

当日配布分に、若干の追補・訂正を加えております (12/20)

新時代の目録規則 (続) (まとめ)

—新しい目録規則から得られるもの：機械可読性の視点から

渡邊隆弘 (帝塚山学院大学)

watanabe@hcs.tezuka-gu.ac.jp

1. 今年度の「情報組織化研究グループ」
2. 新しい目録規則から得られるもの：機械可読性の視点から
 2. 1. 目録規則の機械可読性：その背景
 2. 2. RDA の機械可読性
 2. 3. MARC21 と RDA
 2. 4. 実装に向けたその他の動き
 2. 5. 展望？

1. 今年度の「情報組織化研究グループ」¹

●昨年度の研究大会発表

松井純子「RDA 改訂に見る FRBR の具体化」²

- ・ RDA 草案中の FRBR/FRAD 関連記述から
- ・ 従来の目録規則との違いを具体的に整理

●今年度の研究テーマ： 「新時代の目録規則 (続)」

2009 年度の研究テーマとして「新時代の目録規則」を表明した文章に、「10 年以上にわたる見直し作業が一応の結実をみた年として、本 2009 年は記録・記憶に残るでしょう」と記しました。2009 年は確かに、パリ原則 (1961) に変わる「国際目録原則覚書」の刊行 (2 月)、FRAD (典拠レコードの機能要件) の刊行 (5 月)、図書館目録や書誌コントロールの今後をめぐる議論はここ数年非常に盛んで、FRBRoo (オブジェクト指向版 FRBR) の Ver1.0 発表 (6 月) と、目録規則関係の大きな出来事が続きました。

しかしながら、「結実」の大きな一つになるはずだった RDA (Resource Description and Access : AACR2 の後継規則) はスケジュール通り刊行されず、2010 年 6 月完成予定にずれこんでいます。また、FRSAD (主題典拠データの機能要件) や ISBD (国際標準書誌記述) 改訂など、IFLA 関係でもなお動きがあります。VMF (Vocabulary Mapping Framework) などのように様々なメタデータスキーマ間の相互運用性という視点から目録規則をとらえることも必要でしょう。

そして、わが国の標準目録規則である NCR (日本目録規則) も抜本改定に向かっていくことは間違いないく、非基本記入方式や書誌構造 (書誌階層) など独自の特徴を含めて、今後のあり方が問われてきます。

当グループでは本年も引き続き「新時代の目録規則」を年間テーマとし、情報環境の変化の中での目録規則の今後を考えていきたいと思えます。

¹ グループ活動記録等は <http://www.tezuka-gu.ac.jp/public/seiken/> に蓄積

² 松井純子「RDA 改訂に見る FRBR の具体化」『図書館界』62(2), 2010.7. pp.182-192

●科学研究費

基盤研究 (C) 「情報環境の変化に適切に対応する目録規則の在り方に関する研究」

課題番号 22500223

研究代表者：渡邊隆弘 研究分担者：吉田暁史、田窪直規、松井純子、研谷紀夫

●グループ月例研究会

2010.4.24 統制形データの概念モデル (FRAD と FRSAD) の概要について

和中幹雄 (同志社大学)

2010.5.15 情報組織化をめぐる最近の動向

渡邊隆弘 (帝塚山学院大学)

2010.6.19 インターネットへ対応するための図書館・システム：

Aleph+Primo の導入について

入江伸 (慶應義塾大学)

2010.7.17 ネットワーク文化情報資源で活用する人名典拠情報構築に関する研究とその成果

研谷紀夫 (東京大学)

2010.9.18 学術情報流通と OAI-PMH

横谷弘美 (大手前大学図書館)

2010.10.23 中国と韓国における目録をめぐる動向

小島浩之 (東京大学)、高橋菜奈子 (国立情報学研究所)

2010.11.13 RDA の完成とこれからの目録

古川肇 (近畿大学)

2010.12.18 まとめ (本日)

渡邊隆弘 (帝塚山学院大学)

(予) 2011.1.8 東アジアの目録規則 (拡大月例研究会)

李常慶 (北京大学)、崔錫斗 (漢城大学)、渡邊隆弘

●日本図書館研究会研究大会発表 (予定)

・研究大会： 2/19 (土) ~20 (日) 於：相愛大学

・「グループ研究発表」： 2/19 午前 (発表 20 分・質疑 10 分)

・渡邊隆弘「新しい目録規則から得られるもの：機械可読性の視点から」

概要 (11 月時点) は別紙 (付 1)

『図書館界』63(2), 2011.7 に論文執筆予定

2. 新しい目録規則から得られるもの：機械可読性の視点から

2. 1. 目録規則の機械可読性：その背景

●目録規則改訂の動向

- ・概念モデル
 - 1997 FRBR（書誌レコードの機能要件）
 - 2009.5 FRAD（典拠データの機能要件）
 - 2010.7 FR SAD（主題典拠データの機能要件）
 - 2009.6 FRBRoo Ver.1.0（オブジェクト指向モデル版）
- ・国際原則・標準
 - 2009.2 国際目録原則覚書（パリ原則の後継）
 - 2011? ISBD 統合版（World Wide Review が終わり最終段階）
- ・目録規則
 - 2009.1 REICAT（イタリア目録規則）
 - 2010.6 RDA（Resource Description and Access）AACR2 の後継
 - 2010.9 『日本目録規則』改訂に向けて

◆新しい目録規則に求められる要件

- パリ原則（1961）から半世紀
 - 対象資料の変化、多様化への対応
 - 情報組織化環境の変化
- インターネット普及の15年
 - OPAC 高度化（次世代 OPAC）への対応
 - メタデータの開放性（相互運用性）
 - メタデータの付加価値性

●目録情報の外部開放

これだけではないが、非常に重要な方向性

- 2009.5 LC Authorities & Vocabularies（当面、件名データ等）
- 2009.12 OCLC WorldCat Basic API（一般公開）
- 2010.4 ハンガリー国立図書館（書誌・典拠データ）
- 2010.4 ドイツ国立図書館 Linked Data Service（名称・主題典拠データ）
- 2010.11 JISC OpenBibliography（BNB：英国全国書誌）

- 2010.5 W3C Library Linked Data Incubator Group³
- 2010.11 JISC “Open Bibliographic Data Guide”⁴

³ http://www.w3.org/2005/Incubator/lld/wiki/Main_Page

⁴ <http://obd.jisc.ac.uk/>

2010.6 Web NDLSH (国立国会図書館)
典拠データ、書誌データへ拡大？

◆セマンティックウェブと Linked Data

- コンピュータに理解可能な、相互運用性を備えたデータ開放
- URI の確実な付与
- 標準的な構文形式 (RDF/XML など)
- エレメントの明確な弁別
- スキーマの明示 (エレメントの定義、ルール)

◆要するに、目録情報の「機械可読性」

- データのもつ意味的内容が機械可読性をもつこと
- 「次世代 OPAC」など他の動きにおいても重要

2. 2. RDA の機械可読性

●RDA の完成

- ・ 2010.6 公開 (RDA Toolkit)
- ・ 現在はウェブ版のみで有料。間もなく冊子体が刊行される予定
- ・ 2008.11 の「全体草案」と、基本的には同じ
(用語や章題の変更が若干)
- ・ RDA の全体像は古川肇氏の研究⁵に詳しい

●RDA の特徴⁶

- (1)FRBR/FRAD モデルへの密着
「実体の属性」「実体の関連」の2部構造からなる規則構成
- (2)典拠コントロールの明確な位置づけ
標目・参照だけでなく、「実体」として管理
- (3)資料の内容的側面と物理的側面の整理
内容種別とメディア・キャリア種別による資料種別表現
内容的側面 (著作、表現形) を現行より重視
- (4)「関連」の重視
属性とは別立てで、「関連」を管理
「関連指示子 (relationship designator)」による関連の種類管理

⁵ 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. <<http://ojs.info.gsec.osaka-cu.ac.jp/TS/index.php/TS/issue/view/7>>
古川氏には RDA の策定過程を丹念に追った一連の論考もあるが、書誌事項は省略させていただく (上記論文の注を参照されたい)。

⁶ 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

(5)機械可読性の向上

エレメントの増強（従来の注記や MARC フォーマットから、など）
 語彙リストからの選択によるエレメント
 策定段階で「エレメント分析」⁷

(6)意味的側面と構文的側面の分離

構文的側面は扱わない（付録に示すのみ）

◆機械可読性

上記(5)以外の特徴も、機械可読性に関わるもの

FRBR モデル（実体関連モデル）に沿った構造

実体の異なりを明確に区別し、「関連」を設定（リンク付け）

●真の機械可読性：実装による影響

・意味的側面と構文的側面の分離

意味的側面：エレメントの定義と格納されるべき値のルール
 構文的側面：記述文法やエレメントの配列等

・RDA で扱われないこと

エレメントの順序
 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

「関連」表現の具体

複数の方式： 識別子、典拠形アクセスポイント、注記的記述
 方式選択や実装方法を決定する必要

「レコード」「ファイル」の単位

規則構造と目録データベース設計とは直結しない

FRBR に密着したレコード構造を意味しない

・RDA の意味的側面の複雑さ

単純・平板なモデルではなく、実体間の縦横の関連が絡み合った構造
 構文的側面の選択が意味的側面を制約しかねない

・AACR2 との継続性

「コアエレメント」（必須エレメント）は比較的限定

新しいエレメントは、ほとんどオプショナル

個々の条項規定は、AACR2 からの継承が多い

◆実装面も見ないと、本当のところはわからない

構文的側面： データベース実装、OPAC 表示等の出力形式

出力形式はともかく、データベース実装は重要
 （意味的側面を制約することもある）

実際運用： 実際にどのようなデータが作られるのか

⁷ *RDA Element Analysis* <<http://www.rda-jsc.org/docs/5rda-elementanalysisrev3.pdf>>

2. 3. MARC21 と RDA

●まだ MARC なのか？

- ・ *On the Record* (2008.1) ⁸

3.1.1 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

- 3.1.1.1 LC: Z39.2 (MARC) がもはやその目的にふさわしくないことを認識し、図書館が関わりを持つデータのすべての範囲を表現でき、図書館間および関連コミュニティ間のデータ交換を促進することができる書誌情報の記録形式を特定し、実装するため、図書館および他の関係コミュニティと作業を行うこと。

●それでも MARC21 を継続する

- ・「MARC は、AACR とのかなりゆるやかな関係と、目録法の世界に対するその構造的ドキュメンテーションの重要性を勘案すると、「データ構造標準」と「技術的交換標準」の両面の機能を果たしているといえよう (MARC 自身のドキュメントでは後者が第一に強調されているが)。... 図書館員、とりわけカタログラーは通常、AACR2 の条項番号よりも MARC タグを用いて会話しているが、それは書誌レコードに対する彼らの構造的概念が典型的な MARC タグ形式表示に基づいているからだ。」⁹
- ・ 2002～ MARC21 と FRBR に関する検討¹⁰
特に、著作・表現形の扱い (書誌・典拠)
- ・ 2006 JSC (AACR 合同改訂委員会) による、最初のマッピング¹¹
2007 にかけて、JSC 側でマッピング検討
- ・ 2008.3 RDA/MARC ワーキンググループ発足¹²
主導は BL, LC, LAC
さまざまな Discussion paper, Proposal を提出
却下・修正される提案も少なくない
- ・ 大きな修正は 2009 までにほぼなされたが、2010 に入っても若干

◆現時点での主な修正点¹³

<属性>

- ・ 内容種別、メディア種別、キャリア種別のフィールド新設 (336～338)
- ・ 個人・団体・家族に関わる属性のフィールドを典拠に新設 (370 番台)
- ・ 著作・表現形に関わる属性のフィールドを書誌・典拠に新設 (380 番台)

⁸ 国立国会図書館の日本語訳に拠った。

<http://www.ndl.go.jp/jp/library/data/kokusai.html#01>

⁹ 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>>

¹⁰ *Functional Analysis of the MARC 21 Bibliographic and Holdings Formats* (2002)
<<http://www.loc.gov/marc/marc-functional-analysis/functional-analysis.html>>

Using MARC21 with FRBR: Record Configurations (2005)

<<http://www.loc.gov/marc/marbi/2005/2005-report02.pdf>>

¹¹ *RDA and MARC21* <<http://www.loc.gov/marc/marbi/2007/5chair12.pdf>>

¹² <http://www.rda-jsc.org/rdamarcwg.html>

¹³ *RDA in MARC (May 2010)* <<http://www.loc.gov/marc/RDAinMARC29.html>>

<関連>

- ・グループ1とグループ2の間の関連指示子（RDAの付録I：役割表示を表す）は、既にある Relator code を使用（手直し？）
- ・グループ1同士の間での関連指示子（RDAの付録J：書誌的関連を表す）は、書誌（7XX）及び典拠（4XX、5XX）にサブフィールド\$iを新設（コードは？）

◆著作・表現形をどう表すか

- ・いくつかの提案
 - 典拠レコードで表現
 - 書誌フォーマットにフィールドを新設して、独立レコードで表現
- ・ドラスティックな案は却下
 - 著作・表現形に関わるフィールドは、書誌・典拠両フォーマットに設営

2. 4. 実装に向けたその他の動き

●メタデータ標準との接合

- ・2007.5 RDA関係者とDCMI（ダブリンコア）関係者で「データモデル会議」（Data Model Meeting）
- ・2007.7 DCMI/RDA Task Group 結成¹⁴
- ・RDA語彙プロジェクト（RDA Vocabulary Project）¹⁵
 - セマンティックウェブで利用可能な資産に「語彙」にURI付与、RDFによる表現（メタデータレジストリ）
 - Class list 実体
 - Element list 属性、関連（関連指示子はこちら）
 - Concept list 各エレメントに設定された値リストやコード

●実装テスト

- ・LC等によるテスト（パートナー館との協同）¹⁶
 - 適用細則等の各種文書も
 - テストレコードのダウンロードも

¹⁴ <http://www.dublincore.org/dcmirdataskgroup/>¹⁵ 注9 文献参照*The RDA Vocabularies* <<http://metadataregistry.org/rdabrowse.htm>>¹⁶ *Testing Resource Description and Access (RDA)*<<http://www.loc.gov/bibliographic-future/rda/>>

2. 5. 展望？

●検証したいこと

- ・MARCにおけるRDAの実装（とテスト）でFRBR/FRADモデルは過不足なく表現できているか？

RDAはFRBR/FRADに忠実 → RDAを表現できているか？と同じ

- （・RDFによる表現と比べてどうか？）

◆FRBR/FRAD（=RDAの規則構造）とMARCの構造

そのままのレコード構造である必然性はない

エレメント単位で位置づけが明確であれば、意味的構造は伝わる

●考察したいこと

- ・機械可読性の観点から、従来の書誌データとの異なりを整理

RDAに沿ったデータの利用可能性（と限界？）

- ・「日本目録規則」への意識も

付 1. 11 月時点で作成した「要旨」(下線は今回付した)

FRBR (「書誌レコードの機能要件」) と国際目録原則をベースとして完成された RDA (Resource Description and Access) について、その実装がどのように OPAC や書誌情報流通を変えていく可能性があるのかを、特徴点の一つである「機械可読性の向上」という視点から考察する。

1990 年代にはじまった目録規則の抜本的見直しの動きは、パリ原則に代わる国際目録原則の刊行 (2009.2)、AACR2 の後継である RDA の刊行 (2010.6) によって、大きな節目を迎えたといえる。また、NCR (日本目録規則) についても日本図書館協会目録委員会によって改訂に向けての方針が明らかにされ、新たな段階に入ろうとしている。

当研究グループでは昨年度から「新時代の目録規則」をメインテーマとし、これらの動向を追ってきた。昨年度のグループ発表 (松井純子「RDA 改訂に見る FRBR の具体化：新時代の目録規則を考える」) では、RDA の最大の特徴点である FRBR 及び FRAD (「典拠データの機能要件」) モデルへの密着を分析し、従来の目録規則との相違を具体的に整理した。この発表は全体草案 (2008.11) に依拠しており、RDA の完成はその後になったが、全体草案と完成版との間にそれほど大きな相違はないため、完成版に対しても適用できる考察になっていると考えている。

これからの書誌データや目録システムを考えるうえで、「機械可読性」は重要な視点である。 流通する情報量が増大の一途をたどるインターネット時代において、人間の目と判断で構造や文脈が理解できるだけでなく、コンピュータがシステムティックに処理できる構造のデータが求められている。それは、「次世代 OPAC」など図書館コミュニティの中でのシステム構築・サービス構築にも重要であるし、図書館コミュニティが作成したデータがより広い世界で存在感を発揮するためにも不可欠である。データの持つ意味的構造が機械可読であることは、いわゆるセマンティックウェブの根幹となる考え方でもある。

その点、FRBR/FRAD モデルに則った RDA の構造は、こうした可能性に向かうものと前向きに捉えることができる。 著作・表現形・体现形・個別資料・個人・団体・家族等の諸実体の「属性」と実体間の「関連」によって表される構造、特にそして資料の物理的側面 (体现形・個別資料) と内容的側面 (著作・表現形) を整理して設定された段階的に具体化されていく構造は、従来の目録規則によるデータにはない可能性を秘めている。また、エレメント (データ要素) の大幅な増強や語彙リストの多用なども機械可読性を増し、より柔軟な利用可能性につながるものである。

しかしながら、RDA の規則構造や条文の検討には、従来の目録規則に対する検討の態度では対処しきれない問題がある。 それは、RDA が今日のメタデータ規則の世界の潮流に従って、セマンティクス (意味的側面：データ要素の定義と格納されるべき値のルール) のみを扱い、シンタックス (構文的側面：記述文法やデータ要素の配列等) は扱わないとしていることである。すなわち、RDA に準拠した書誌データは、従来の ISBD や MARC など図書館伝統の方式でもセマンティックウェブの基盤をなす RDF などでも表現 (入出力) しようし、FRBR/FRAD モデルに密着した規則は直ちにその通りのレコード構造を意味するものではない。本来、シンタックスの選択はセマンティクスに影響を与えないはずであるが、単純・平板なモデルではなく、実体間の縦横の関連がからみあった複雑な意味的構造をもっている RDA においては、シンタックス (実装方式) の選択がセマンティクスを制約しかねない側面を持っているように思われる。

本発表では、RDA に規定されたセマンティクスに加えて、MARC21 フォーマットの改訂やレコード作成テストなど現在進みつつある実装の動きも視野に入れた検討を行い、これから本格的に作られる RDA 準拠の書誌データが「機械可読性」の視点から従来のデータとどれほど異なりがあるか、活用可能性の高いデータを生成するためにはシンタックスも含めて運用上どのような点が必要になるのか、等を考察したい。