある
- レジストリを利用したソフトウェアの開発
 - 横断検索、メタデータの変換
 - ボキャブラリ作成支援
 - メタデータを利用したソフトウェアの構築支援

2004/10/29

39

翻訳を利用するまでの長い道のり

- これまで多くの方々がDCMIのメタデータスキーマを翻訳して下さった
- しかしながら送られてきたものは...
 - ファイルフォーマットがまちまち
 - 文字コードは、翻訳した言語固有のもの
 - UTF-8に変換するために手間がかかる
 - 送られたものを扱うことができない場合もあった(アルメニア語、パンジャブ語)

2004/10/29

40

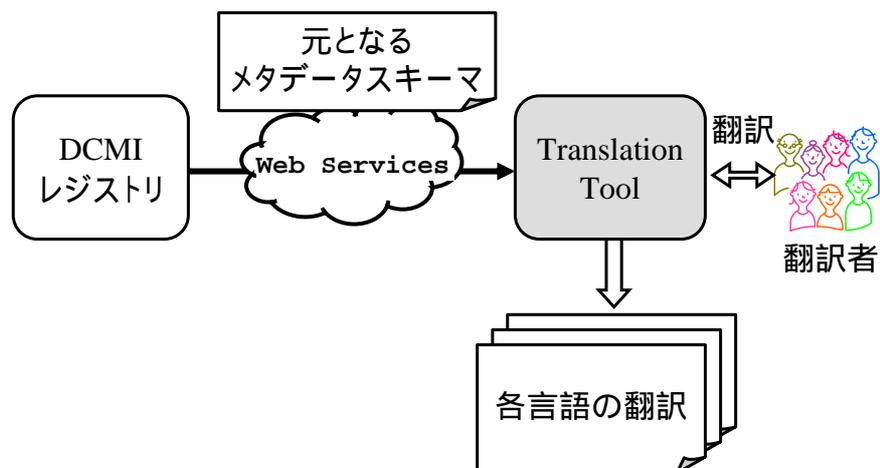
Translation Tool

- メタデータスキーマの翻訳を支援するためのツール
- Webブラウザを使ってスキーマの翻訳を編集することができる
- DC Registry の Application Interface を利用したアプリケーション
- 結果、翻訳されたものを簡単に利用できるようになった

2004/10/29

41

Translation Tool の概要



2004/10/29

42

<http://wip.dublincore.org/translate/>



翻訳する対象を選択



Termの一覧



翻訳の編集



おわりに

<http://dublincore.org/dcregistry/>

<http://dublincore.org/groups/registry/>