

アーカイブ Data Report

NO. 46

(2020年9月23日)

〒500-8813 岐阜県岐阜市明德町10番地 杉山ビル5F
E-mail: shikaku@npo-nak.com URL: https://npo-nak.com

デジタルアーカイブを世界中にリンクする 2

—ADEACメタデータのNDLフォーマットへの変換とAPIによる連携—

井上 透 (岐阜女子大学)

1. メタデータの流通の流れ

ADEACシステムで使用されているファイルフォーマットやメタデータ(目録データ)は、当然のことではあるが、NDLサーチのファイルフォーマットやメタデータと異っており、そのままデータを提供しても利用できない。ADEACは、連携の前提としてNDLサーチが定めたファイルフォーマットによって、Open Archives Initiativeが定めたOAI-PMHプロトコルにメタデータを記述し変換する必要がある。以下、具体的に解説する。

2. メタデータの連携

国会図書館NDLサーチは、ブリンコアに準拠した項目(Simple-DC)のフォーマット(図1)を定めている。「和田家おうらい」を含んだADEACのファイルフォーマットを、NDLのフォーマットのどの項目に該当するかを確認し、メタデータをOAI-PMHプロトコルで記述したものに交換する、つまり機械的にマッピングするソフトが必要となる。この過程を経て、ADEACの全てのメタデータはすべて別ファイルに蓄積され、利用を待つことになる。

内容	エレメント名	説明
タイトル	Title	資料のタイトル
作成者	Creator	資料の作成に主たる責任を負う主体
主題	Subject	資料の主題を示すキーワードやフレーズ
記述	Description	内容の説明。要約、目次、自由形式の説明文など
公開者	Publisher	資料を利用可能にする責任を負う主体
寄与者	Contributor	資料への寄与に主として責任を負う主体
日付	Date	資料のライフサイクルに関連する日付
タイプ	Type	資料の性質又はジャンル
形式	Format	資料のファイルフォーマット、物理媒体等
資源識別子	Identifier	ある文脈内での資料への曖昧でない参照

情報源	Source	その資料が由来する情報資源への参照
言語	Language	資料の記述言語
関連	Relation	関連する資料
対象範囲	Coverage	資料の時間的、空間的範囲
権利	Rights	資料に関連する権利情報

図1. NDLサーチのメタデータ概要 (Simple-DCの15のエレメント)

この蓄積された情報を、API (Application Programming Interface) として国立国会図書

館に技術的にADEACのサーバー公開することによって、国会図書館のNDLサーチ側が1週間おきにメタデータを収集に行くことで、NDLサーチの膨大な最新のメタデータの一部として蓄積される。そのメタデータがNDLサーチの利用者の検索にヒットすれば、直接、ADEACシステム内の「和田家おうらい」のテキストや画像が提供され、分散型データベースとして連携することになる。このことで、小規模デジタルアーカイブ開発運用を行う組織にとって技術的に障壁がある分野横断型総合ポータルへの参画を、TRC-ADEACがアグリゲーターとして取りまとめて実現していることとなる。

3. ジャパンサーチとの連携

「和田家おうらい」を含めたTRC-ADEACのメタデータは、国会図書館NDLサーチ経由でジャパンサーチに提供されており、直接、ジャパンサーチのメタデータに変換して提供していない。ここではNDLサーチがアグリゲーターとして、ADEACのメタデータを含んでジャパンサーチと連携している。

なお、ジャパンサーチ開発時に国会図書館のNDLサーチのメタデータはモデルになったが、Simple-DC（シンプルなダブリンコア）と命名されていても、Description等に膨大な情報を求めていることから、多くのアグリゲータ、データを取りまとめて提供する機関が参加しやすいような、図2のシンプルな形式になった。

必須	ID	データベースのID	各機関で付与。ジャパンサーチ上のデータベース紹介ページのURLに使用
必須	データベース名（日/英）	データベースの名称（日本語/英語）	
必須	データベース名（ヨミ）	データベースの名称のヨミ（全角カナ）	DB一覧の表示順に使用
	データベースの説明（日/英）	データベースの説明（日本語/英語）	100字以上は折り畳み表示
必須	カテゴリ	データベースが扱うコンテンツの種別。選択式	データベース当たり一つが推奨だが、複数選択可
	サブカテゴリ	データベースが扱うコンテンツの種別。選択式のカテゴリを補うための情報。テキストによる自由記入。	
	メタデータの権利表示	メタデータの権利情報/二次利用条件について、例外の場合のみ記載。原則はCC0（著作権のある項目のみCC BYでも可だが、その場合も項目名とCC BYの記載が必要。）	自由記入。クリエイティブ・コモンズライセンス、政府標準利用規約等。権利情報について記述されている外部資源へのリンク（URL）でも可
	サムネイル画像の権利表示	対象のサムネイル画像がある場合のその権利情報/二次利用条件	
	コンテンツの権利表示	対象のデジタルデータがある場合の権利情報/二次利用条件。例外がある場合は（メタデータの）共通項目で定義することが可能。	
	コンテンツの権利区分	コンテンツの権利情報/二次利用条件だが、検索・絞り込み用に選択式になっている	
	データベースのURL	データベースのURL	
	提供機関（データベースの管理者）	データベースの管理者（オーナー）が自動で入る	
	コンテンツ公開状況	デジタルコンテンツのアクセス範囲を選択式（「ウェブ公開」「限定公開」「デジタルコンテンツなし」）で設定。アイテムにより公開状況が異なる場合は（メタデータの）共通項目で定義することが可能。	
	代表画像	データベースの代表画像	

図2 ジャパンサーチ・データベースの基本情報・権利表示の設定

4. 小規模デジタルアーカイブの連携策

NDLサーチ、ジャパンサーチはアグリゲーター経由でのデジタルアーカイブ連携を求めており、小規模な大学や地域のデジタルアーカイブを直接、接続するのは困難になっている。したがって、デジタルアーカイブを世界中でリンクさせるには、アグリゲーターとしての機能を持つ大規模な大学、国の研究機関やTRC-ADEACと連携せざるを得ない。

しかし、分野横断型ポータルに接続しなくとも、オープンデータとして単独でメタデータをAPI公開することにより他機関、ボランティアがアプリを開発し、社会に有用なサービスとしてデータを提供することが実現している。したがって、広くデータが利用されるためには、デジタルアーカイブの開発条件として、オープンデータ化は必須ではなからうか。

参考：「WebAPIによるシステム連携ガイドライン ver.1.4(2018.12.28)」

https://iss.ndl.go.jp/information/wp-content/uploads/2019/01/WebAPIguideline_ver.1.4_20181228.pdf