

[E13] アーカイブズ記述の新標準：

Records in Contexts (RiC) の実装方法

○元ナミ¹⁾, 橋本陽²⁾

1) 東京大学文書館, 〒113-8654 東京都文京区本郷 7-3-1

2) 京都大学大学文書館, 〒606-8305 京都市左京区吉田河原町 15-9

E-mail: won.nami@mail.u-tokyo.ac.jp; hashimoto.yo.4a@kyoto-u.ac.jp

The New Standard of Archival Description: The Implementation Methods of Records in Contexts (RiC)

WON Nami¹⁾, HASHIMOTO Yo²⁾

1) The University of Tokyo Archives, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8654 Japan

2) Kyoto University Archives, 15-9 Yoshidakawara-cho, Sakyo-ku, Kyoto, 606-8305 Japan

【発表概要】

本報告ではアーカイブズ記述の新標準である Records in Contexts (RiC) の実装方法について検討する。

RiC とは、アーカイブズ記述のための概念モデル (Conceptual Model) である RiC-CM およびそのオントロジー (Ontology) である RiC-O から構成される。他国を見れば、すでに RiC-O を用いたテスト段階の検索システムが少数ではあるが公開されている。それらを紹介しつつ、実装のための具体的なワークフローを構想する。

提示するのは次の方法である。Access to Memory (AtoM) を利用し、XML 形式でアーカイブズ記述のメタデータを作成する。それをフランス国立公文書館が開発した RiC-O Converter によって RiC-O 準拠の RDF/XML に変換する。最後に RDF/XML を Omeka-S などのアプリを使い、検索とその視覚化を実現させる。

1. はじめに

Records in Contexts (RiC) とはアーカイブズ記述のための概念モデル (Conceptual Model) である RiC-CM およびそのオントロジー (Ontology) である RiC-O から構成される。RiC は、既存の国際記述標準である General International Standard Archival Description (ISAD(G))、International Standard Archival Authority Records—Corporate Bodies, Persons, and Families (ISAAR(CPF))、International Standard Description of Functions (ISDF)、International Standard Description of Institutions with Archival Holdings (ISDIAH) を統合することを目的として作成されている。RiC-CM および RiC-O の作成と更新は国際アーカイブズ評議会 (International Council on Archives, ICA) 設立のアーカイブズ記述専門家グループ (Expert Group on Archival Description, EGAD) が担当し、ともに 0.2 番が現在までに公表されている[1]。

国内の研究では、寺澤正直が、RiC-O 公表以前の段階において、RiC-CM の概要説明および既存の 4 つの標準と RiC-CM の照合などの重要な作業を行なっている[2]。また、谷合佳代子が RiC の中の典拠データに着目し、その作成実験について報告している[3]。一方で、海外では RiC-O 公表後に、それに準拠したメタデータを使用したシステムの実装に着手した研究事例がいくつか確認される。これらの事例を参照することで、RiC に則ったアーカイブズ記述を国内で実現するための手がかりを得られるはずである。

そこで本報告では、RiC-CM および RiC-O の概要を確認した上で、フランスとブラジルの実装事例を紹介し、実装のための具体的なワークフローを構想する。なお、国内では、4 つの国際標準に準拠した記述を適切にマークアップできる Encoded Archival Description (EAD)、Encoded Archival Context (EAC) によって作成された xml ファイルを用意できていないことも、RiC 実装の障害の 1 つとなる。そのため、EAD.xml、EAC.xml の生成

も含めたワークフローを提示したい。

2. RiC-CM と RiC-O

RiC-CM は、アーカイブズに関する事象を抽象的に表現したモデルで、アーキビストの専門的な視点から作成されている。これは、記録資料群、記録資料群を作成し使用する組織や人々、その組織や人々が行う業務・活動を特定、記述するためのモデルである。

既存の 4 つの国際記述標準においては、ISAD(G)が記録資料群、ISAAR(CPF)が組織や人々、ISDF が業務・活動、ISDIAH が記録資料群の保存を請け負う組織を記述するという役割を果たしていた。また、それぞれの標準に準拠して記入された記述項目は、それぞれリンクさせることが意図されていた。同一の標準内の場合、ISAD(G)であれば、記録資料群内の関係性、つまりフォンド—シリーズ—ファイル—アイテムの各階層の記述がリンクさせられる。ISAAR(CPF)は、組織内の各部署といった組織間の関係性や親類などの個人間関係性を結びつけることができる。同様に ISDF は、機能—業務・活動のそれぞれの関係性を紐づける。一方、異なる標準の記述項目もリンクさせる。このリンクによって、例えば、ISAAR(CPF)で書かれる一つの部署が、ISDF 内のある業務を遂行する中で、ISAD(G)内のシリーズを作成したという状況を表現することができる。ISDIAH の記述は、アーカイブズの保存する記録資料群を記述した ISAD(G)の記述項目に紐づけられる。さらに、リンクの関係性について、記録資料群とその作成者でいえば作成や収集、作成者間でいえば前後・上下といったその種類も明示できる。

RiC-CM は、記録資料群、業務・活動および組織・人々を実体 (Entity) として特定、記述し、実体とそれぞれの関係性を示しながら可視化するという点で、既存の標準を含むモデルである。ここに、場所、出来事、権限などの実体が追加され、さらに記述を充実化および細分化させることができる。RiC-O は、RiC-CM の正式な実装様式であり、Resource Description Framework (RDF) の表示形式として、アーカイブズ記述を表すための語彙と規則を定めている

一方、RiC を実装させるためのガイドラインである Records in Contexts-Application Guidelines (RiC-AG) は現在のところ未公表である[4]。

しかし、そのような現状の中、実装に取り組む事例が確認できる。ここでは、フランスとブラジルの事例を紹介する。

3. フランスとブラジルにおける RiC-O 実装の事例

RiC-O 実装を先進的に進めているのが、フランス国立公文書館 (Archives nationales de France, ANF) である。ANF は、既存の国際記述標準に準拠して目録をウェブ上に公開してきたが、利便性を高めることを 1 つの目的として RDF で表示するためのプロジェクトを開始し、RiC の採用を決定した。その一環として、これまで作成してきた大量の EAD と EAC の xml ファイルを、RiC-O 準拠の rdf ファイルに変換する RiC-O Converter を開発した。それにより変換したファイルを、GraphDB を通じて、視覚的に表示する検索システムの試行を進めている[5]。

また、リオデジャネイロ州立大学の LABOGAD 研究所 (Laboratório de Preservação e Gestão de Acervos Digitais) は、「Atlantic Slave Trade」サイトに掲載されている目録データと筆写テキストを RiC-O で入力するプロジェクトを行なっている。具体的には、ISAD(G)をベースとしたブラジルのアーカイブズ記述標準 Nobrade[6]に準拠して記述されている Arquivo Geral da Cidade do Rio de Janeiro (AGCRJ) 所蔵の大西洋奴隷貿易関係文書 (Slavery Series) 目録とコンテンツの筆写テキストを RiC-O で記述し、「Atlantic Slave Trade」サイトと Linked Open Archives (LOA) [7]が相互に接続、検索できるようにする試みである[8]。

LOA は RiC で提供されるデジタルアーカイブのハブであり、そのプラットフォームは、デジタルコレクション管理ソフトウェアである Omeka S、仮想知識グラフシステムの Ontop、オブジェクト関係データベースの Virtuoso Open Source (VOS) で構成されている。将来的には、ブラジル、アンゴラ、ポルトガルにそれぞれ保管されている大西洋奴

隷貿易関係文書に、RDF フォーマットと SPARQL エンドポイント提供を通じて、エンドユーザーがアクセスできる Linked Open Data (LOD) 環境の構築を目指している。

4. 日本における RiC-O の実装

本研究では、既存のアーカイブズ目録記述を RiC-O に移行する検証作業を進めている海外の知見を利用し、現段階において RiC を利用した日本語のアーカイブズ記述をどこまで実現できるのかを試してみた。

4.1 EAD.xml と EAC.xml の生成

まず、多くのアーカイブズ機関などにおいてスプレッドシートに入力済みの資料目録を機械可読な形に変換する作業を行う。その際に、4 つの国際記述標準に準拠したデジタルアーカイブのアプリケーション、Access to Memory (AtoM) を利用し、AtoM に入力されているデータを EAD および EAC でコード化した XML 形式のメタデータとしてエクスポートした。

4.2 EAD.xml と EAC.xml を RDF/xml に変換

次に、ANF の RiC-O Converter (v1.0) を使って、EAD 2002 形式で書き出された XML ファイルを RiC-O 準拠の RDF/XML ファイルに変換する。

変換したい EAD.xml または EAC.xml ファイルを Github よりダウンロードした RiC-O Converter フォルダの中の「input-ead」または「input-eac」フォルダに入れる。その後、「ricoconverter」実行ファイル[9]を立ち上げ、ファイル変換に必要なコマンド[10]を実行する。

「convert_ead」コマンドを実行した場合、自動に「output-ead-8 桁の日付半角数字」のフォルダが作成され、フォルダ内に RiC-O オントロジーに変換された RDF ファイルが生成される。

また、「convert-eac」コマンドを実行した場合、同じく「output-eac-8 桁の日付半角数字」のフォルダが作成されるが、フォルダ内にはさらに「agents」「places」「relations」の 3 つのフォルダが作成され、それぞれに該当する RDF ファイルが生成される。

4.3 Omeka S への適用と視覚化

次に、ブラジルの事例に学んで RiC-O に変換された RDF/XML ファイルを Omeka S を利用し、LOD への対応を試みた。また、Ontop によって Omeka S のデータベースから RDF/Turtle アーカイブにエクスポートする工程を経て、VOS でグラフ化するという可能性を模索した。

しかし、RiC-O コンバーターで生成された RDF/XML ファイルが予想に反し Omeka S の Bulk Import モジュールで適切に処理できなかった。RDF データモデルを視覚化するウェブアプリ、zazuko Sketch[11]で実体記述と関係性などを確認したところ、変換前の EAD データの記述が不十分であることや、RDF への変換方式などに問題があるなど、変換前後の記述を再検討する必要があることがわかった。そのため、Omeka S で RiC-O オントロジーの実装とグラフによる可視化は実現できていないのが現状である。

5. 終わりに

フランスとブラジル以外に、ベルギー、ドイツ、オランダ、スイスとアメリカなどにおいても、既存のアーカイブズ記述標準から新標準の RiC に移行するための実装研究が進められている。その結果、まだ少数ながら、これまでアーカイブズが十分に取り組んでこなかったセマンティックウェブ技術と LOD 環境を活用した検索システムの成果が公開され始めている。これらの変化はアーカイブズ資料のエンドユーザーに検索性と利用可能性をより高める可能性を提示している。

今回、日本のアーカイブズ記述を RiC に変換し、検索システムに実装するまでには至らなかったため、完全なワークフローの提示ができなかった。今後は RiC-O Converter v2.0 の公開と、ICA-EGAD による RiC-AG が公開されることを期待しつつ、改めて問題を把握し、実装の検証を試みていく。

参考文献

- [1] Expert Group on Archival Description. Records in Contexts - Conceptual Model. <https://www.ica.org/en/records-in-contexts-conceptual-model> (参照 2022-09-30).
- [2] 寺澤正直. アーカイブズ記述の国際標準に関する近年の動き. アーカイブズ学研究. 2016, 25, p. 79-90; 新たなアーカイブズ記述の国際標準 Records in Context (RiC) への対応に係る課題の抽出. アーカイブズ学研究. 2017, 27, p. 4-31.
- [3] 今野創祐. 谷合佳代子報告. アーカイブズ組織化の新国際概念モデル RiC と図書館目録のマッピング: 典拠データから考える. 情報組織化研究グループ月例研究会報. 2021-04. <http://josoken.digick.jp/meeting/2021/202104.html> (参照 2022-09-30).
- [4] International Council on Archives Expert Group on Archival Description. Records in Contexts Conceptual Model Consultation Draft v0.2. 2021, 122p https://www.ica.org/sites/default/files/ric-cm-02_july2021_0.pdf (参照 2022-09-30).
- [5] Florence Clavaud and Pauline Charbonnier. Sparnatural demonstrator, Archives nationales: documentation. 2022-09-09. <https://sparna-git.github.io/sparnatural-demonstrateur-an/presentation-en.html> (参照 2022-09-30).
- [6] BRASIL. Conselho Nacional de Arquivos. Nობრადე: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Arquivo Nacional, 2006. <https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/nობრადე.pdf> (参照 2022-09-30).
- [7] Linked Open Archives. <https://www.linkedopenarchives.com> (参照 2022-09-30).
- [8] Jair Martins de Miranda. Records in Contexts (RiC): Analysis of its Application in Archives, in Light of the Linked Open Data (LOD) technologies. Acervo, Rio de Janeiro. 2021, 34, 3, p. 1-26 <https://revista.an.gov.br/index.php/revistaacervo/article/view/1745/1676> (参照 2022-09-30).
- [9] ricoconverter.bat と ricoconverter.sh が用意されている。
- [10] convert_eac, convert_eac_raw, convert_ead, test_eac, test_ead, version, help の7つのコマンドが用意されている。
- [11] SKETCH. <https://sketch.zazuko.com/> (参照 2022-09-30).



この記事の著作権は著者に属します。この記事は Creative Commons 4.0 に基づきライセンスされます(<http://creativecommons.org/licenses/by/4.0/>)。出典を表示することを主な条件とし、複製、改変はもちろん、営利目的での二次利用も許可されています。