

新しい目録規則(RDA) から得られるもの： 機械可読性の視点から

渡邊 隆弘(情報組織化研究グループ)

Evaluation of new cataloging rules, RDA: From the viewpoint of machine-readability, by WATANABE Takahiro.

1. はじめに

2010年6月、AACR2(英米目録規則第2版)の後継規則であるRDA(Resource Description and Access)がついに刊行された¹⁾。FRBR(書誌レコードの機能要件)/FRAD(典拠データの機能要件)モデルに密着した構成をもつRDAの完成は、1997年のFRBR発表から十数年を経て、英語圏を中心とする目録法の改革の動向の、一つの大きな節目といえる。

今次の改革は、「パリ原則」とISBD(国際標準書誌記述)によって1960~70年代(即ちカード目録が主流の時代)に確立された枠組みを抜本的に見直すとともに、インターネット時代における情報環境の大きな変化にも対応しようとするものであり、様々な側面を含んでいる。その全体的な概括は先行研究でも試みられている²⁾ので、本稿では今後の目録法において非常に重要な点と筆者が考える、「機械可読性」に焦点を絞ってRDAを評価したい。

目録情報の機械可読性、すなわち人間だけではなくコンピュータに理解(解析)できるデータ保持は、これからの図書館目録に欠かせない重要点である。それは、「次世代OPAC」のように図書館コミュニティ内で付加価値の高い新たなサービスを構築していくためにも重要であると同時に、目録情報を外部のコミュニティに広く開放していく場合の基盤ともなる。いわゆる「Web2.0」時代にあっては、閉じられた世界で正しく動作するだけでなく、外部に開放され他のシステムと協働しうるシステムやデータが重んじられる。特に近年は、セマンティックウェブの文脈において“Linked Data”(リンクされる

データ)の重要性が説かれるようになってきた。Linked Dataとは、適切に機械可読化され(多くはRDF形式)、一意性と永続性を持った識別子(identifier)を付して広く公開されるデータを指す。様々なコミュニティからこうしたデータが公開されることによって、ウェブ世界が豊かになり、セマンティックウェブの実現に近づくこととなる。図書館コミュニティにおいても、分類・件名表やレファレンス事例など蓄積してきた専門性の結晶を可視化することの重要性は大きいですが、とりわけ目録情報は、最も基盤的な資産である。そして、目録情報がLinked Dataとなりうるためには、機械可読性は必須の要件である。

前述のように本稿では、それに沿って作られた目録情報の機械可読性という観点から、RDAによって得られるものを検討・評価する。その際、後述する理由からMARC21フォーマットにも焦点をあてることとする。以下、2章で目録規則の機械可読性を整理した後、その視点から3章でRDA自体の、4、5章でRDAを実装するMARC21フォーマットの検討・評価を行う。

2. 目録規則と「機械可読性」

メタデータに関わるルールには、エレメントの定義と格納されるべき値のルールという「意味的側面」と、記述文法やエレメントの配列等の「構文的側面」がある。

2.1. 意味的側面の機械可読性

このうち意味的側面の機械可読性としては、次のような事項があげられよう。

- ・データ化される対象の種類と対象間の関連が明確に整理されていること。例えば、実体関連モデル

July 2011

に沿っているといたことである。

- データ要素が十分に弁別できること(分節化されていること)。現行の目録規則において、「注記」や「その他の形態的細目」といったエレメントには様々な要素のものが必要に応じて入れ込まれる。こうした方式では人間が読めば理解できても、コンピュータには正しい要素識別はできない。
- エレメント間の意味的構造が明確であること。書誌情報の各エレメントは平板に羅列されているわけではなく階層構造(入れ子構造)や相互の対応関係を持っており、それが明確に識別できる必要がある。
- 意味的側面の表現に開放性があること。その意味的構造が機械可読性をもって定義されている必要がある。

2.2. 構文的側面の機械可読性

両者のうちより本質的なのは意味的側面であり、構文的側面の機械可読性としては意味的側面のそれを損なわないことがまず重要である。

- 意味的構造を十分表現できていること。エレメントの弁別やデータ間の対応関係が損なわれることがあってはならない。
- 構文形式に開放性(相互運用性)があること。他のコミュニティでも扱いやすい形式が望ましい。

渡邊：新しい目録規則(RDA)から得られるもの

2.3. 意味的側面と構文的側面の分離

これまでの目録規則は、「区切り記号法」に代表される構文的側面のルールを、意味的側面のルールとともに扱ってきた。ダブリン・コアなど近年のメタデータ規則では、意味的側面のみを守備範囲とし、構文的側面は別途の標準に委ねる(複数の構文形式も許容する)のが一般的である。このことは機械可読性という観点から絶対の要件とまではいえないが、相互運用性のための柔軟性を考えると、望ましい方向である。

3. RDAの機械可読性

RDAの構成の概略を図1に示した。AACR2とは大きく様変わりし、FRBR/FRADモデル(図2)に沿って、全体を「実体の属性」「実体間の関連」の2部立てとし、10セクション37章から成る構成である。その特徴はいくつかあるが³⁾本稿では「機械可読性」という視点に絞って3点を指摘する。

3.1. エレメントの弁別とデータ管理

AACR2と比べ、エレメントが大幅に増強されている。例えば3章「キャリアの記述」は従来の「形態に関する事項」にほぼ相当するが、数量・大きさに加えて、Base material(書写材料), Applied material(筆記材料), Mount(台質), Production method(製造方法)等20個近いエレメントが定義されてい

序論	15p	セクション5：著作～表現形の主要な関連 17章 ガイドライン	10p
セクション1：体現形・個別資料の属性		セクション6：資源と個人・家族・団体の関連 18～22章	計 55p
1章 ガイドライン	15p	セクション7：著作と主題の関連 23章 (未刊)	
2章 体現形・個別資料の識別	133p	セクション8：著作～個別資料どうしの関連 24～28章	計 24p
3章 キャリアの記述	88p	セクション9：個人・家族・団体の間の関連 29～32章	計 13p
4章 入手・アクセス情報	5p	セクション10：概念～場所の間の関連 33～37章 (未刊)	
セクション2：著作・表現形の属性		付録	計 204p
5章 ガイドライン	7p	用語集	44p
6章 著作・表現形の識別	163p	索引	49p
7章 内容の記述	37p		
セクション3：個人・家族・団体の属性			
8章 ガイドライン	11p		
9章 個人の識別	60p		
10章 家族の識別	12p		
11章 団体の識別	75p		
セクション4：概念・物・出来事・場所の識別 12～16章 (場所以外は未刊)	計 17p		

図1 RDAの構成(概略)

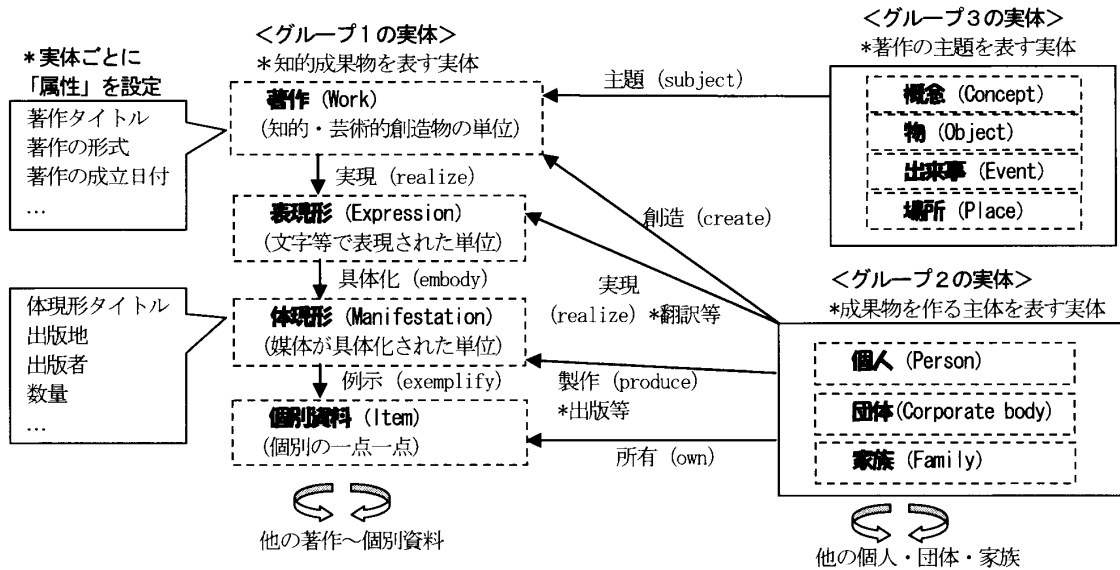


図2 FRBR/FRAD モデルの概略

注)「グループ3の実体」の部分は RDA 刊行時に未完成で、本稿でも扱わないので簡略化して示した。「書誌的実体」のみを示し、FRAD で規定されている典拠作業に関わる実体は省略した。

る。これらの多くは、従来の目録規則では注記や「その他の形態的細目」に記録されていたものであり、独立したエレメントとなることによって、明確に弁別できるようになった。

また、エレメント **Base material** の項では、**paper**, **canvas**, **leather** など24個の語彙リストが示され、ここから選択しての入力が基本となっている。刊行された RDA には含まれていないが、策定段階では「エレメント分析」⁴⁾ という文書が作られ、エレメントごとにその性質や値の準拠すべきスキーマ等が整理された。この際、情報源からの転記によらないフィールドでは、できる限り値の語彙リストを設定することとされた。これもまた、機械可読性の向上に資する仕組みである。

3.2. FRBR/FRAD モデルへの密着

FRBR/FRAD モデル(図2)はいわゆる「書誌的世界」を実体関連分析(E-R 分析)に基づいてリレーショナルな構造にモデル化したものであり、これに密着した規則構造によって、データ化される対象の種類と対象間の関係が明確化されている。

また、さらにいくつかの具体的な点でも、その利点を指摘することができる。まず、AACR2が書誌レコードに付す「標目」「参照」を規定するだけだったのに対して、RDA では著作や個人などを「実体」として管理することによって典拠コントロールに規則上明確な位置づけを与えている。また、FRBR

モデルの最大の特徴である、著作・表現形・体现形・個別資料と順次具体化される構造を取り入れたことで、資料の内容的側面と物理的側面を切り分けて扱うことが可能となった。

また、書誌的世界のリレーショナルな構造をとらえるものとして、実体間の「関連」が実体に付属する「属性」とは別立てで管理されることとなった。この際、一部の関連については「関連指示子(relationship designator)」を設けてさらに詳細な関連の種別を管理しようとしている。付録Iには「資料と個人・家族・団体の間の関連」(セクション6)に対応した約120種の関連指示子一覧があり、これは役割表示をカテゴライズしたものである。付録Jには「著作~個別資料どうしの関連」(セクション8)に対応した約270種の関連指示子一覧があり、翻案・翻訳・階層構造など種々の書誌的関係がカテゴライズされている。さらに付録Kには「個人・家族・団体の間の関連」(セクション9)に対応した約20種の関連指示子一覧があり、個人や団体に関わる各種関係がカテゴライズされている。

FRBR/FRAD モデルへの準拠に伴うこれらの特徴は、いずれもこれに従って作られた書誌情報の機械可読性を向上させるものである。

3.3. 意味的側面と構文的側面の分離

RDA では、AACR2の持っていた構文的側面を規則から排除し、意味的側面に特化した設計が行われ

July 2011

た。従って ISBD 区切り記号やエレメントの配列は規則本体には現れず、付録で ISBD や MARC21 とのマッピングが示されるのみである。

3.4. RDA の機械可読性向上とその「限界」

以上のように、RDA は AACR2 に比べてその機械可読性を大きく向上させていると評価することができる。

ここで留意すべきは、本章で述べてきた評価点を 2 章で述べた目録規則の機械可読性の枠組みに照らすと、3.1~3.2 は 2.1 の「意味的側面の機械可読性」に、3.3 は 2.3 の「意味的側面と構文的側面の分離」にあたり、2.2 の「構文的側面の機械可読性」は評価できていないことである。そもそも意味的側面に特化し、構文的側面を排除した規則なのであるから、当然のことではある。

RDA で扱われていない構文的側面のルールには、エレメントの順序や構造表現の手法(例えば、入れ子構造が繰り返す場合の対応管理)のほか、関連表現の具体や「レコード」「ファイル」の単位などが含まれる。RDA では関連の表現方法として、何らかの識別子による方法、従来の統一標目に相当する「典拠形アクセスポイント」による方法、注記的な記述(引用記述)による方法が挙げられており、どれを選択しどのように運用するかは実装面に委ねら

渡邊：新しい目録規則(RDA)から得られるもの

れている。また、FRBR/FRAD モデルに密着した規則構造はあくまで意味的な構造を示すもので、実際のデータベース設計・レコード構造と直結するものではない。すなわち、「著作」「表現形」「体现形」といった種類ごとにレコードが作成される構造を意味するものではない。

そう考えると、RDA 自体のみからその機械可読性を軽々に評価することはできないように思われる。意味的構造が相当複雑な RDA では、構文的側面、すなわちデータベース実装のありかたをも評価の視野に入れる必要がある⁵⁾。

こうした観点から次節以下では、目録情報の構文的側面を担う⁶⁾ MARC21 フォーマットにおける RDA への対応を検討・評価する。

4. MARC21 フォーマットにおける RDA への対応

図 3 に MARC21 フォーマットの概観を示した⁷⁾。2008年に米国議会図書館(LC)の「書誌コントロールの将来ワーキンググループ」が作成した報告書 *On the Record*⁸⁾では、MARC は「もはやその目的にふさわしくない」と断じて LC に「より柔軟で拡張性の高いメタデータ記録形式を開発すること」を提言している。1960年代の情報環境を背景として作られた MARC フォーマットは、「MARC21」

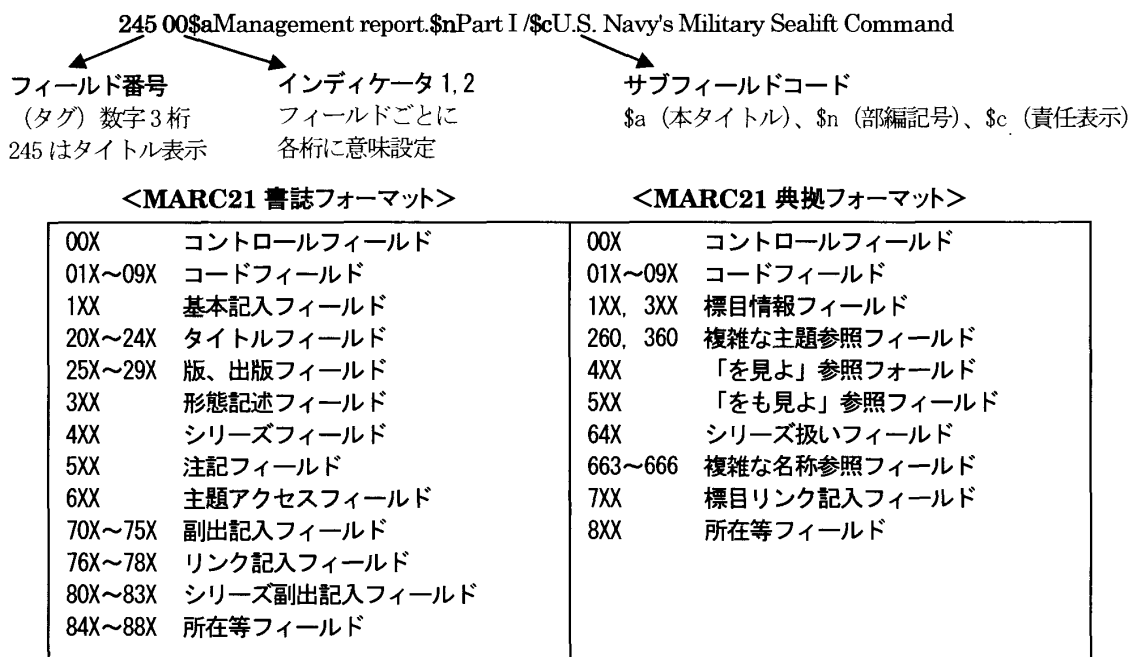


図 3 MARC21 フォーマット概観

注) 他に、所蔵・分類等のフォーマットもあるが省略した。

(1999) の名称となっても根本的には変わっておらず、表現性や相互運用性の観点からこれを批判する声は少なからずある。しかしながら、MARC21 フォーマット(以下、MARC21)を手直しすることで RDA に対応しようというのが少なくとも英語圏における方向性であり、本章ではその対応方式を検証したい。

4.1. MARC21改訂の道程

LCは2002年にMARC21とFRBRとの最初のマッピング結果を発表しており、その後もいくつかの動きがあったが、MARC21改訂に向けての本格的検討が進むのは、米英加の3国立図書館の主導でRDA/MARC Working Groupが発足(2008年3月)してからである⁹⁾。MARC21コミュニティにおける改訂事項の検討は、特定の問題について関係者から提案された文書が会議での検討に付され、採否が決定され公表されるという形をとる。文書には Discussion paper と Proposal という2つの段階がある。

2008年から2011年3月までに出された27本の Discussion paper のうち16本が、同時期に出された37本の Proposal のうち15本が、それぞれ RDA への対応に関係するものである。すべての決着がついているわけではなく、これからの検討・決定が待たれる事項もあるが、以下では現時点までに決まっている範囲での整理検討を行う。

4.2. エレメントの弁別とデータ管理

まず3.1で述べた RDA の特徴点、特に大幅なエレメントの増強への対応についてである。

新たなエレメントに対しては、基本的にフィールドもしくはサブフィールドの追加で対応が行われている。一例をあげれば、従来の資料種別を再編して設けられた属性「内容種別」「メディア種別」「キャリア種別」はそれぞれフィールド336～338として書誌・典拠両フォーマットに追加された。また、個人・団体・家族について新たに設けられた属性「住所」「活動分野」「所属」「性別」等がフィールド370番台としてそれぞれ典拠フォーマットに追加された。一方、「レイアウト」「世代」「フォントサイズ」など RDA 3章(「キャリアの記述」)に規定された属性の一部は、書誌フォーマットのフィールド340(物理媒体)にサブフィールドを新設する形で設定された。

この際、既存フィールド等の大きな異同は極力避けられ、空いたフィールド/サブフィールドを使う形になっている。その結果、RDA の規則構造上は同じ位置づけのエレメントが異なった扱いをされる場合が出てくる。例えば RDA の3章「キャリアの記述」に規定されているエレメントのうち、数量(条項3.4)と大きさ(3.5)はフィールド300(形態的記述)のサブフィールドとして、書写材料(3.6)と筆記材料(3.7)等はフィールド340(物理媒体)のサブフィールドとして定義されているが、音声特性(3.16)はフィールド344に新設する方向で検討が進んでいる。音声特性は下位にサブエレメントが定義されていることと関連しているが、横並びに定義されたエレメント群にフィールドレベルと位置づけられるものとサブフィールドレベルと位置づけられるものが混在していることとなる。結果として、フィールド/サブフィールドの構造は、RDA の意味的構造を必ずしも反映していない。

4.3. FRBR/FRAD モデルへの密着

次に、3.2で述べた FRBR/FRAD モデルに密着した RDA の規則構造が、MARC21でどのように表現されようとしているかを検討する。

2007年に、RDA の編集者(editor)であったデルジー(T. Delsey)によって「RDA のデータベース実装シナリオ」という短い文書が発表されている¹⁰⁾。RDA 策定側がどのような実装を想定していたかがわかる文書である。以下の3つの「シナリオ」が示されている。

シナリオ1

リレーショナル/オブジェクト指向データベース構造：FRBR モデルの書誌的実体ごとにレコードが作成され、相互にリンク付けされた構造

シナリオ2

リンクされた書誌・典拠レコード：書誌レコードと典拠レコードで表現し、相互にリンク付けを行った構造

シナリオ3

「フラットファイル」データベース構造：典拠管理はなされているが、書誌・典拠の明示的なリンクは行われない状態

シナリオ1が RDA の規則構造に最も忠実な方式であるが、必ずしもそれに限るわけではないことが、他ならぬ RDA 策定側から示されている。すなわち、

July 2011

RDA の規則構造と実装上のデータベース設計は直結したものとはとらえられていない。

現行の MARC21 の枠組みを保ちつつ RDA に対処しようとするならば、上記「シナリオ 2」が中心的想定となる。これは、FRBR/FRAD モデルにおける個別資料(「所蔵」フォーマットで表現される)以外の諸実体を「書誌」「典拠」両フォーマットで表現することとなり、機械可読性という観点からは「実体(とその属性)をどのように識別するのか」「実体間の関連(と関連指示子)をどのように表現するのか」の 2 つの問いをクリアしなければならない。

4.3.1. 実体(とその属性)の識別

体現形が書誌レコードとして、個人・家族・団体がフィールド等で識別されながら典拠レコードとして表現される¹¹⁾のは、これまで通りの運用である。残る著作・表現形については、典拠レコードとして表現する方式と、書誌レコード内で体現形と合わせて表現する方式を選択可能とし、著作・表現形の属性を収める多くのフィールドを両フォーマットに同一フィールド番号で並立させている。

4.2 で述べたエレメントの弁別がきちんと行われていれば、同一レコードが複数実体を表現していても、フィールド/サブフィールドごとに仕分けすれば FRBR/FRAD モデルに則したデータ処理が可能である。ただし、従来から存在するフィールド/サブフィールドは原則として変更していないため、MARC21 の構造は RDA の意味的構造を必ずしも反映しない。例えば書誌フォーマットの 3XX フィールドは形態的記述などを収めるところで、多くのフィールドは RDA の 3 章「キャリアの記述」に規定された体現形の諸属性と対応しているが、その中に RDA では表現形の属性として 7 章「内容の記述」に規定されているフィールド 306(再生時間)が混じっている。AACR2 までは形態事項に含まれていたことを反映しているが¹²⁾、あえてフィールド番号を変更することはせず、データを処理するアプリケーションに委ねた格好となっている。

なお、FRBR/FRAD モデルにおける「グループ 1」の各実体ごとにレコードを作成し、上記「シナリオ 1」の表現も可能とする改訂案も提案され、現在議論が続けられている¹³⁾。フィールド 883(実体種別)を新設して、各レコードが表現している実体を明示する仕組みである。

渡邊：新しい目録規則(RDA)から得られるもの

4.3.2. 関連(と関連指示子)の表現

実体間の関連は実体識別の方式に依存し、①書誌レコード同士、②書誌レコードー典拠レコード間、③典拠レコード同士、④同一レコード内に複数実体を表現、の 4 種の表現方式がありえる。いずれも既存のレコード間リンク、フィールド間リンクの仕組みを用いて表現される。

3.2 で述べた関連指示子のうち、「資料と個人・家族・団体の間の関連」(セクション 6)に対応した付録 I(役割表示をカテゴライズ)は、既存の MARC Relator Code を手直しすることで表現される。「著作～個別資料どうしの関連」(セクション 8)に対応した付録 J(種々の書誌的關係)と「個人・家族・団体の間の関連」(セクション 9)に対応した付録 K は、書誌フォーマットの 7XX フィールド(副出記入等)、典拠フォーマットの 4XX～5XX フィールド(参照)にサブフィールドコードを新設して表現される。いずれも既存の枠内での、一定の拡張による表現といえる。

5. MARC21 フォーマットによる RDA 表現の評価

MARC21 の改訂作業はいまだ進行中であるが、前章で整理した現時点での RDA 対応について、目録情報の機械可読性という観点から評価を行いたい。

筆者は 2001 年の時点で、メタデータの構文表現方式としての MARC フォーマットを検討し、フィールド番号やサブフィールドコードによる分節化を行う MARC は一種の「マークアップ言語」であるが、SGML 等の標準的なマークアップ言語に比べて表現力に問題があると指摘した¹⁴⁾。そして、目録規則が大きく変わることがあれば、それを契機として XML ベースの RDF(Resource Description Framework)等に移行していくのではとの予測も述べた。この予測は外れ、RDA という新たな目録規則の誕生に際して、既存の MARC フォーマットの手直しによる実装がめざされているわけである。

前章で整理した通り、エレメントの弁別と FRBR/FRAD モデルの表現という両面において、無理はある(合理的な表現とは言い難いところもある)が、RDA の意味的構造を表現することは一応果たされていると評価する。MARC21 による書誌・典拠レコードをエレメントごとに解体し、再構成すれば、FRBR/FRAD モデルに沿った RDA の意味的構造

を組み立て直すことが可能であり、例えば FRBR/FRAD モデルに沿った新たな形の OPAC を実現することもできるだろう。ただ、それには「組み立て直す」機構(プログラム等)を別途要する点に留意が必要である。それは即ち、プログラムの作り手(=人間)による RDA の意味的構造の理解が必須だということであり、この点において「意味的構造が機械可読」とはいえない状況にある。このことは、図書館コミュニティ内での目録情報の利用(新たなサービス構築など)にはそれほどの不都合とはならないが、外部コミュニティでの利用を想定したメタデータの開放性という文脈では大きな障壁になりうる。当該メタデータの構造に関わる情報が機械可読性を持った形で得られないことは、データ利用にかかるコストを上げるといわざるをえない。

ところで、MARC21による実装への対応とは別に、RDAとメタデータ標準との接合をめざす「DCMI/RDA タスクグループ」¹⁵⁾の活動がある。これは、2007年にダブリンコアを維持管理する DCMI (Dublin Core Metadata Initiative) 関係者と RDA 関係者が会して行われた「データモデル会議」を契機として作られた、両コミュニティの共同作業グループである。

その中心的な活動の一つに、RDA をセマンティックウェブで利用可能な資産とすることをめざす「RDA 語彙プロジェクト」(RDA Vocabulary Project)¹⁶⁾がある。RDA に現れる「語彙」に識別子(URI)を付し、RDF 表現によるメタデータレジストリ登録を行うものである。本プロジェクトで扱う「語彙」は、実体を表す「クラス(Class)」, 属性や関連(関連指示子を含む)を表す「エレメント(Element)」, 3.1で述べた語彙リストに現れる語やコード類などを表す「コンセプト(Concept)」の3種から成る。エレメントについては、エレメント間の上下関係やクラス、コンセプトとの関係も示されるので、ある属性がどの実体に関わるもので、他のエレメントとどのような関係を持ち、どのような値を取りうるのか、といったことが表現される。すなわち、MARC21による実装対応では十分果たされていない、意味的構造の機械可読性という部分を担うものとして重要である。

6. おわりに

以上、これからの目録規則には不可欠の視点であ

る「目録情報の機械可読性」という観点から、RDA 自体と MARC21フォーマットによる実装表現を検討・評価した。簡単にまとめれば、① RDA は意味的側面において大きな前進が見られる、②ただし、実装面(構文的側面)も視野に入れないと真の機械可読性は評価できない、③ MARC21における RDA 表現は、RDA の意味的側面における向上を生かす、一定の表現力を持っている、④ただし、意味的構造そのものの機械可読性に不十分さがあり、RDA 語彙プロジェクト等による補完が求められる、ということである。

本稿における考察をやや異なる視点から眺めると、RDA は機械可読性を重視して設計されたと評価できるが、目録情報の機械可読性の十分な発揮には MARC21や「語彙プロジェクト」といった外部の機構に依存する部分も大きいということである。このことは、FRBR/FRAD モデルに沿った意味的構造を別途管理する外部機構が必要ならば、RDA のように規則自体の構成を概念モデルに密着させる必要があるのか、など目録規則のあるべき構成・構造に関する問題意識も派生させるが、これは今後の課題としたい。

注

- 1) 冊子体も2011年に入って刊行されたが、頒布の中心は有料のウェブ版である。
RDA Toolkit. <<http://www.rdatoolkit.org/>> [引用日: 2011-04-29]
- 2) RDA 刊行後の文献としては、次のものがある。
古川肇「書誌レコードおよび典拠レコードに関する規則の成立: RDA の完成」『資料組織化研究-e』59, 2010.12, p.13-32. <<http://ojs.info.gsecc.osaka-cu.ac.jp/TS/index.php/TS/article/viewFile/37/75>> [引用日: 2011-04-29]
- 3) 渡邊隆弘「目録法の再構築をめざして」『図書館雑誌』103(6), 2009.6, p.376-379.
松井純子「RDA 改訂に見る FRBR の具体化」『図書館界』62(2), 2010.7, p.182-192.
- 4) RDA Element Analysis, 2009 (Last update) <<http://www.rda-jsc.org/docs/5rda-elementanalysisrev3.pdf>> [引用日: 2011-04-29]
- 5) 目録情報の構文的側面には、OPAC 表示等の出力形式も含まれるが、本稿ではデータ管理のためのデータベース実装のみをとりあげる。
- 6) MARC フォーマットでは、目録規則で規定されたエレメントよりも細かくフィールド・サブフィールドの弁別がなされているなど、必ずしも構文的側面にとどまるものでは

July 2011

ないが、本稿では RDA を十分に表現できるかどうかという視点からのみ取り扱う。

7) MARC21の維持は米国議会図書館(LC)で行われている。

MARC Standards.<<http://www.loc.gov/marc/>>[引用日: 2011-04-29]

8) Working Group on the Future of Bibliographic Control, Library of Congress. *On the Record: Report of The Library of Congress Working Group on the Future of Bibliographic Control*. 2008, 44p. <<http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf>> [引用日: 2011-04-29]

9) 2011年1月現在の状況の概観が、以下で把握できる。

RDA in MARC, January 2011.<<http://www.loc.gov/marc/RDAinMARC29.html>> [引用日: 2011-04-29]

なお、RDA の付録 D, E として、RDA のエレメントと MARC21(書誌・典拠フォーマット)とのマッピング表がある。

10) Tom Delsey. *RDA Database Implementation Scenarios*. <<http://www.rda-jsc.org/docs/5editor2rev.pdf>> [引用日: 2011-04-29] なお、現在閲覧できるのは2009年の改訂版だが、本質は変わっていない。

11) 統一標目を表すフィールドとして、100があれば「個人」、110もしくは111があれば「団体」(111は会議を指す)である。ただし100のインディケータ1が“3”のものは「家族

情報組織化研究グループ月例研究会報告

を表す。

12) 従来は慣行的に「数量」の一部とみなされてきたが、ページ数が判型や活字の大きさなど表現形レベルの条件に左右されるのに対し、再生時間は表現形レベルの変更がない限り変わらない情報である。

13) MARC Discussion Paper No. 2011-DP03: Identifying Work, Expression, and Manifestation Records in the MARC21 Bibliographic, Authority, and Holdings Formats. <<http://www.loc.gov/marc/marbi/2011/2011-dp03.html>> [引用日: 2010-04-29]

14) 渡邊隆弘「図書館目録とメタデータ」『図書館界』53(2), 2001.7, p.126-133.

15) DCMI/RDA Task Group Wiki. <<http://www.dublincore.org/dcmirdataskgroup/>> [引用日: 2010-04-29]

16) *The RDA Vocabularies* <<http://metadataregistry.org/rdabrowse.htm>> [引用日: 2010-04-29]

Diane Hillmann et al. "RDA Vocabularies: Process, Outcome, Use" *D-Lib Magazine*, 16(1/2), 2010. <<http://www.dlib.org/dlib/january10/hillmann/01hillmann.html>> [引用日: 2010-04-29]

本研究は、科研費基盤研究(C) 課題番号22500223「情報環境の変化に適切に対応する目録規則の研究」の成果である。

情報組織化研究グループ月例研究会報告

◎2011年3月月例研究会

日時: 3月12日(土) 14:30~17:00

会場: 大阪市立弁天町市民学習センター

発表者: 大向一輝氏(国立情報学研究所)

テーマ: 学術情報サービスのメタデータ・デザイン

出席: 36名

国立情報学研究所(NII)が運営する論文情報ナビゲータ CiNii におけるメタデータの設計・運用について発表された。

1. CiNii の展開

- CiNii は、NII で構築された NII-ELS(電子図書館)・CJP(引用文献索引)に NDL, JST, 機関リポジトリ由来の情報を加え、約1,300万件を検索可能な国内最大規模の論文検索サービスである。NII-ELS の約350万件等は、本文情報も提供している。
- 複数のソースから論文情報が集まるため、書誌同

定・統合作業が重要である。自動・手動を組み合わせた処理を行っている。

- 2007年4月の Google 連携, 2009年4月のインタフェース一新等により、アクセス数は大きく伸び、継続的な右肩上がり傾向にある。学術専門家だけでなく一般の人々にも利用が広がっており、それに対応したインタフェース改善を行ってきた。

2. セマンティックウェブ対応とメタデータ

- 極めて多くの論文が生産される分野でのマイニングの必要性や厳密な研究評価の要請等により、今日の論文情報は人間だけでなく「コンピュータ」にも扱えるものであることが求められる。
- すなわち、ウェブ情報の構造化・組織化を目指す「セマンティックウェブ」への対応である。より具体的には、様々な操作対象に一意な ID を与えることと、構造化されたフォーマットで表現することが必要である。
- 学術情報分野は、情報の構造化を生業とする職業