

BIBFRAME 2.0 の意義を問い直す

谷口 祥一 (慶應義塾大学文学部)
taniguchi@z2.keio.jp

1. はじめに

米国議会図書館が主導する新たなメタデータスキーマの策定は BIBFRAME と名付けられ、2016年4月にバージョン 2.0 案が公開されている¹⁾。これまでのバージョン 1.0 からかなり修正された部分も見受けられるが、基本的な考え方や方針など (たとえば Linked Data の原則に従って RDF を採用) は変更されていない。

こうした段階で、本研究は BIBFRAME の意義、すなわち目的の設定とそれを実現する規定内容、目的設定の妥当性等を再検討することを目的としている。同スキーマは、a) 図書館目録のメタデータである RDA メタデータを作成し交換・共有するためのスキーマ、かつ b) RDA メタデータを超えて出版流通業界作成のメタデータや利用者作成のメタデータなど、情報資源に対する多様な記述メタデータの受け皿、すなわち多様なメタデータからの変換先スキーマそして交換・共有スキーマとして用いられることを意図しており、この両者の観点から再検討を試みる。

なお、発表者は以前に BIBFRAME 1.0 について、主に RDA 作成用メタデータスキーマの観点からその問題点等を論じた²⁾。

2. RDA メタデータ作成用スキーマとしての BIBFRAME

RDA 運営委員会は、RDA にそのまま対応させた RDF のクラスとプロパティを RDA Registry に登録し公開している³⁾。RDA メタデータ作成用とする意図の有無については定かではないが、Linked Data の考え方に沿った登録内容とされている。これを BIBFRAME のクラスおよびプ

ロパティと対比し、BIBFRAME へのマッピングを検討する。適切にマッピングできなければ、RDA メタデータを十全に表現できないことになる。

2. 1 主要な RDF クラスとその関連

RDA Registry では、情報資源を表すクラスとして FRBR/RDA の実体「著作」、「表現形」、「体現形」、「個別資料」に対応するクラス rdac:C10001 (<http://rdaregistry.info/Elements/c/C10001> の略記。以下、略記形式を使用) (ラベル "Work")、rdac:C10006 ("Expression")、rdac:C10007 ("Manifestation")、そして rdac:C10003 ("Item") が登録されている。

他方、BIBFRAME 2.0 ではクラス Work (著作; <http://id.loc.gov/ontologies/bibframe/Work>)、Instance (インスタンス)、Item (アイテム; 個別資料) が定義されている (図 1)。著作は FRBR/RDA の著作に、インスタンスは体現形に、アイテムは個別資料にそれぞれ対応する。

従来の BIBFRAME 1.0 では、著作とインスタンスのみ定義され、アイテムはなく、代わって他のクラス Annotation (アノテーション) のサブクラス HeldMaterial (所蔵) が存在した。これにより FRBR/RDA における個別資料を表現することはできたが、直截な対応づけではなかった。BIBFRAME 2.0 はこの点で RDA スキーマの観点からは改善されているといえよう。

次に、FRBR/RDA の「表現形」の BIBFRAME における扱いであるが、以前のバージョン同様、2.0 においても表現形に対応するクラスはない。

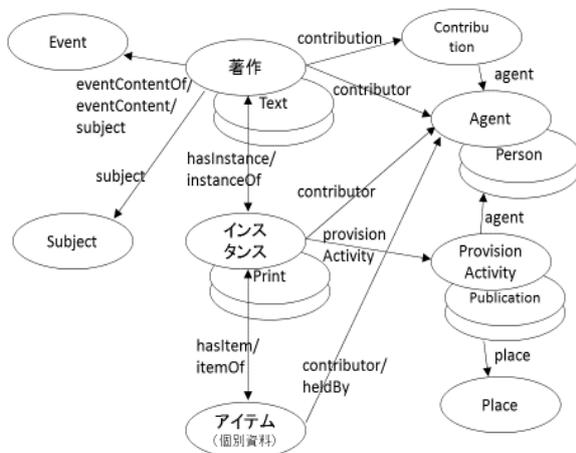


図 1 主要な RDF クラスとその関連

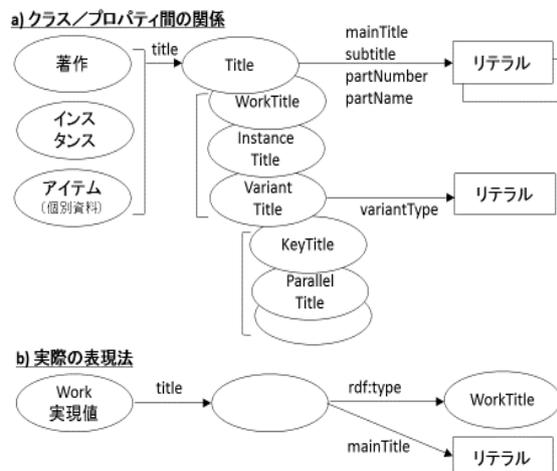


図 2 タイトルにかかわるクラスとプロパティ

他方、付随するドキュメント等によれば、クラス「著作」が表現形をも包含したクラスであるとしている。つまり、クラス「著作」を用いて表現形にかかわる事項を記述することになる。実際には、著作に対応するプロパティに加えて、表現形に相当するプロパティを用いれば、表現形を記述したものとなる（ただし、著作等のプロパティ設定に伴う問題については後述する）。逆にいえば、対象とするリソースが著作か表現形かを明示的に区別するプロパティはない。それゆえ、対象とするリソースすなわち単一 URI を、ある場合には著作として扱い、他の場合には表現形として扱うことが結果的に可能となる。

上記に合わせて、著作とその表現形との関連を表すプロパティ `hasExpression` と `expressionOf` は、いずれもその定義域（主語）と値域（目的語）がクラス「著作」となる。それに対して、RDA Registry では、プロパティ `rdaw:P10078("has expression of work")` は定義域が著作、値域が表現形となり、逆方向のプロパティ `rdac:P20231("has work expressed")` は定義域が表現形、値域が著作とされ、明確に著作と表現形が区別されている。加えて、表現形間の関連を表すプロパティ、たとえば `rdac:P20141("is translation of")`、`rdac:P20171("is translated as")` などは、定義域、値域とも表現形となる。このような表現形間の関連は BIBFRAME では適切に表現しがたい。

FRBR 第 2 グループ実体に相当する RDA クラスとして `rdac:C10004("Person")`、`rdac:C10005("Corporate body")`、`rdac:C10008("Family")` とともに、これらの上位クラス `rdac:C10002("Agent")` を登録している。BIBFRAME もクラス `Agent` とそのサブクラス `Person`、`Family`、`Organization`（組織・団体）、`Jurisdiction`（行政区分）、`Meeting`（会議）を登録している。両者において大きな齟齬はないといえよう。（第 1 グループ実体との関連については後述する。）

なお、こうした主要クラスとその関連が構成するものは、それぞれのスキーマの概念モデルに相当する。

2. 2 主要な RDF クラスに対するプロパティ

RDA Registry には RDA エlement（サブエlement、エlementサブタイプを含む）に対応するプロパティが網羅的に登録されている。一方、BIBFRAME も幅広くプロパティを設定しているが、その採択基準については不明である。

RDA から BIBFRAME プロパティへのマッピングを行ったときに、一対一と一対多の対応となるものは問題ないが、多対一と多対多の関係にあるものが問題となる。後者の場合 BIBFRAME に値をマッピング後、元の RDA プロパティとそ

の値に戻すことができない。以下、問題となるものを例示的に取り上げる。

1) タイトルにかかわるプロパティ：RDA Registry には表現形のタイトルとして、プロパティ `rdam:P30134("has title")` と、そのサブプロパティ `rdam:P30156("has title proper")`、`rdam:P30203("has parallel title proper")` など RDA エlementのすべてが表現できるよう登録されている。著作のタイトルについても、プロパティ `rdaw:P10088("has title of the work")` とそのサブプロパティ `rdaw:P10086("has variant title for the work")`、`rdaw:P10223("has preferred title for the work")` とがある。

それに対して、BIBFRAME では、プロパティ `title` があり、定義域として著作、インスタンス、アイテムのいずれも可とし、値域はクラス `Title` としている（図 2）。なお、プロパティの定義域をたとえば著作としたときには、その値域として `WorkTitle` (`Title` のサブクラス)をとる。同様に、`Title` のサブクラスには、`WorkTitle`、`InstanceTitle`、`VariantTitle` があり、`VariantTitle` はさらにサブクラス `KeyTitle`、`AbbreviatedTitle`、`ParallelTitle`、`CollectiveTitle` をとる。また、クラス `Title` を定義域にもつプロパティには `mainTitle`、`subtitle`、`partNumber`、`partName` があり、これらは値としてリテラルをとる。クラス `VariantTitle` を定義域とするプロパティ `variantType` もある。このようなクラスとプロパティを組み合わせることで表現できるため、全体としてかなりの表現力を有するが、RDA のすべてのプロパティを適切に表現できるわけではない。たとえば、`rdam:P30129("has later title proper")` や `rdam:P30130("has earlier title proper")` は適切に表現できない。

2) 責任表示、版表示、シリーズ表示にかかわるプロパティ：責任表示について RDA では、プロパティ `rdam:P30117("has statement of responsibility")`、そのサブプロパティ `rdam:P30105("has statement of responsibility relating to title proper")`、`rdam:P30116("has parallel statement of responsibility relating to title proper")` がある。一方、BIBFRAME はプロパティ `responsibilityStatement` のみがあり、その結果 RDA との対応づけでは多対一となる。

版表示やシリーズ表示についても同様であり、RDA のプロパティから BIBFRAME のプロパティへは多対一のマッピングとなる。

3) 出版・頒布・製作等にかかわるプロパティ：RDA では、`rdam:P30110("has production statement")`、`rdam:P30111("has publication statement")`、`rdam:P30108("has distribution statement")`、`rdam:P30109("has manufacture`

statement")があり、さらにそれぞれに対して、たとえば rdam:P30111 の場合 rdam:P30088 ("has place of publication")、rdam:P30092 ("has parallel place of publication")、rdam:P30176 ("has publisher's name")、rdam:P30011 ("has date of publication") が登録されている。

他方、BIBFRAME では、プロパティ provisionActivity がクラス ProvisionActivity を値域としてとり、そのサブクラスに Production、Publication、Distribution、Manufacture があり、さらにこれらクラスにはプロパティ place、agent、そして date があるという幾分複雑な構成をとる (図 1)。加えて、インスタンスに対するプロパティ provisionActivityStatement もある。4)メタレベルのデータ項目に相当するプロパティ: RDA Registry では、RDA そのものとは若干異なったかたちで 6 つのメタエレメントプロパティ rdaz:P60001 ("has source consulted")、rdaz:P60002 ("has cataloguer's note")、rdaz:P60003 ("has status of identification")、rdaz:P60004 ("has explanation of relationship")、rdaz:P60005 ("has date of usage")、rdaz:P60006 ("has scope of usage") が、すべて定義域と値域が指定されない形式で登録されている。つまり、RDA 自体では著作・表現形のエレメント、個人等のエレメント、第 1 グループ実体間の関連に伴うエレメントという区別があり、いずれにも共通するものと、いずれかに限定されるものがあつたが、そうした区別をここでは撤廃している。

BIBFRAME は、定義域の指定をもたないプロパティ adminMetadata が値域としてクラス AdminMetadata をとり、そのサブクラスには DescriptionConventions、GenerationProcess、DescriptionAuthentication がある。また、これらのクラスを定義域とするプロパティ assigner、derivedFrom、changeDate、creationDate、descriptionConventions、descriptionLanguage、generationProcess、descriptionModifie、descriptionAuthentication、generaitonDate があり、それぞれの値域となるクラスは異なる。プロパティ adminMetadata は、定義域が指定されていないため、いかなるクラスにも適用可能という柔軟性を備えたかたちでの登録である。

5)第 2 グループ実体のプロパティ: RDA Registry には個人・団体など、それぞれのクラスごとに多数のプロパティがある。たとえば、個人にかかわるプロパティには、rdaa:P50106 ("has real identity")、rdaa:P50107 ("has date associated with the person")、rdaa:P50108 ("has other designation associated with the person")、

rdaa:P50111 ("has name of the person") などがある。一方、BIBFRAME では Agent 等にかかわるプロパティを備えていない。それらについては、他のスキーマに委ねるとの立場であるのか、詳細は不明である。

6)第 1 グループ実体と第 2 グループ実体の関連に相当するプロパティ: RDA では、著作と Agent の間にはプロパティ rdaw:P10065 ("has creator") と rdaw:P10047 ("has other person, family, or corporate body associated with a work") があり、表現形と Agent の間には rdae:P20053 ("has contributor") がある。さらに、それぞれのプロパティは、RDA 付録 I に掲載されている関連指示子 (relationship designator) に対応して、多様な役割ごとのプロパティ、たとえば rdaw:P10051 ("has designer") などが定義されている。

それに対して BIBFRAME では、著作、インスタンス、アイテムのいずれにも適用できるプロパティ contributor (値域はクラス Agent) と contribution (値域はクラス Contribution) が定義され、後者の Contribution はさらにそれ自身がプロパティ role と agent をとるとされている。

上記以外にも、第 1 グループ実体間の関連に相当するプロパティなど、RDA Registry に登録されているプロパティ群と BIBFRAME のプロパティ群では、いろいろと異なるところも多く、多対一や多対多の対応関係となるものもある。

以上のことより、RDA メタデータ作成用には RDA Registry に登録されている内容からなるスキーマを使用し、必要な時点でメタデータを BIBFRAME スキーマに機械的に変換することが適切と考える。なお、変換の際には、多対一や多対多の対応関係にあるものについては便宜的な処置を採用せざるをえない。

3. 多様な記述メタデータの交換・共有用スキーマとしての BIBFRAME

RDA メタデータを超えて、情報資源に対する多様な記述メタデータからの変換先スキーマまたは交換・共有用スキーマとしての BIBFRAME 2.0 を検討する。多様なスキーマや作成基準で作成されたメタデータを BIBFRAME スキーマに機械的に変換することが前提となる。

3. 1 著作とインスタンスの区分の導入

前述の通り、FRBR/RDA の実体「著作」・「表現形」と「体現形」に相当する区分を、BIBFRAME はクラス「著作」と「インスタンス」によって導入している。多様なメタデータの受け皿となることを意図したスキーマにおいて、こうした区分を採用することは稀である。その点で画期的であるが、それに伴う課題も大きい。

著作の定義、すなわち何をもって著作と見なすか、インスタンスとどのように区別するのか、さらにはその実現値の単位設定はどのようなものかなどについて共通理解がない状況、あるいは対象とする情報資源を著作とインスタンスの両者または一方のクラスのみで表現するなど多様な方式が混在する状況で、上記のような区分を導入したときに何が発生するかは不明瞭である。

そもそも a)多様なメタデータを集めて単一のスキーマで表現する、いわゆる aggregation のレベルを意図しているのか、あるいは b)それを超えてメタデータの照合やマージなどを含め、実質的な交換・共有というレベルを意図するのかが未規定である。前者であれば、BIBFRAME スキーマによってメタデータを表現した時点で処置は終わり、同一 URI を介したリンクのみが達成される。それに対して、後者であれば複数のデータ項目値によるメタデータの照合やマージなどを含めた上での交換・共有を意図することになる。

3. 2 著作とインスタンスのプロパティ設定

著作とインスタンスのクラスの区分に加えて、それぞれのプロパティ設定について見てみると、BIBFRAME では、それらのプロパティは例外を除いて著作とインスタンスのいずれにも適用できるよう両者を定義域としている。プロパティ natureOfContent 、 intendedAudience 、 dissertation、summary、capture、notation、contentAccessibility、その他多数は、プロパティの趣旨からいけば本来は著作を定義域とすることになるが、BIBFRAME では著作とインスタンスのいずれも定義域としている。主題を表現する subject と classification についても同様である。一方、本来はインスタンスを定義域とすべきプロパティ custodialHistory、acquisitionTerms、acquisitionSource、coverArt など、同様に著作とインスタンスのいずれも定義域としている。

これらは定義域を著作とインスタンスのいずれかに限定したかたちで採用することには、スキーマの利用層や範囲などの点で困難であるとの判断の現れといえよう。なお、一方のクラスのみを定義域とするプロパティもある。originDate、historyOfWork、historyOfWork などの定義域は著作のみであり、responsibilityStatement、editionStatement、provisionActivityStatement などの定義域はインスタンスのみである。しかし、このように限定した定義域をもつプロパティと、著作とインスタンスの両者を定義域とするものとの区別は判然としない。

同様に、著作間、インスタンス間の関連を表すプロパティも、accompaniedBy、accompanies、hasDerivative、derivativeOf など、殆どのもの

が定義域、値域の両者とも、著作またはインスタンスとしている。これらプロパティは著作間、またはインスタンス間、あるいは著作とインスタンスの間にも適用しようということである。

このように、BIBFRAME は著作とインスタンスのクラスの区分を導入しながら、同時に多くのプロパティについてそれら両者を区別しない、すなわち限定することを回避するという方針をとっている。この点の妥当性には疑問がある。

ここで、著作とインスタンスにかかわる想定可能な選択肢を挙げると、下記の通りである(制約の強い順から記載)。

選択肢 1: 著作とインスタンスのクラスを分け、プロパティは著作またはインスタンスのいずれかのみにも適用できるよう定義する。

選択肢 2 (現行 BIBFRAME 2.0): 著作とインスタンスのクラスを分け、プロパティはいずれのクラスにも適用できるよう定義する。

選択肢 3 (OCLC による schema.org の採用方式): 著作とインスタンスのクラス自体を廃止し、単一のクラスとする(OCLC では CreativeWork)。その上で、クラスに属するリソースのタイプ、たとえば著作かインスタンスかを指定するためのクラスを別途定義する。プロパティは、基本的にこのタイプ指定の有無とは無関係に適用できるよう定義する。これらの結果、より制限の緩い方式となる。

選択肢 4: 著作とインスタンスのクラス自体を廃止し、たとえばクラス「リソース」のみとする。プロパティはこのクラスを定義域としてとることになり、結果的には Dublin Core の方向性に近づける結果となる。

3. 3 汎用的な属性やメタレベル属性に相当するプロパティ設定

前述の通り、タイトルにかかわるプロパティや、メタレベルのデータ項目に相当するプロパティなどは、対象リソース(定義域)を限定せず、そのため柔軟な表現力をもつ。同様に、識別子の使用を表すプロパティ identifiedBy、日付を表す date、言語を表す language、場所を表す place、注記を表す note など、多くの汎用的なプロパティが定義域を限定せずに登録されている。これらの点では、多様なメタデータを想定した柔軟な表現力を備えているといえよう。

注

- 1) Library of Congress. BIBFRAME Model & Vocabulary. <http://www.loc.gov/bibframe/docs/>
- 2) 谷口祥一. BIBFRAME とその問題点: RDA メタデータの観点から. 情報管理. Vol.58, No.1, 2015, p.20-27.
- 3) RDA Registry. <http://www.rdaregistry.info/>