

バーチャル国際典拠ファイル (VIAF) について

川鍋 道子*

キーワード：VIAF, バーチャル国際典拠ファイル, 典拠レコード, VIAF 評議会



本稿は、クリエイティブ・コモンズ表示-非営利-改変禁止 4.0 国際 (CC BY-NC-ND 4.0) ライセンスの下に提供する (<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.ja>)

1. はじめに

図書館資料などの書誌データ検索の手がかりとなる標目は、典拠レコードを用いて統制されている。書誌データ作成場面では、著者の例でいえば、資料の著者である個人や団体の典拠レコードを作成し、資料を表す書誌データにリンクづけている。典拠レコードの作成及びリンクの形成によって、同一著者の資料をもれなく検索できるだけでなく、同一名称の別人別団体を明確に区別できるように典拠レコードの標目形を作成している。典拠レコードの標目形自体を識別子としてとらえることができる。

国立国会図書館では、①個人名、家族名、団体名、地名、著作、統一タイトルという名称、②普通件名、③ジャンル・形式用語の3種の典拠レコードを作成し、国立国会図書館典拠データ検索・提供サービス (Web NDL Authorities. <https://id.ndl.go.jp/auth/ndla>) という典拠ファイルを提供している。

バーチャル国際典拠ファイル (Virtual International Authority File: VIAF) は、各国の国立図書館等から名称典拠レコードの提供を受けて、同一の実体 (データモデルにおける対象) に対する典拠レコードを機械的に同定し、各機関の典拠レコードをリンクする事業である。

本稿では、識別子の機能の観点も加えて、VIAF の概要、課題、今後について紹介する。

2. VIAF の概要

2.1 VIAF の典拠レコードの特徴

VIAF は、各機関の典拠レコードを一つの典拠レコードに統合するのではなく、各機関の各言語で作成された標目形を維持したまま、ひとかたまりのレコード (クラスター) として提供している¹⁾。

クラスターには各国の言語の標目形が表示され、各機関の典拠レコードへのリンク一覧がある。日本を示すリンクからは国立国会図書館の Web NDL Authorities の詳細情報画面とつながっており、VIAF と Web NDL Authorities

が相互にリンクされている。さらに、Web NDL Authorities から、典拠レコードにリンクした国立国会図書館の所蔵資料の書誌データを検索することができる。

2.2 VIAF の経緯

1990年代以降、典拠ファイルの共同構築・共同利用に関する国際的な取組みが進められていた。1998年に、米国議会図書館、ドイツ国立図書館、OCLC (Online Computer Library Center, Inc.) の3機関によって互いの個人名典拠のリンクをテストするVIAFのコンセプト検証がスタートした。2007年にフランス国立図書館が加入し、さらに、2012年4月にVIAFはOCLCが提供するサービスの一つとなった。

日本からは、2012年10月に国立国会図書館がVIAFに参加した。国立国会図書館は、OCLCの書誌データベースであるWorldCatへの書誌データ提供に加え、VIAFへ名称典拠レコードを提供することで、名称典拠レコードの国際的な流通を進めることとした。2017年3月には、国立情報学研究所がVIAFに参加し、NACSIS-CATの著者名典拠レコードのVIAFへの登録が開始されている²⁾。

2021年6月現在、VIAFの参加機関は57機関、各機関から提供された典拠レコードの合計は約8,729万件、同定された典拠レコード(クラスター)は約3,560万件である³⁾。

2.3 VIAF に期待される効果と問題

VIAFに期待される効果は、次の4点と考えられる。第一に、VIAFが存在することにより、利用者は、各機関の提供サービスを利用せずに、各機関が作成・維持管理している典拠レコードに使いやすい形でアクセスできる。典拠レコードのリンクをたどり各機関の書誌データの発見可能性向上が期待できる。第二に、典拠レコードの国際的な共有により、各機関の典拠レコード作成及び維持作業の省力化が図られる。第三に、VIAFを通じて、世界の書誌作成機関が日本の典拠レコードを利用でき、日本人や日本の団体の著者の同定・識別が容易になる。同様に、日本の書誌作成機関も他国の個人、団体の著者の同定・識別が容易になる。第四に、VIAFの典拠レコードは、最大22桁の数字で構成されているIDとパーマリンク (URI) を持っており、ウェブ上の情報資源とのリンクが可能である。

*かわなべ みちこ 国立国会図書館収集書誌部

〒100-8924 東京都千代田区永田町 1-10-1

(原稿受領 2021.8.27)

VIAFに期待される効果の発揮には、典拠レコードの正確な同定が前提となる。VIAFの同定メカニズムは公開されていないが、典拠レコードの文字列やIDなどを元に、機械的に同定が行われている。OCLCでは、メカニズムの改善、全件の再クラスタリングの実施、参加機関に対して典拠レコードの品質の維持向上やLCCN (Library of Congress Control Number)など他機関のIDの典拠レコードへの記録を推奨するなど、同定の正確性を追求している。だが、機械的な同定には限度があり、誤りは生じうる。なお、VIAF典拠レコードの識別子は上述のように最大22桁の数字であるが、その体系は明示されておらず、他の識別子の体系とは関係がない。

宮澤によれば、識別子の機能には、誰が最初に識別子を付与するか、どのようにして識別の衝突を避けるかなどの制度の確立と運用が必要である⁴⁾。VIAFの場合、特定の機関がユニークな「統一標目」を付与しているわけではない。各機関の典拠レコードを機械的に同定しており、VIAFの識別子としての利用には、誤った同定の可能性があることに注意する必要がある。

一方で、渡邊によれば、VIAFが典拠レコードの相互リンクを図る背景には、標目形の異なりは各国(文化圏)の利用者の便宜を考慮した結果であり、無理な国際的な統一は好ましくないという判断がある。VIAFが「バーチャル」の語を冠する所以である⁵⁾。

3. VIAFの活動から

VIAFの参加機関会合は、2012年のOCLCへの移管後、評議会という形で毎年開催されている。評議会では、OCLCからの活動報告、参加機関からVIAF事業の方針や実践、運営に関する重要事項への助言・指導が行われている。国立国会図書館は2012年の評議会から毎年参加し、評議会の概要を刊行物で報告してきた⁶⁾。同報告を基に、近年のVIAFの活動のトピックを紹介する。

3.1 個人情報の扱い

個人名典拠には、氏名のほか生年等の個人情報が含まれている。2012年10月に、当時のVIAF評議会議長であるVincent Boulet氏(フランス国立図書館)から個人のプライバシー問題を扱うワーキング・グループの設置が参加機関に対して呼びかけられた。2018年5月から、EU一般データ保護規則(General Data Protection Regulation: GDPR)の適用が開始され、VIAFとしてGDPRに基づく典拠レコードの修正・削除依頼への対応が必要となり、その方針検討のため、各参加機関の対応事例をアンケートで集約することとなった。2019年のアテネでの評議会⁷⁾で、個人情報保護ワーキング・グループ担当のRicardo Santos Muñoz氏(スペイン国立図書館)からアンケート結果の報告がされた。アンケートには37機関(うち国立図書館18機関)が回答し、大多数の機関が、著者などからの要求があった場合に典拠レコードから生没年や付記事項を削除する一方、半数以上の機関が典拠レ

コードそのものを削除という要求を却下していた。調査結果を踏まえ、VIAFの個人情報の取り扱いに関する方針書とデータ訂正の要望を受けた際の処理手順(protocol)の策定に向けて、検討中である。

典拠レコードにおける個人情報の取り扱いについては、各参加機関はもちろん、VIAFとしても慎重な対応が求められる。(生没年等を付記した)標目形という文字列による同定・識別よりも、ID等の識別子同士のリンクのほうが、個人情報保護の観点からも、図書館間だけでなく図書館関係以外の機関との連携の観点からも、扱いが容易であるといえるだろう。

3.2 外部サイトとのリンク

3.2.1 ISNIとの連携

国際標準名称識別子(International Standard Name Identifier: ISNI)は、創作物や著作物を創作、作成、編集、頒布した公的アイデンティティ(Public Identity)を識別する国際標準識別子である⁸⁾。ISNIとVIAFの間では、それぞれのデータ品質向上のためデータの相互利用等に関する連携が図られてきた。

2012年のヘルシンキでの評議会報告によると⁹⁾、ISNI側がVIAFの典拠レコードをベースとして読み込み、権利関係団体など他のリソースと同定し、VIAFを含めた各データにISNIの識別子を割り当てている。このISNI割り当ての過程で判明したVIAFのエラーは、ISNI側からVIAF側に通知され、VIAFを主管するOCLCからVIAF参加機関に報告される。各機関がエラーを修正し、その修正がVIAFに反映されれば、VIAFの典拠レコードの品質向上につながる。このようにISNIとVIAFは相互補完的であった。

蔵川は、ISNIとVIAFの関係について次のとおり言及している。「ISNIとVIAFは、運営組織はともに国立図書館を主体としており、対象とするデータベースは図書館に所蔵されるような創作物一般であるという点において類似のものである。それゆえ、どちらか一方がグローバルな著者識別子として存在すれば十分といえる典型例である。ISNIはVIAFの結果を最大限に組み込んで識別子の品質を高める方法を補完している。この点を考慮すれば、インターネット上のグローバルな著者識別子としてはISNIが代表し、VIAFはそのデータ構成元である各国図書館の名称典拠とISNIとの中継点として位置付けるのが妥当である」¹⁰⁾。

2017年のヴロツワフでの評議会では、英国図書館とフランス国立図書館による建議をきっかけにISNIとVIAFの連携や運営サービスの在り方が議論された。VIAFがOCLCによって無償で提供されるサービスであるのに対し、国際規格であるISNIはメンバーの基金で維持管理されている。VIAFとISNI両事業の相違も踏まえた連携についてはワーキング・グループを設置し、戦略的・技術的に具体的な方向性とロードマップをまとめることとなった¹¹⁾。

2018年のクアラルンプールでの評議会で、同ワーキング・グループでの検討について経過報告がなされている¹²⁾。同ワーキング・グループは、2017年12月に発足し、メンバーは、スペイン、フランス、ドイツ等の国立図書館、米国の大学図書館、OCLC、PCC (Program for Cooperative Cataloging : 米国議会図書館等による国際的な共同目録プログラム)等の担当で構成されている。ワーキング・グループでは、VIAFは典拠コントロール、ISNIはID管理を目的としたサービスと整理し、それぞれの利用方法を明確にするための各機関における事例やワークフローの確認が行われた。また、VIAFとISNIの間でのデータのやりとり等、相互運用性の向上に関する方法も検討されている。

クアラルンプールでの評議会の参加者からは、音楽セクター、図書館等の様々なコミュニティからなるISNIと、図書館の典拠コントロールをおもな目的としたVIAFが併せて継続することの重要性について意見があった。双方の運営に関与するOCLCから、双方が持続可能なビジネスモデルを引き続き検討する必要があるとのコメントがあった。

なお、OCLCによれば、2020年9月現在、VIAFからISNIへのデータ読み込みは、2016年から中断されたままであり、再開時期は未定である¹³⁾。

3.2.2 Wikipedia との相互リンク

2012年からVIAFとWikipediaとの相互リンクが開始された。VIAFの典拠レコードの関連情報「外部へのリンク」にWikipedia記事へのリンクがあり、Wikipedia記事の「典拠管理」にVIAFへのリンクがある。

石橋によると、この相互リンクプロジェクトは、OCLCリサーチのウィキペディアン・イン・レジデンスであるMax Klein氏らが主導したものである。Klein氏らがWikipediaユーザーへ提案、意見募集を行い、Wikipediaコミュニティから賛同を得て正式なプロジェクト開始に至った¹⁴⁾。

2015年のケープタウンでの評議会では、英語版WikipediaからWikidataへのリンク先の変更が報告された¹⁵⁾。Wikidata¹⁶⁾の作成により、英語版Wikipediaの記事だけでなく、他言語版の記事にもVIAFのIDが追加されるようになってきている。さらに、VIAFを通じてWikipediaの記事にVIAF参加機関の典拠レコードIDも追加されている。

世界でアクセスの多いWikipediaとの相互リンクにより、VIAFへのアクセスも大幅に増加したとのことである¹⁷⁾。Wikipediaの利用者がVIAFのリンクをたどり、各国の国立図書館等の資料検索につながっているであろう。

4. おわりに

VIAFの典拠レコードが機械的に同定されていることから、誤りが生じうることは先に述べた。木村は、VIAFのこの問題に関し、「VIAFには従来の典拠データベースが

保障してきたほどの正確さはなく、既存の典拠データベースをウェブの世界で活用できるように発展させたもの、と捉えるべきである。」と述べている¹⁸⁾。

VIAFのウェブ上での利活用、情報資源とのつながりに関して例を挙げると、図書館、博物館、美術館等のデジタルアーカイブの検索・閲覧・活用プラットフォームであるジャパンサーチの名称正規化レコードにVIAFのIDがリンクされている¹⁹⁾。

また、2021年12月に完成予定のSEMI (Shared entity management infrastructure) のデータソースとしてVIAFが活用されることが、2020年9月のオンラインでの評議会でOCLCから報告された。SEMIは、永続的で一元的な基盤として、著作と個人の実体の記述を公開するOCLCのプロジェクトである。他機関の語彙や典拠ファイルへのリンクを集約させ、ウェブ上の学術資料のつながりを促進し、発見しやすくすることが目的である²⁰⁾。

VIAFは、書誌作成機関による典拠ファイルの共同利用を意図して始まり、VIAFのID及びVIAFがまとめた識別子を通じた様々な取組の結果、参加機関が作成する名称典拠レコードのハブとして利活用されるようになった。そして、VIAFとの連携対象は図書館の典拠レコードにとどまらず、ウェブ上の様々な情報資源に広がってきている。VIAFのウェブ上の情報資源との連携の発展を引き続き注視していきたい。

注・参考文献

- 1) 収集書誌部収集・書誌調整課. ことばの壁をこえる典拠—バーチャル国際典拠ファイル (VIAF) への参加. 国立国会図書館月報. 2013, no.623, p.30-34.
- 2) 国立情報学研究所総務部企画課. CiNii Books にバーチャル国際典拠ファイル (VIAF) との連携機能を追加—OCLC と VIAF 参加の協定を締結. 平成 29 年 6 月 28 日. <https://www.nii.ac.jp/news/release/2017/0628.html>, (参照 2021-08-26)
- 3) 2021年8月20日にオンラインで開催されたVIAF評議会でのOCLCからの報告による。
- 4) 宮澤彰. 識別子序論. 情報の科学と技術. 2020, vol.70, no.5, p.262-265.
- 5) 渡邊隆弘. 典拠コントロールの現状と将来. 情報の科学と技術. 2010, vol.60, no.9, p.371-377.
- 6) 大柴忠彦. 典拠の国際流通—バーチャル国際典拠ファイル (VIAF) への参加 (3). NDL 書誌情報ニュースレター. 2013, vol.25, p.16-19.
- 7) 村上一恵. 世界図書館・情報会議 (第 85 回 IFLA 大会)・VIAF 評議会会議報告. NDL 書誌情報ニュースレター. 2020, vol.52, p.1-8.
- 8) 蔵川圭. 国際標準名称識別子 ISNI. 情報の科学と技術. 2020, vol.70, no.7, p.376-379.
- 9) 大柴・前掲注 6) NDL 書誌情報ニュースレター
- 10) 蔵川圭. 方法としての著者識別子. 情報の科学と技術. 2017, vol.67, no.2, p.67-72.
- 11) 津田深雪. 世界図書館・情報会議 (第 83 回 IFLA 大会) と VIAF 評議会会議参加報告. NDL 書誌情報ニュースレター. 2017, vol.43, p.1-6.
- 12) 柴田洋子. VIAF 評議会会議参加報告. NDL 書誌情報ニュースレター. 2018, vol.47, p.8-11.
- 13) 2020年9月18日にオンラインで開催されたVIAF評議会でのOCLCからの報告による。
- 14) 石橋恵. 図書館データとWikipediaをつなぐVIAFbot. カ

- レントアウェアネス-E1517. 2013, no.251.
<https://current.ndl.go.jp/e1517> (参照 2021-08-26)
- 15) 津田深雪. 世界図書館・情報会議 (第 81 回 IFLA 大会), VIAF 評議会会議報告. NDL 書誌情報ニュースレター. 2015, vol.35, p.1-6.
- 16) 大向一輝. 識別子としての Wikidata. 情報の科学と技術. 2020, vol.70, no.11, p.559-562.
- 17) 石橋・前掲注 14) カレントアウェアネス-E1517
- 18) 木村麻衣子. “典拠コントロールの動向”. メタデータとウェブサービス. 日本図書館情報学会研究委員会編. 勉誠出版, 2016, p.41-56. (わかる! 図書館情報学シリーズ, 3)
- 19) ジャパンサーチの SPARQL エンドポイント. <https://jpsearch.go.jp/rdf/sparql-explain/>, (参照 2021-08-26)
- 20) OCLC, ウェブ上の学術資料の発見可能性を高めることを目的とした “Entity Management Infrastructure” 構築のための助成金を獲得: 他機関の典拠ファイルや語彙へのリンクを集約. 2020 年 1 月 14 日.
<https://current.ndl.go.jp/node/39947>, (参照 2021-08-26)

Series: Identifiers used in information science and technology, 18: About Virtual International Authority File—VIAF. Michiko KAWANABE (Acquisitions and Bibliography Department, the National Diet Library, Japan, 1-10-1, Nagata-cho, Chiyoda-ku, Tokyo, Japan)

Keywords: VIAF / Virtual International Authority File / Authority record / VIAF Council