

# 典拠形アクセス・ポイント、管理番号としての ID、レコード概念の再検討 図書館目録メタデータのリンクトデータ化の観点から

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

図書館目録のメタデータを、現行の MARC 書誌・典拠レコードによるものからリンクトデータ (LD) へと移行するには、採用する概念モデルとそれに基づく語彙定義などが主要な検討課題となる。本研究では、それら以外の残された検討事項のうち、①典拠形アクセス・ポイント、②管理番号としての ID、そして③レコードの概念について、LD 化の観点から機能の確認と必要性の検討、事項相互の関連の確認、加えて LD 化における選択肢の整理を試みる。

## 1. はじめに

図書館目録のメタデータを、現行の MARC 書誌・典拠レコードによるものからリンクトデータ (LD) へと移行することを想定した場合、採用する概念モデルとそれに基づく語彙定義などが主要な検討課題となる。概念モデルとは、対象とする情報資源とそれに関わる事象をどのように捉えるべきかをモデルとして表現したものである。こうしたモデリングは RDF によるものに限る必要はなく、実体関連モデル (ER モデル) を用いる場合も多い。その上で、LD では、採用した概念モデルを構成する要素群を RDF による語彙 (RDF クラス、RDF プロパティ) として定義し、さらに必要な語彙を追加して定義することが通常行われる。RDF 語彙定義とは、RDF クラスの場合、ラベル、定義文、上位・排他関係など他クラスとの関係指示などから構成される。また、RDF プロパティ定義とは、ラベル、定義文、定義域と値域、プロパティ種別、逆・上位プロパティなど他のプロパティとの関係指示などが含まれる。現在、図書館目録のメタデータについて、複数の概念モデルや語彙が並行して提案されており、大筋の合意形成にも至っていない状況にある<sup>1)</sup>。

LD 化に当たっては、上記以外にも多様な検討すべき事項が残されている。これらのうち、本研究では①典拠形アクセス・ポイント (典拠形 AP)、②管理番号としての ID、そして③レコード概念について取り上げる。①から③のそれぞれに関して LD 化の観点から機能の確認と必要性の検討、事項相互の関連の確認などを試みる。換言すれば、これらに関する検討の構図の提示、そして LD 化における選択肢の整理を意図している。表 1 に、LD 化における選択肢の一覧を簡略版ながら示した。なお、本研究では、図書館目録のメタデータを当初から LD

として設計し作成するのか、既存 MARC レコード等から機械的に変換し LD とするのかについては問わない。いずれにも当てはまる議論と考える。

## 2. 典拠形アクセス・ポイント

### 2. 1 NCR2018、ICP、現行 RDA における位置づけ

NCR2018 は、典拠形 AP に関して、“当該実体を他の実体と一意に判別する典拠形アクセス・ポイントは、...” (#0.5.4)、“特定の实体に対する典拠形アクセス・ポイントは、他の实体に対する典拠形アクセス・ポイントと明確に判別される必要がある。” (#21.1) と規定し、形としての一意性を求めている。

それに対して、国際目録原則覚書 (ICP; 2016) は、典拠形 AP について、“ある実体を同一の名称の他の実体と区別することが必要な場合には、その実体に対する典拠形 AP の部分として、さらに識別のための特性を含めるものとする。” (5.3.4.5) とし、必ずしも形の一意性を求めている。現行 RDA も同様な扱いである。

また、NCR2018 は、典拠形 AP の機能を“実体を発見、識別し、または関連する他の実体を発見する手がかりとなる。” (用語解説。より詳細な記載は #21.1.1 に見られる) とし、ICP は典拠形 AP を含めて“統制形 AP は、一群のリソースの書誌データを集中する (collocating) ために必要な一貫性を与える。” (5.1.1) としている。

### 2. 2 典拠形 AP の機能の再規定

典拠形 AP については、機能を含めて見直しの議論がある<sup>2), 3)</sup>。こうした議論を踏まえつつ、本研究での検討の対象は、主題表現に用いる普通件名等は除いた、行為主体や著作などの典拠形 AP とする。

上記の範囲の典拠形 AP について、その機能を本研究では以下の 3 つに整理する。

- a) 一意性による識別子。ただし、典拠形 AP に一意性を持たせたとしても、関連先の指示（リンク）などを担う識別子として適切ではない。具体的には、典拠形 AP は長い文字列となり視認性などの問題、そして一意性を保つための後からの更新（識別要素の後の追加など）がありうる点が問題となる。特に著作や表現形の典拠形 AP の場合には、こうした問題は顕著である。RDA は典拠形 AP を識別子ではなく、「構造記述」による値と位置づけている点は妥当である。
- b) 対象リソース群のコロケーション（集中）。典拠形 AP による当該機能の実現の意義は残されているが、そこでは一意性は必ずしも必要ではない。同一形の典拠形 AP が存在しても、コロケーションは実現可能と考えるべきである。メタデータ作成者および利用者の双方とも、典拠形 AP のみで対象リソースの識別を含めてコロケーションを実現しているわけではない。一意の形にした無理にした典拠形 AP によっても判別等には必ずしも有効とはいえない。
- c) 対象リソースの情報提示。簡潔な形で対象リソースの特徴を表現するという機能である。ここにおいても、一意性は不可欠ではない。

次に、LD においては、a)一意性による識別子の機能は URI へと移行することになり、実際には現行（LD 化以前）でもローカル ID がそれを担っていることが常態である（この点は「3. 管理番号としての ID」で取り上げる）。

それ以外の機能である b)と c)については、その必要性の判断が、実際には LD 化にかかわらず以前から求められている。判断に当たっては、i)著作や行為主体等の認定基準（単位設定）と典拠形 AP の構成の問題は切り離し、かつ ii)典拠形 AP に一意性を求めない方式が前提とされるべきと考える。異なる典拠形 AP 構築規則、または同一構築規則内での異なる選択肢（別法など）に基づき構築し付与された典拠形 AP をもつメタデータとの混在利用という状況をも前提として想定すべきと考える。

さらには、典拠形 AP を特権的な位置づけをもつラベルとはせず、単純な識別子（URI など）のみによって典拠データおよびそれとのリンクを維持・管理するという典拠コントロールのあり方も考えられる。そこでは、典拠形 AP は他の多様なラベルと同じ位置づけとなる。旧来の典拠コントロールから新たな「identity management」に移行するには、識別子とそれらに結びつけられたラベルとを切り離すべき

との議論となる<sup>4)</sup>。

### 3. 管理番号としての ID

#### 3. 1 ローカル ID の機能

NDL による MARC21 書誌レコードフォーマットのフィールド 001「レコード管理番号」、015「全国書誌番号」、そして 6XX「典拠形 AP」や 70X-75X「AP ブロック（創作者等）」などのサブフィールド 0「典拠レコード管理番号」に記録されているものが、管理番号として付与されたローカル ID に該当する。また、上記の「典拠レコード管理番号」は、同じく MARC21 典拠レコードフォーマットの 001「レコード管理番号」に記録されている番号に対応する。

こうした ID は、a)ローカルな識別子として、レコードを特定し、他とリンクづける（あるいは他からのリンク先となる）ために機能する。それと同時に、b)ID によって当該レコードの構成範囲（情報表現単位）を指示するという機能を担っている。

LD では、a)識別子とそれによるリンク先の指示機能は URI に移行するため、URI を構成する要素としてローカル ID を捉え直す必要がある。NDL による上記の ID 群は、それぞれ固定のドメイン名とパスを用いて URI の構成要素とされている<sup>5)</sup>。一方、b)レコードの構成範囲指示機能については、「4. レコード概念」において整理する。

#### 3. 2 LD における URI 付与

ローカル ID をベースに URI を構成し付与するとしたときに、付与対象が対象リソースか、そのメタデータかという区別が必要になる。これまでの ID がメタデータに対する ID であり、これを転用して URI を構成するとした場合、LD においてはそれとは別にデータ記述（トリプル群）の主語となる対象リソースの URI が必要になる。つまり、メタデータのエレメント（プロパティ）の定義域は、通常、メタデータではなく、対象リソースそのものである。そして、この問題は URI における「RWO（実世界オブジェクト；非情報リソース）」と「情報リソース」の区別にはほぼ対応する。ただし、対象リソースがリモートアクセス可能な電子資源等である「情報リソース」であったとしても、それ自体の URI とそれに対するメタデータ URI を区別することは必要である。

URI 付与に関して追加される検討課題、そして選択肢には下記のものがある。

a) 対象リソース URI とメタデータ URI を区別する範囲。個人等の行為主体は URI のこうした区別を適用することが多いが、著作・表現形・体現形・個別資料 (WEMI) についてはいずれも区別するのか、あるいは著作・表現形のみ区別するのかといった選択肢がある。他方、主題表現用の普通件名など概念自体とそのメタデータを区別することは稀である。

Web NDL Authorities では、「個人、家族、団体、著作」について2つの URI を区別している。著作「源氏物語」の URI は <http://id.ndl.go.jp/auth/entity/00633493>、その典拠データの URI は <http://id.ndl.go.jp/auth/ndlna/00633493> となる。一方、「地名、統一タイトル、普通件名、ジャンル・形式用語」はこうした区別をせず、単一の URI によって両者を表すとしている。なお、「著作」と「統一タイトル」はここでは異なる扱いであり、後者は主題表現にのみ用いるものを指す。

他方、JLA 目録委員会が公開している NCR2018 データ作成事例の RDF データでは、WEMI そして個人等の行為主体のいずれも RWO と捉え URI を付与するが、それとは異なるメタデータ URI は適用していない (ただし、データ管理情報 URI を別途適用している)。

上記の URI の区別を適用するとしたときに特に問題となるのは、WEMI の扱いであろう。この URI の区別の問題は、それぞれの URI を主語リソースとするプロパティの定義域と整合させる必要があるため、プロパティの問題でもある。

b) メタデータ URI の付与方式。対象リソースではないメタデータに URI を付与するときには、多くの場合、この URI がデータ記述の主語リソースとなる。それゆえ、主語リソースが属するクラスを定義域とする (あるいは、それとは矛盾しない) プロパティを適用してメタデータを構成することになる。Web NDL Authorities における「典拠情報 URI」がこれに該当し、そこでは外部定義 (skos, skos-xl, dcterms など) の既存プロパティ群と独自定義 (dcndl) のプロパティ群が適用されている。

もう1つの方式は、対象リソースに対するデータ記述 (記述セット) 全体を RDF グラフと捉え、それに URI (グラフ URI) を付与する方式である。グラフ内のデータ記述の主語リソースは、通常、対象リソース URI となる。

### 3. 3 Nomen とデータ管理情報・由来情報

## への URI 付与

LD における URI 付与の問題において、Nomen とデータ管理情報・由来情報への URI 付与という選択肢が追加される<sup>6),7)</sup>。

a) Nomen への URI 付与。Nomen とは実体インスタンスを指し示す呼称 (appellation) であり、識別子、AP、タイトルなどを包含する。こうした Nomen に独立した URI を付与し表現または管理するという選択肢がある。この URI は URI 種別としての「情報リソース URI」となる。RDA では、記述ステートメント「対象リソース URI – appellation of RDA entity – Nomen URI」となり、Nomen URI を主語にしてプロパティ nomen string (Nomen 文字列)、note on nomen (Nomen に関する注記)、category of nomen (Nomen のカテゴリ)、context of use (利用の文脈)、date of usage (使用期間) などが適用される。

それに対して、Nomen に URI は付与せず、空白ノードで代用する、あるいは対象リソース URI に対して nomen string など、先に挙げたプロパティを直接適用することが、多くの場合に採用されている。その結果、「対象リソース URI – appellation of RDA entity – "nomen string"」となり、RDA ではこの方式を「ショートカット」と呼んで許容している。

b) データ管理情報・由来情報への URI 付与。データ管理情報・由来情報の主語となるものに URI を付与し、その下にデータ管理情報・由来情報に属するプロパティを適用し記録する方式がある。この場合、先の URI は、対象リソース URI (あるいはそのメタデータまたは名称データ) とリンクさせることになる。

JLA 目録委員会による NCR2018 の RDF 語彙を用いたデータ事例では、こうした「データ管理情報 URI」を採用している。エレメント「確定状況」、「名称未判別標示」、「出典」、「データ作成者の注記」に該当するプロパティが当該 URI に適用される。

他方、他の記述データと同じ位置づけで管理情報・由来情報を記録する、すなわち管理情報・由来情報自体に URI は付与せず、i)管理情報・由来情報プロパティの主語リソースは対象リソース URI (またはそのメタデータ URI や名称 URI) あるいはグラフ URI とする方式、または ii)空白ノードとする方式もありうる。

## 4. レコード概念

### 4. 1 レコードという情報表現単位の必要性

1つのまとまった括り (単位) で対象リソ

スを表示するデータ記述や名称データが「レコード」となる。単一の記述ステートメント（空白ノードによる構造化を含む）を超えて、構成体としての意味表現を求めるのか、すなわち個別記述ステートメントの単なる集合以上の意味をもたせるのかという選択となる。あるいは、同一対象リソースに対する複数の異なるデータ記述を相互に区別する必要性を認めるのかという選択でもある。

そのためには、レコードごとに異なる対象リソース URI を採用し、それらの関係を別途表現する方式がある。同一のリソースであれば、URI 間に owl:sameAs を適用し関連づけることができる。あるいは、レコードごとにグラフ URI を付与する方式もある。

それに対して、そうした URI を付与せず、対象リソース URI を主語とする記述セットを、そのまま「特定データセット内でのレコード」と捉える方式は、可能とはいえ、脆弱な表現方式といえよう。

#### 4.2 データ管理情報・由来情報が対象とする範囲指示の必要性

対象リソースのデータ記述や名称データに対する管理情報・由来情報について、それが対象とする範囲を指示することの必要性について判断が求められる。すなわち、複数の異なる管理情報・由来情報のそれぞれが対象とする範囲を明示できるようレコードの概念を採用するのかという選択である。

この場合にも、管理情報・由来情報の対象となる範囲を明示的に示すには、対象とするデータ記述や名称データにグラフ URI を付与し、その URI を主語にして管理情報・由来情報を記録する方式となる。RDA において新たに導入されたクラス metadata work の適用もこれ

に該当する。

他方、i)前述した通り、他の記述データと同じ位置づけで管理情報・由来情報を記録する方式や、あるいは ii)レコード単位ではなく、ステートメント単位で管理情報・由来情報を記録する方式があり、後者は典型的には Wikidata における「ステートメント記述（statement description）」によるものが該当する。

#### 引用文献

- 1) 谷口祥一. 知識資源のメタデータへのリンクトデータ・アプローチ. 勁草書房, 2023.
- 2) Gordon Dunsire, and Mirna Willer. Authority versus authenticity: the shift from labels to identifiers. In: *Authority, Provenance, Authenticity, Evidence*. Ed by Mirna Willer, Anne J. Gilliland and Marijana Tomić. Zadar : Sveučilište u Zadru, 2018, p.87-113. ISBN 978-953-331-220-0
- 3) Gordon Dunsire. Reconstructing authorities: New approaches to the management and use of authority data. In: *Mirna Willer: Festschrift*. Zadar : Morepress, 2020, p.81-98. ISBN 978-953-331-275-0
- 4) Karen Smith-Yoshimura. *Transitioning to the Next Generation of Metadata*. OCLC Research, 2020. <https://doi.org/10.25333/rqgd-b343>
- 5) 国立国会図書館. 国立国会図書館データの URI. <https://www.ndl.go.jp/jp/dlib/standards/lod/uri.html>
- 6) Shoichi Taniguchi. Data provenance and administrative information in library Linked Data: Reviewing RDA in RDF, BIBFRAME, and Wikidata. *Cataloging & Classification Quarterly*. Vol. 61, No. 1, 2023, p.67-90. <https://doi.org/10.1080/01639374.2023.2178048>
- 7) 谷口祥一. 図書館目録のリンクトデータにおけるデータ由来情報と管理情報：RDA in RDF、BIBFRAME、Wikidata の再検討. 三田図書館・情報学会 2022 年度研究大会発表論文集. 2022, p.41-44.

表1 リンクトデータ化における選択肢一覧（簡略版）

<b>典拠形 AP</b>
i) 一意性をもたせるか
ii) 対象リソース群のコロケーション、対象リソースの情報提示をもつばら典拠形 AP に担わせるのか
<b>LD における URI 付与</b>
i) 対象リソース URI とは異なるメタデータ URI を付与するか
ii) 対象リソース URI とメタデータ URI を区別する範囲をどのように設定するのか それぞれの URI を主語リソースとするプロパティがあるか（整合しているか）
iii) メタデータ URI は、それを主語リソースとするメタデータ記述か、対象リソースのメタデータ記述全体を指すグラフ URI とするか
iv) Nomen に独立した URI を付与するか
v) データ管理情報・由来情報に独立した URI を付与するか
<b>レコード概念</b>
i) レコードごとに異なる対象リソース URI を付与するか、情報表現単位としてグラフ URI を採用するか
ii) データ管理情報・由来情報の対象にグラフ URI を付与するか