

実装指向の表現形優先モデルの提案: 全体部分関連および aggregate の問題を介して

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

[抄録] 発表者が以前から提案している「表現形優先モデル」(表現形を基盤にした概念モデル)の一部を修正・更新し、実装指向のモデルとして再提案する。主に全体部分関連と aggregate (集合表現形) の問題を取り上げ検討することで、①表現形インスタンスの設定単位(粒度)のさらなる明確化が図られ、また②上記の両者が一貫した構図で扱える、かつ③設定が省略される実体インスタンスがあることを配慮した場合にもよりシンプルな構図で扱えるという、表現形優先モデルに対する優位性を示した。

1. はじめに

発表者は「表現形優先モデル」(表現形を基盤にした概念モデル)の検討と提案を以前から行っている¹⁾⁵⁾。当初は「テキストレベル実体を基盤にした概念モデル」と呼称していたが、FRBR や IFLA LRM においてテキストレベルの実体は「表現形(Expression)」と名付けられており、それに従い「表現形優先モデル」と呼ぶ。現行処理方式に対応する表現形を基盤とする概念モデルとそれに依拠したメタデータ作成から、表現形を基盤に採用したモデルおよびメタデータ作成への転換を意図したものであった。

FRBR や IFLA LRM をはじめとして、情報資源全般あるいは特定種別に対するメタデータの概念モデルは、対象とする情報資源を複数レベル(複数種)の実体によって表すモデルとしており、その際、実質的にいずれかの実体が優先されている。しかしながら、この点がモデルにおいて必ずしも認識され明示されているわけではない。優先される実体が異なれば、実体群の構成が同じでも、属性等の付与が異なる。また、個別事象のメタデータ記述において、実際に設定され記述される実体インスタンスも異なる場合がある。これらにより、いずれの実体を優先するかは、メタデータの相互運用性の点においても重要な問題である。

本発表は、以前に提案したモデルの一部の更新(修正)を含めた上で、実装指向のモデルとして表現形優先モデルを再提案する。主に全体部分関連と aggregate (集合表現形) の問題を取り上げ検討することで、①表現形インスタンスの設定単位(粒度)のさらなる明確化が図られ、また②上記の両者が一貫した構図で扱える、かつ③設定が省略される実体インスタンスがあることを配慮した場合にもよりシンプルな構図で扱えるという、表現形優先モデルに対する優位性を示す。

なお、FRBR や IFLA LRM に沿った議論とするため、実体関連モデル(ERモデル)を用いた議論とする。かつ、モデルに即したメタデータ記述、すなわち実体インスタンスとそれらの関連づけによってメタデータを構成することを検討の前提とする。すなわち、MARC フォーマットなど、モデルとは全く異なるスキーマによる記録とすることを前提とはしない。

2. 表現形優先モデルとは

2.1 表現形優先モデル

FRBR や IFLA LRM は表現形優先モデルに該当し、それらに依拠した記述規則である NCR2018 や RDA なども表現形優先モデルに即した記述規則である。

これらモデルにそのまま従ったときには、個別事例の記述において、基本的には「著作—表現形—表現形—個別資料」のインスタンスがすべて必要になる。著作者など個人等の行為主体との関連づけ、そして主題概念との関連づけには、著作インスタンスが不可欠となるからである。また、NCR2018 と RDA では、表現形を介せずして著作と表現形とを直接結びつける関連が登録されているが、表現形に関わる行為主体(翻訳者など)を記録する場合には、表現形インスタンスも併せて必要になる。

表現形は自立的とは言えず、著作に従属して存在する弱実体と位置づけられる。表現形には著作インスタンスとの関連なしに自立できるだけの属性が与えられていない。

他方、現行 MARC フォーマットを用いた NCR2018 や RDA による実装の選択肢においては、著作・表現形ともインスタンスの明示的な設定が求められず、表現形インスタンスと、著作者など行為主体インスタンスとを直接関連づけるような、曖昧な構図が許容されていることになる。主題概念との関連づけについても同様である。

2. 2 表現形優先モデル

表現形優先モデルとは、実体「表現形」を優先させ、物理的な媒体やフォーマットから独立して「コンテンツ」を捉え、それに基づくメタデータ作成を可能にすることを意図している。同時にそれは、手元の個別資料から出発し、表現形まで安定して抽象化できることを意味する(表現形優先モデルでは表現形への抽象化にとどまる)。こうしたモデルにより、利用者の要求に応えるコンテンツレベルの組織化に近づけることができると考える。端的には、多様な媒体やフォーマットによる同一表現形(冊子体の図書と電子書籍、冊子体の雑誌と電子ジャーナルなど)の異なる表現形を単一の表現形インスタンスにまとめることができる。なお、表現形優先モデルが有効である資源タイプは、冊子体、電子媒体等の媒体を問わず、いわゆる刊行物(出版物)である。他方、書写資料や博物資料など非刊行物には有効ではない。

表現形優先モデルは、構成する実体の点では表現形優先モデルと同じであるが、属性設定において相違する。表現形を自立した実体とするために、タイトル、責任表示、版表示など、表現形に対応するものをすべて表現形の属性として付与する。よって、版表示が形態上の相違などを示す場合(「大活字版」など)は、ここに含めない。これに合わせて、表現形は表現形に從属する弱実体とする。表現形の属性には、表現形に割り当てた要素は基本的に含めず、表現形の属性値と異なる場合のみ記録する。

上記は、テキストの微細な変更などは捨象し、タイトル等の外形的な手がかりによってテキストの同一性が示されることを仮定している。つまり、表現形ごとに通常、タイトルと責任表示等が付与される、逆にいえば同一表現形が表現形ごとに異なるタイトル等を有することは再出版など例外的なケースに発生する以外は稀であることに依拠している。

併せて、著作者等などの行為主体や主題は、表現形と直接関連づけるモデルとしている。これに加えて著作レベルで行為主体と関連づけても無論問題はない。

3. 全体部分関連と aggregate

3. 1 全体部分関連と aggregate との区別

定義上では、全体部分関連の場合、全体と部分の関係は緊密かつ本質的であり、部分は常に全体が想定され存在している。部分は全体にとって不可欠であり、部分が欠けると基本的に全体が機能しないことになる。

それに対して、aggregate の場合には、個別の要素が「集められる」のみであり、要素が全体にとって不可欠とはいいがたい。一部の要素を取り除いても、全体は理解可能である。逆に、個別要素はそれ自身で理解可能・自立可能である。

こうした区別に基づき、IFLA LRM は aggregate (「複数の表現形を具体化している 1 つの表現形」; 集合表現形) を導入している。また、その該当する範囲を、「表現形の集合コレクション」、「増補による集合表現形」、「並列的表現形の集合表現形」の 3 種に限定している。

3. 2 表現形優先モデルによる扱い

IFLA LRM、そしてより詳細を規定した現行 RDA を、ここでは取り上げる。RDA は、FRBR に依拠する NCR2018 とは異なる部分も多い。

a) RDA では全体と部分という全体部分関連は、著作間、表現形間、そして表現形間の各実体レベルにそれぞれ定義されている(図 1)。表現形インスタンスは設定が省略されうることを、図では点線で囲み表している。また、表現形は部分に該当するインスタンスを設ける場合と、設けない場合とがあることを併せて示した。

b) aggregate の場合(図 2)、aggregating (集合化する)側の実体は、「集合表現形」(aggregate manifestation)と、それに対応する「集合化表現形」(aggregating expression)、「集合化著作」(aggregating work)となり、集合化された個別の要素とは、表現形間の関連「aggregated by」(集められた) / 「aggregates」(集めた)によって関連づけられる。ここには著作レベルの関連はない。

c) これらにより、全体部分関連の構図と aggregate による構図とは大きく異なることになる。それゆえ、単一の記述対象において両者が組み合わせられた事例の場合には、複雑な構成となる。

3. 3 表現形優先モデルによる扱い

a) 全体部分関連は、基本的に表現形優先モデルの場合と同じ構図となる(図 3)。ただし、インスタンス設定が任意となるのが、表現形ではなく、著作レベルである点が異なる。著作者等の行為主体や主題が表現形インスタンスと直接関連づけられる点については、前述した通りである。

また、部分の表現形インスタンスから全体に該当する表現形インスタンスへの関連づけも任意となる点が異なる。部分の表現形は、表現

形間の全体部分関連を經由して、全体に該当する表現形そして体現形につながるものが基本的に想定されている。体現形優先モデルと同様、部分に該当する体現形インスタンスを設けない場合もありえることが、ここに含まれる。

b) aggregate については、以下の構図および議論となる(図4)。

・表現形が実質あるもの、かつ主たる実体とされていることを受けて、「集合表現形」(aggregate expression)に呼称を変更する。

・「増補による集合体現形」は、本文に加えて、挿絵、序論、解説・注釈などが併せて収載されている事例を指す。本文の著作者とは異なる行為主体による場合も多い。これらは広く捉えれば、パラテキスト(パラテキスト; paratext)の一種と見ることもできる。こうした要素の扱いが以前の表現形優先モデルでは曖昧であった。本発表では、こうした要素を含めて表現形を構成すると捉える。その結果、表現形と体現形の関連のカーディナリティは、「多対多」ではなく、あくまでも「1対多」となる。そして、集合表現形内の挿絵や解説など個別要素を独立して記録する場合には、それぞれが表現形インスタンスとなる。他方、表現形を著作に近い、「作品としてのテキスト」によって粒度を設定してしまうと、パラテキストの要素を含めることができず、個別要素をインスタンスとして別途設定することができないことになる。

・個別要素の表現形インスタンスから集合体現形への関連づけも任意とする。個別要素に相当する表現形は、表現形間の関連(「aggregated by」)を經由して集合表現形に至り、それが集合体現形につながることを想定している。

・著作のインスタンス設定が省略されうことは、表現形優先モデルの特徴である。著作者等の行為主体や主題は表現形インスタンスと直接関連づけることができる。

c) これらにより、全体部分関連として通常扱われる全体と各章との関係の構図と、本文と挿絵・解説などといった aggregate の構図とが基本的に同じモデル図で表すことができる。つまり、両者を同じ扱いとすることができ、両者が組み合わせあった事例においてもシンプルな構図で対処できる。これは、すべての体現形を aggregate として扱う可能性に言及した Coyle の議論に通じる結果となる⁶⁾。

3.4 その他の追加的議論

a) 表現形の属性「代表体現形」の採用

提案モデルでは、表現形を識別するための支

援要素として属性「代表体現形」(representative manifestation)を追加して採用する。多くの場合、当該表現形の記録に用いた体現形の出版者、出版日付などを記録することを想定している。表現形に対する「情報源」に近い要素でもある。この属性によって、類似する表現形が多数存在する場合に、インスタンス設定や照合などを確信あるものとすることができる。IFLA LRM や RDA における著作の属性「代表表現形」(representative expression)と同じ考え方を表現形に適用したものである。

類似する表現形、たとえば日本語の旧仮名づかいと新仮名づかい、旧字体と新字体、漢字表記と仮名表記などの区別による表現形については、それぞれを独立した表現形として認定するかが問題となる。こうした場合にも、上記の属性「代表体現形」が有効であろう。

b) 現行 MARC レコードによる構図からの移行

前述の表現形インスタンスの設定単位(粒度)は、現行の MARC レコードからの移行をより容易とする。具体的には MARC 書誌レコードを表現形と体現形に分割した上で、同一表現形インスタンスの機械的な照合と統合のみ実行することで移行がほぼ実現できるからである。それを超えた表現形の照合等には人手による介入が不可欠となる。

引用文献

- 1) Taniguchi, Shoichi. "A conceptual model giving primacy to expression-level bibliographic entity in cataloging," *Journal of Documentation*. Vol. 58, No. 4, 2002, p. 363-382.
- 2) Taniguchi, Shoichi. "Conceptual modeling of component parts of bibliographic resources in cataloging," *Journal of Documentation*. Vol. 59, No. 6, 2003, p. 692-708.
- 3) Taniguchi, Shoichi. *A Conceptual Modeling Approach to Design of Catalogs and Cataloging Rules*. ひつじ書房, 2007, 317p.
- 4) Taniguchi, Shoichi. What does giving primacy to a certain entity cause in a conceptual model for cataloging?: An Expression-entity dominant model revisited. *Library Resources and Technical Services*. Vol.61, No.4, 2017, p.212-225.
- 5) 谷口祥一. 実体「表現形」優先モデル再論: 特定の実体を優先することの帰結とは何か. 2016年度日本図書館情報学会春季研究集会発表論文集. 2016, p.55-58.
- 6) Coyle, Karen. *FRBR, Before and After: A Look at Our Bibliographic Models*. American Library Association, 2016, p.130-135.

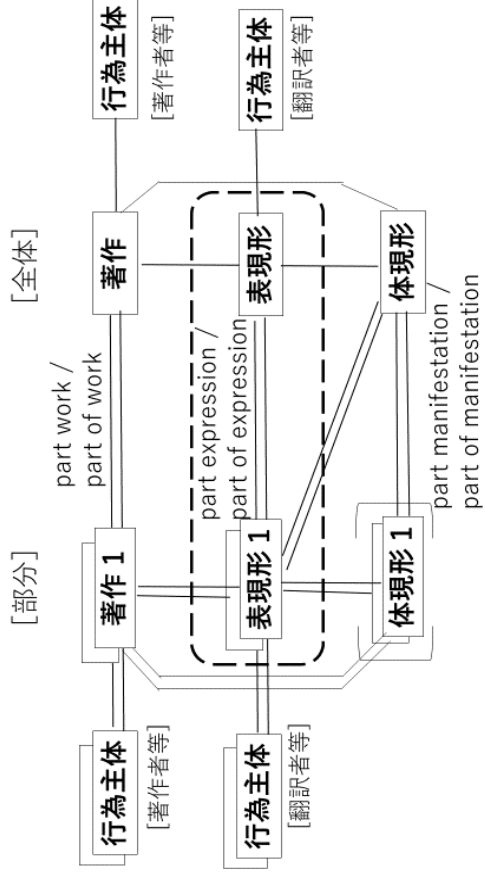


图 1 体现形优先モデル (RDA) : 全体部分関連

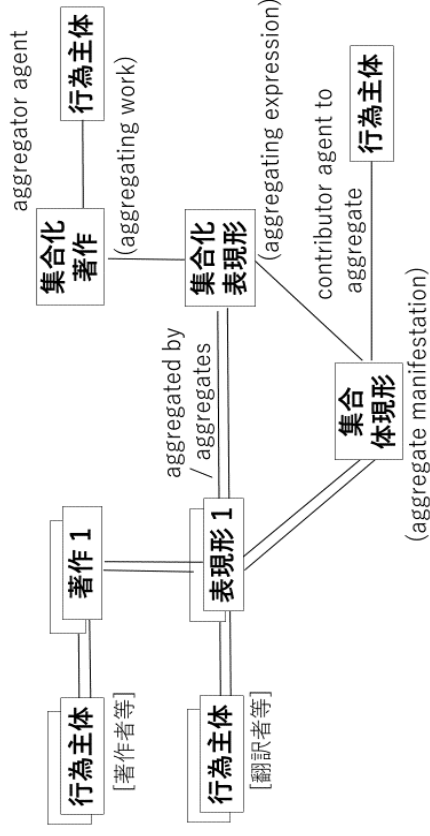


图 2 体现形优先モデル (RDA) : aggregate

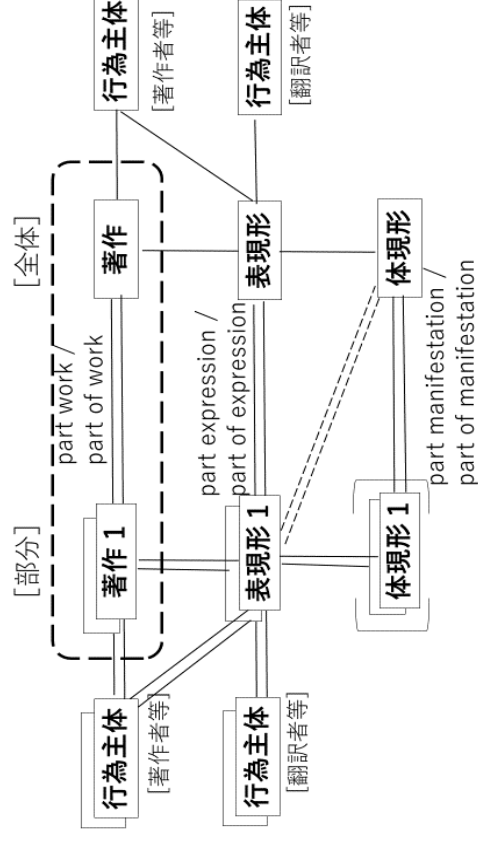


图 3 表现形优先モデル : 全体部分関連

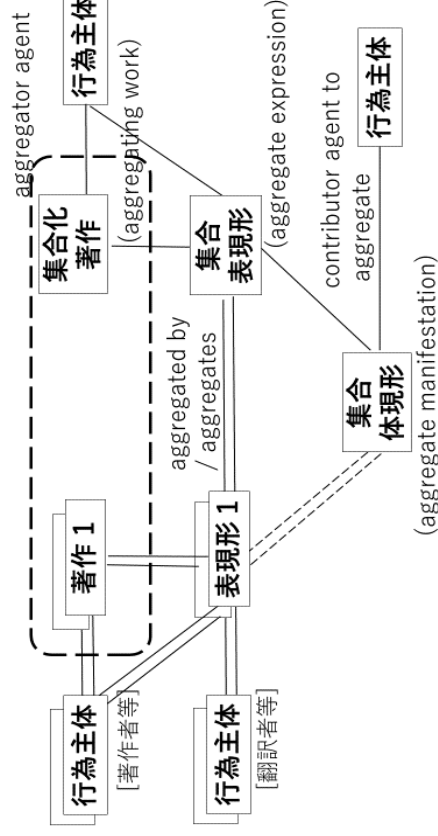


图 4 表现形优先モデル : aggregate