

地域デジタルアーカイブ CMS の構築と 歴史資料の LOD 化による活用

奥野 拓

公立はこだて未来大学
システム情報科学部

山田 亜美

公立はこだて未来大学大学院
システム情報科学研究科

高橋 正輝

エヌ・ティ・ティ・
アドバンステクノロジー株式会社

川嶋 稔夫

公立はこだて未来大学
システム情報科学部

公立はこだて未来大学では、函館市中央図書館と連携して歴史資料の登録から公開までの業務をサポートするデジタルアーカイブ CMS を開発し、デジタル資料館として運用している。蓄積された膨大な歴史資料のファインダビリティを改善し、多目的に活用する目的で、LOD 化に取り組んでいる。他の歴史コンテンツと関連付けるため、日付、人物、場所をリソースとして抽出し、絵葉書資料 RDF データセットを作成している。また、歴史的エピソードを仲介させることにより、間接的に関連する絵葉書資料を閲覧できるシステムを構築している。さらに、デジタル資料館の歴史資料を引用して特集コンテンツを作成するデジタルテーマ展や、市民が歴史資料にアノテーションを行う市民編纂 SNS、様々な歴史資料 RDF データセットを利用する観光アプリの構築を検討している。

Development of Regional Digital Archives CMS and Utilization of Historical Records by Making LOD

Taku Okuno

School of Systems
Information Science
Future University Hakodate

Ami Yamada

Graduate School of Systems
Information Science
Future University Hakodate

Masaki Takahashi

NTT Advanced Technology
Corporation

Toshio Kawashima

School of Systems
Information Science
Future University Hakodate

Future University Hakodate, in cooperation with Hakodate City Central Library, has developed a digital archives CMS. For improving the findability of vast numbers of historical records, LOD approach is adopted. For associating the records with other historical contents, dates, persons, and places have been extracted as resources from the records, and a picture postcard RDF dataset has been developed. Through the intermediation of historical episodes, users can browse indirectly related picture postcards. The applications of the digital archives include: the digital special exhibits that cites the archives' contents, the digital social editing site on which people make annotations to the contents of the archives, and the tourist's applications that utilize some RDF datasets of historical records.

1. まえがき

近年、地域の歴史資料のデジタルアーカイブ化が広く進められている。デジタルアーカイブを構築する際には、貴重な歴史資料の保全だけでなく、蓄積された膨大な歴史資料をいかにして活用するかという点についても考慮が必要である。函館市では、函館市中央図書館と公立はこだて未来大学の連携によりデジタルアーカイブを構築し、歴

史資料のデジタル化、蓄積および公開を進めてきている。学術目的では全国から多くの利用があるが、一般市民や観光客にとって価値のある歴史資料が多数含まれているにもかかわらず、ファインダビリティの点では十分ではなく、十分に活用できる状態にはなっていないのが現状である。

公立はこだて未来大学では、地域の歴史資料を市民文化活動や観光などに活用することを目的

表1 デジタル資料館で公開中の資料・画像枚数
(2014年10月22日現在)

Table 1 The numbers of documents exhibited in the digital archives as of Oct. 22, 2014.

資料形態	資料点数	画像枚数		
		公開	館内限定	合計
古文書・地図	1,005	76,791	35	76,826
絵葉書	3,637	15,816	2,173	17,989
ポスター	1,658	1,776	0	1,776
軸装・額装等	17	92	0	92
写真	258	1,468	0	1,468
合計	6,575	95,943	2,208	98,151

として、デジタルアーカイブや Web で公開されている様々な歴史コンテンツを LOD (Linked Open Data) 化し、連係させることにより活用する仕組み作りに取り組んでいる。本稿では、始めに継続的に取り組んでいるデジタルアーカイブの構築と運用について説明する。次に歴史資料の LOD 化とアプリケーションによる活用の取り組みについて説明する。

2. デジタルアーカイブ CMS

2.1 函館市におけるデジタルアーカイブ

函館市の図書館は多くの歴史資料を所蔵しており、「北方資料の宝庫」と呼ばれてきた。しかし、二次資料の作成は遅れており、資料保存の観点では整備が不十分であった。函館市中央図書館では、情報通信設備を備えた現在の施設への移転に合わせて、2003年度から歴史資料のデジタルアーカイブ化を開始した。資料の画像データ化は、利用が多く劣化が進んでいた古地図と古文書から進められた。写真・絵葉書については、研究対象として公立はこだて未来大学が主導して画像データ化が行われた。また、ポスターや軸装・額装資料なども画像データ化の対象とされた。これらの資料は、図書館で扱う通常の書籍とは形態が大きく異なり管理が難しい。また、貴重な資料を含むため、閲覧には注意を要する。そのため、デジタルアーカイブによる一般公開が期待されていた。

2007年度より、函館市中央図書館と公立はこだて未来大学の連携により、それまでに蓄積されてきた画像データを管理・公開するデジタルアーカイブ CMS (Content Management System) の構築を開始し、2008年度に「デジタル資料館」として運用を開始した[1]。2010年度には公立はこだて未来大学内にデジタルアーカイブ研究センターが設置され、デジタルアーカイブ CMS の継続的な開発とデジタル資料館の運用を行っている。

表2 デジタル資料館の2013年度アクセス統計
Table 2 Access statistics of the digital archives in FY 2013.

サイト全体		形態別ページビュー	
セッション	28,849	古文書・地図	63,421 (49.7%)
ユーザ	17,016	絵葉書	38,415 (30.1%)
ページビュー	417,477	写真	12,980 (10.2%)
ページ/セッション	14	軸装・額装等	2,001 (01.6%)
平均セッション時間	06:25	ポスター	1,700 (01.3%)
都道府県別ページビュー		市区町村別ページビュー	
1 北海道	13,908 (49.6%)	1 函館	6,601 (23.5%)
2 東京	5,548 (19.8%)	2 札幌	6,409 (22.9%)
3 神奈川	1,211 (04.3%)	3 港区	1,289 (04.6%)
4 大阪	1,046 (3.7%)	4 横浜	906 (3.23%)
5 埼玉	805 (2.9%)	5 渋谷	839 (3.00%)

2.2 デジタルアーカイブ CMS の構築

デジタルアーカイブ CMS は、長期に渡って低コストで保守運用と機能拡張を継続する必要がある。そのため、開発チームを多学年の学生から構成し、年度を跨いだ技術移転を行えるようにしている。また、利用技術についても、メンバーの入れ替わりがあっても引き継ぎが容易になるように、広く普及し、ノウハウが豊富な技術を中心に採用している[2]。

CMS の基盤として、実績の豊富な Web アプリケーションフレームワークである CakePHP を用いている。データベースには、任意の目録項目を持つ資料形態の追加に柔軟に対応するため Key-Value Store 型の MongoDB を用いている。また、画像ビューワとして、多段階の解像度に分割した画像を動的に再構成し、シームレスにズーム可能な Zoomify を用いている。基本的には PC からの閲覧を想定しているが、2014年度からの館内におけるデジタル資料館閲覧・印刷用端末の iPad 化に合わせて、タブレットデバイスからの閲覧および印刷に対応させている[2]。

2.3 デジタル資料館の概要

2014年10月の時点で、デジタル資料館には、古文書・地図、絵葉書、ポスター、軸装・額装等、写真の資料形態(カテゴリ)に分類された資料 6,575 点、画像数としては 98,151 枚が公開されている。2014年9月の改修により、館内限定公開の資料についてもデジタルアーカイブ CMS で扱えるようになっている。

資料形態ごとの資料点数、画像枚数の内訳を表1に示す。古文書・地図の画像枚数が資料点数と比較して極端に多いのは、数百ページにわたる文書が多数含まれ、1ページ毎に1枚の画像として掲載されているためである。2013年度は月に100点のペースで資料が追加され、2014年度の上期は、館内限定公開資料を中心に300点程度の資料が追加された。

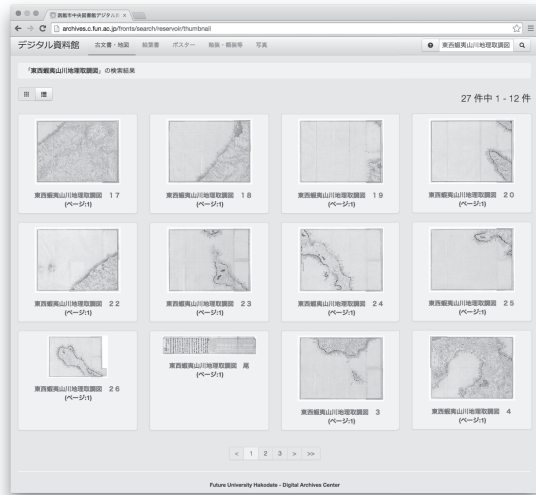


図1 デジタル資料館の資料一覧画面(「古文書・地図」タブのサムネイル表示)
Figure 1 A listing view of the digital archives (thumbnail view of “historical documents and maps” tab)

2013年度のデジタル資料館へのアクセス統計を表2に示す。サイト全体では約42万ページビューであり、ユーザは約1万7千人であった。形態別では、半数が古文書・地図であり、続いて、絵葉書が約30%、写真が約13%であった。地域別では、半数が北海道外からのアクセスであり、全国からの利用がある。

デジタルアーカイブCMSにより運用中のデジタル資料館の資料一覧画面を図1に示す。上部のタブで切り替えることにより、表1で示した資料形態毎に資料のサムネイル画像が一覧できる。左上部のボタンでサムネイル表示と画像付き目録リスト表示を切り替えることができる。図1は、「東西蝦夷山川地理取調図」でキーワード検索した結果の画面である。

資料一覧画面から画像を選択すると、資料詳細画面に遷移する。資料詳細画面には資料画像と目録データが表示される。資料画像はズーム可能なビューワにより表示される。図1の右下角の画像を選択して遷移した画面が図2である。画像は松浦武四郎による「東西蝦夷山川地理取調図」の一部であり、北海道の噴火湾(内浦湾)の周辺を含む古地図である。この資料の左端の大沼・駒ヶ岳付近を最大ズームインした画面が図3である。微細な文字や地形を表現するテクスチャを鮮明に読み取ることが可能である。

2.4 資料データ作成・出力・登録業務フロー

函館市中央図書館では、歴史資料のうち古文書・地図については、一般の書籍と同様にOPACにより管理している。しかし、書籍と同様の管理が困難な形態である絵葉書、ポスター、軸装・額装等、写真については、デジタル資料館での公開を目的としてOPACとは独立したスタンドアロ

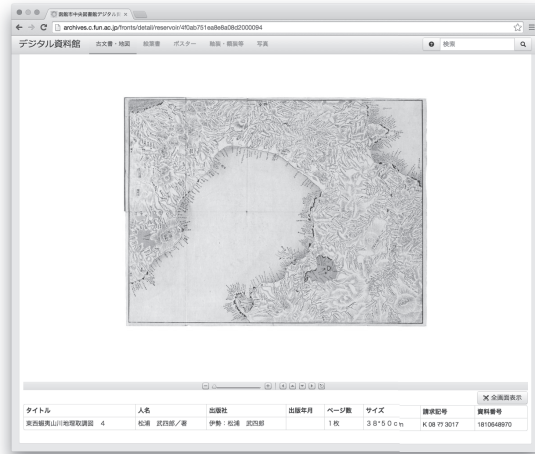


図2 デジタル資料館の資料詳細画面(最大ズームアウト表示)
Figure 2 A detail view of the digital archives (most zoomed out).

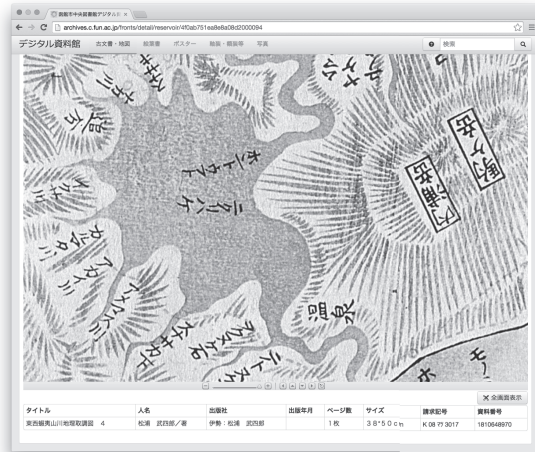


図3 デジタル資料館の資料詳細画面(最大ズームイン表示)
Figure 3 A detail view of the digital archives (most zoomed in).

ーンのデータベース管理ソフトウェアにより管理している。現在は、資料の撮影から画像データに対応付けられた目録データの作成までを図書館で行い、デジタル資料館への画像データと目録の登録については大学で行っている。これらの業務フローを図4と図5にそれぞれUMLのアクティビティ図により示す。

データ作成のプロセス(図4上側レーン)では、最初に図書館職員が資料を撮影し、資料画像に一意な資料番号を付与する。次に、分割撮影した画像の接合、トリミングなどの補正を行い、図書館内LANに接続されたNASに保存する。また、資料番号をキーとして目録情報を作成する。古文書・地図についてはOPACに、他の形態の資料については、非OPAC資料データベースに目録情報(メタデータ)を登録する。なお、通常

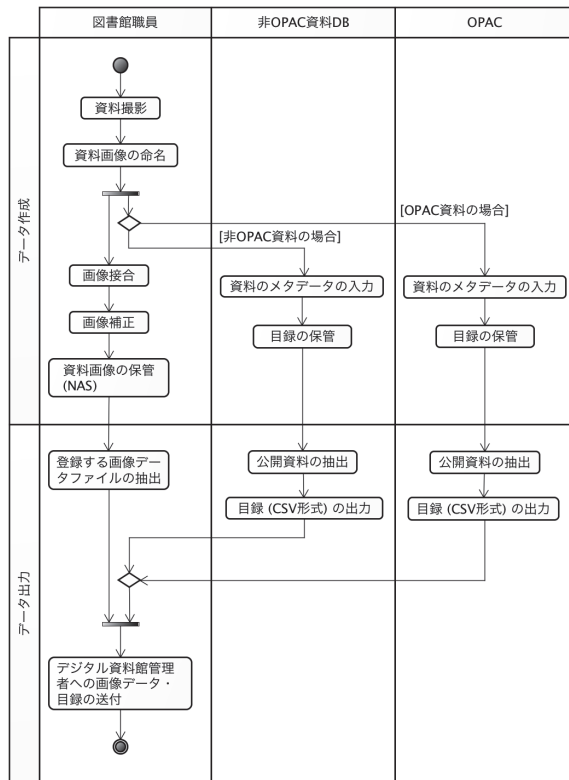


図 4 資料画像・目録データの作成・出力業務フロー

Figure 4 Workflow of making and exporting images and catalog data of documents.

は、ここまでのプロセスを一定数の資料に対して繰り返し行った後、デジタル資料館への登録のためのデータ出力のプロセスに移る。

データ出力のプロセス(図4下側レーン)では、最初に非 OPAC 資料データベースと OPAC から、それぞれ公開する資料を抽出する。抽出された資料の目録データを CSV 形式で出力する。同時に、新たにデジタル資料館に登録する画像データファイルを NAS から抽出する。これらをまとめてポータブル HDD にコピーし、大学のデジタル資料館管理者に送付する。

データ登録のプロセス(図5)では、図書館から送付された画像データを、NAS と外付け HDD に、目録データを Web ストレージにそれぞれ保管する。次に、ターミナルから画像データをデジタル資料館サーバにアップロードし、Web ブラウザでデジタル資料館の管理画面にアクセスし、Zoomify 形式への変換処理を行う。変換処理の完了後、分割画像をデジタル資料館サーバと NAS に保存する。その後、管理画面より CSV 形式の目録データをインポートすると、目録データと画像データが関連付けされる。最後に、グルーピングされた資料の代表サムネイル画像を設定するスクリプトをデジタル資料館サーバで実行する。以上のプロセスによりデジタル資料館に資料が

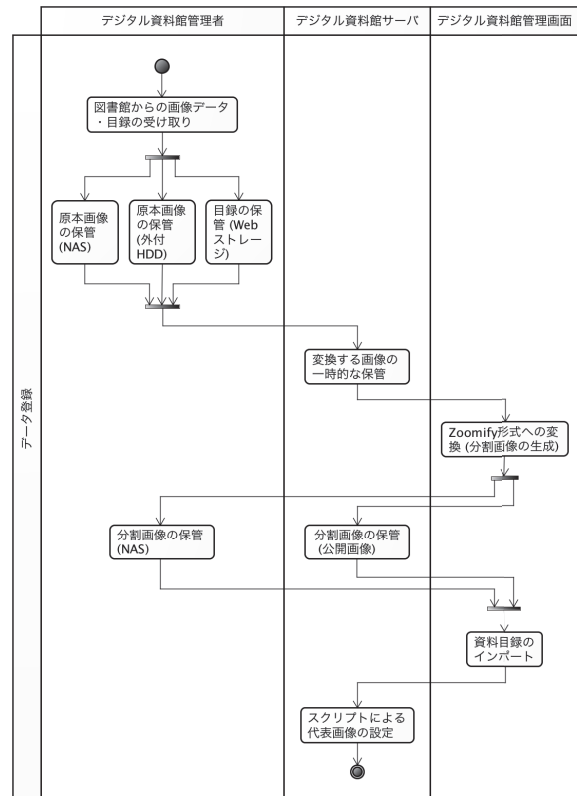


図 5 資料画像・目録データのデジタル資料館への登録業務フロー

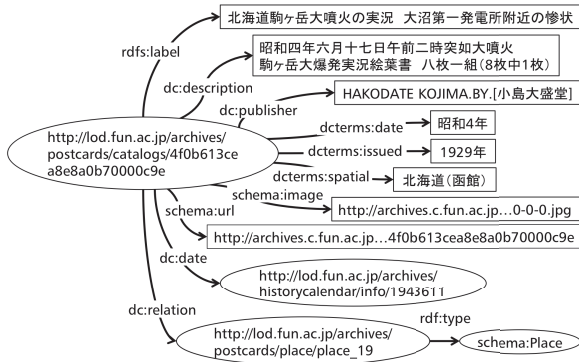
Figure 5 Workflow of registration of images and catalog data of documents to the digital archives.

公開される。

2.5 業務フローにおける課題

2.4 で示したデータ作成から登録までの業務フローについては、現時点でいくつかの課題がある。デジタルアーカイブ CMS は、最終的には資料の画像データ作成からデジタル資料館での公開までを、完全に図書館職員が行えるようなシステムとなることを目指している。しかし、現状では、デジタル資料館サーバにターミナルから接続して行わなければならない作業がいくつか残っており、完全に管理画面からのみの操作による処理は実現できていない。

また、資料をデジタル資料館に登録した後の目録データの修正についても課題がある。現在は、OPAC 資料と非 OPAC 資料のいずれについても、図書館のデータベースのデータがマスターであり、デジタル資料館のデータはコピーという扱いである。目録に修正が必要な場合は、OPAC または非 OPAC 資料データベース上で修正し、形態毎に全資料の目録データが含まれる CSV ファイルを出力し、デジタル資料館にインポートする必要がある。



xmlns:rdfs="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
 xmlns:dc="http://purl.org/dc/elements/1.1/"
 xmlns:dcterms="http://purl.org/dc/terms/"
 xmlns:schema="http://schema.org/"
 xmlns:foaf="http://xmlns.com/foaf/0.1/"

図6 絵葉書資料 RDF データモデル
Figure 6 RDF data model of picture postcards.

非 OPAC 資料については、現在スタンドアロンのデータベースアプリケーションにより管理されているため、同時並行作業ができないという問題もある。

このような現状を改善するためには、デジタル資料館の目録データをマスターとして一元管理することが望ましい。しかしデータベースでは資料の寄贈者等の個人情報も管理されているため、館外での管理については慎重に検討する必要がある。

一方、OPAC 資料については、OPAC とデジタル資料館で重複した目録の管理が不可避であるが、OPAC 上の目録の変更を自動的にデジタル資料館に同期するような仕組みについては検討の余地がある。

以上の課題については、今後、図書館との協議により解決を図っていく予定である。

2.6 資料閲覧における課題

デジタル資料館の主な利用は学術目的であり、出版・報道業者による利用も多い。しかし、インターネット経由で無償アクセス可能な歴史資料には、これらの用途に留まらない潜在的利用価値がある。特に、絵葉書や写真については、地域の歴史的建造物や歴史的エピソードをビジュアルに伝える貴重な資料である。このような歴史資料は市民の文化的活動において有用であるだけでなく、歴史的要素が観光資源として重要である函館のような観光都市においては、観光資源の魅力を引き出す貴重な材料となり得る。

一方で、一般にデジタルアーカイブから目的の資料を探し出すのは容易ではない。デジタル資料館では、画像資料を探す方法として、一覧画面よりサムネイル画像をブラウズするか、目録テキストの全文検索を行うことが可能であるが、これらの方法は膨大な数の資料を収蔵するデジタルアーカイブから興味のある資料を見つけ出す方法

としては不十分である。そこで、本取り組みでは、既にデジタル化されて Web 等で公開されている他の歴史資料や歴史コンテンツを利用し、それらに共通に含まれるキーワードにより歴史資料を意味的に結びつけ、関連する資料を抽出する仕組みの構築を試みている。以降ではその取り組みについて述べる。

3. 歴史資料の LOD 化

3.1 リソース抽出

デジタル資料館の目録に含まれる項目は資料形態により異なるが、絵葉書、写真、ポスターについては、ほぼ共通した項目が与えられている。それらのうち、他の歴史資料と意味的に関連付け可能な項目は、タイトル、内容説明、西暦、和暦である。しかし、内容説明に記載されている内容は、資料の枚数、付属物等の構成、タイトルと重複した説明記述が多く、被写体の説明としては不十分である。

一方、公立はこだて未来大学では、函館市において Web 等で公開されている地域史、観光などの様々なコンテンツを RDF データセット化し、LOD としての公開を目指す活動を行っている。現在、RDF 化を進めている歴史コンテンツはテキスト形式の「函館市史年表編」と「はこだて人物誌」である。

函館市史は、函館の歴史をまとめた全 7 編 11 巻から構成される書籍であり、年表編を含む一部がデジタル化されている [3]。年表編は主に新聞記事を基礎資料として作成され、個々の出来事が簡潔にまとめられている。また、出来事には基本的に日付が付与されているため、他の歴史資料と日付で関連付けることで有効活用できると考えられる。今回は他資料と関連する可能性が高い昭和期の 6,808 件の出来事を含む RDF データセットを作成した。

はこだて人物誌は、函館市文化・スポーツ財団の月刊広報誌「ステップアップ」に連載中の「函館ゆかりの人物伝」を Web コンテンツ化したものである [4]。函館に縁のある人物について、生い立ちから亡くなるまで、年代ごとに体験した出来事が記述されている。今回は、2013 年 9 月時点で公開されていた 271 人のコンテンツを含む RDF データセットを作成した [5]。

これらの RDF データセット作成に当たり、相互に関連付けるためのリソースとして、日付、人物、場所をテキストより抽出することとした。日付は年表の出来事と関連付ける目的で使用し、人物ははこだて人物誌の人物と関連付ける目的で使用し、場所は、歴史資料を観光資源として活用することを想定し、観光スポットとの関連付けに使用する。

これら 2 種類の RDF データセットをデジタル資料館の歴史資料の目録と関連付けることにより、史実が補完され、ファインダビリティの向上



図7 「エピソードでつながる函館歴史写真」の資料詳細画面（関連画像一覧表示）

Figure 7 A photo detail view of “The Historical Photographs of Hakodate Linked by Episodes” (with a catalog view of related photos).

が期待できる。画像資料に関連付けられたメタデータとしてテキスト検索の対象が広がるという効果もあるが、目録だけでは直接的に関連付けることが困難な複数の画像資料を、歴史上の人物に関する記述や歴史年表のエピソードを仲介させることで、関連付けることが可能となる。

次節では、これらの関連付けを行うために行った、デジタル資料館の絵葉書資料の RDF データセットの作成について述べる。

3.2 絵葉書資料 RDF データセットの作成

デジタル資料館の画像資料の中でも、絵葉書資料はイベント等を記念して作成されたものが多いため、一枚毎に歴史的エピソードと関連付く可能性が高いと考えられる。そこで、最初に絵葉書資料 13,667 件の RDF データ化を行った。最初の作業として、目録から日付、人物、場所をリソースとして抽出した。

日付は、目録の「西暦」、「和暦」の項目に年のみが記されているが、「タイトル」と「内容説明」の項目に年月日が記されている場合がある。これらより、漢数字の「年号年月日」形式の文字列を抽出した。なお、写真に関連する出来事、印刷、発行、検閲など、意味の異なる複数の日付が一つの資料に含まれる場合は、写真に関連する出来事の日付を用いることとした。以上の作業により日付が抽出された絵葉書資料は 338 件である。

人物は、はこだて人物誌 RDF データセットに含まれる人物名 271 人分を用いて抽出した。人物が抽出された絵葉書資料は 461 件である。

場所の抽出は、歴史資料を観光資源として活用することを想定し、観光スポットを対象として行った。観光スポットは、函館地域の観光スポットを網羅的に掲載している函館市公式観光情報サイト「はこぶら」[6]から抽出した。歴史的建造物などを含む「見る」カテゴリの観光スポット 198 件を用いた。場所が抽出された絵葉書資料は 805 件である。

絵葉書資料 RDF データセットは、目録に含まれるタイトル、内容説明、印刷発行、和暦、西暦、地域分類に加え、サムネイル画像、資料画像の URL、そして、上記の方法により抽出した日付、人物、場所を含むようにモデリングを行った。モデリングは、国立国会図書館ダブリンコアメタデータ記述 (DC-NDL) を参考にして行った。作成した RDF データモデルを図 6 に示す。

作成したデータセットは、作業開始時点でデジタル資料館に登録されていた絵葉書資料 13,667 件のデータから構成される。このうち人物、場所、日付のいずれかを含む絵葉書資料は 1,521 件であり、絵葉書資料全体の 11% となった。

3.3 歴史的関連性に基づく資料閲覧

作成した地域史 RDF データセットを用いて、ユーザが写真間の歴史的関連性に基づき、興味のある写真を選択していくことで発見的に閲覧することができる Web アプリケーション「エピソードでつながる函館歴史写真」を構築している。このアプリケーションでは、絵葉書資料 RDF データセットと函館市史年表編 RDF データセットを使用している。

最初に興味のある写真 1 枚をサムネイル画像一覧から選択すると、写真とその目録、その写真に関連する年表のエピソードが年代順に表示さ

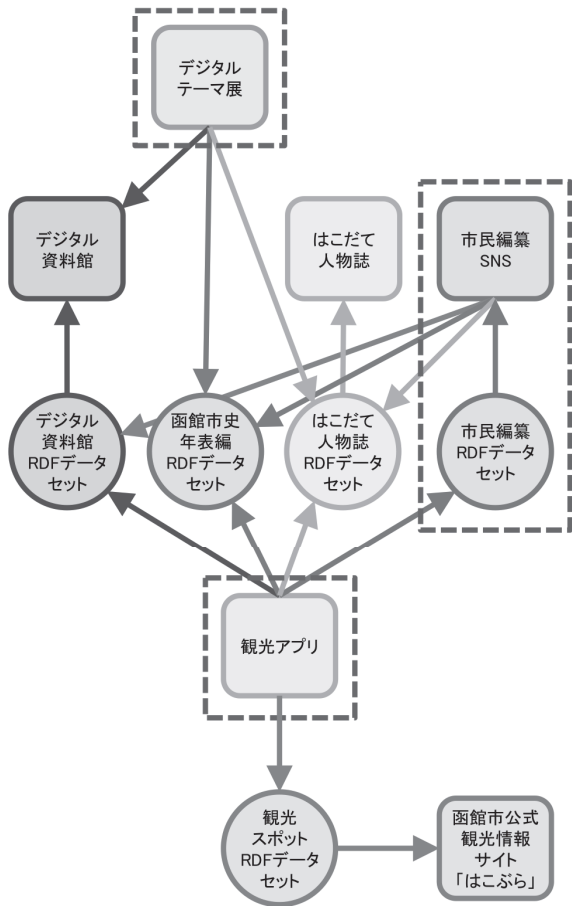


図 8 函館地域史 RDF データセットと Web コンテンツおよびアプリケーションとそれらの間のデータ参照関係

Figure 8 Hakodate historical records RDF datasets, Web contents and applications, and data reference relationships among them.

れる。ここで、エピソードに関連する写真が存在する場合、すなわち、写真とエピソードを関連付けるリソースが存在する場合にはリソースの文字列が水色に着色されて表示される。その場合、エピソードを選択すると、関連付けられた写真の一覧がエピソード直下に展開表示される(図7)。一覧の中から1枚の写真を選択すると、左側に最初に選択した写真、右側に関連写真、下にはそれらを関連付ける年表のエピソードが表示される。

図7は「(北海道公園) 大沼より駒ヶ岳を望む View of Onuma Hokkaido park」というタイトルの絵葉書を選択した例である。タイトル中に「駒ヶ岳」という地名が含まれているため、函館市史年表編の「昭和4年6月17日 駒ヶ岳が爆発し、近郊の8町村に総額838万円の大きな被害がでる。(以下略)」という歴史のエピソードが提示される。さらに、「昭和4年6月17日」という日付から、この日付が含まれる写真が多数表示される。提示された写真から、最初に選択した写真の被写体となっている駒ヶ岳の噴火の様

子などを記録した様々な写真を見ることができ

3.4 函館歴史資料 LOD によるデジタル資料館の活用

本取り組みで作成している地域史 RDF データセットを LOD として公開し、それらを用いてデジタル資料館の歴史資料を活用するいくつかのアプリケーションを検討している。RDF データセット、Web コンテンツ、アプリケーションと、それらの間のデータ参照関係を図8に示す。破線で囲まれた部分が検討中のものである。

検討中のアプリケーションのうち、「デジタルテーマ展」(仮称)についてはすでにプロトタイプシステムの開発を行っている。デジタルテーマ展は、学芸員が特定のテーマを設定して、デジタル資料館より資料を引用し、解説文を加えてコンテンツ化し、公開する Web サイトであり、汎用のオープンソース CMS (Drupal) をベースに構築している。試験運用コンテンツとして、函館山周辺の古地図を中心に歴史的エピソードを交えて解説する「古地図資料館」、松前藩家老であり「夷酋列像」で知られる画家の蠣崎波響を特集した「蠣崎波響特別展」などが作成されている。このようなテーマ展コンテンツにおいて、関連する函館市史年表編のエピソードやはこだて人物誌の人物紹介を参照可能とすることにより、コンテンツに広がりを持たせることが可能になると考えられる。

デジタル資料館の目録における内容記述を補い、歴史資料を市民の文化活動や観光へ活用することを目的とする Web サイトが「市民編纂 SNS」(仮称)である。これは、デジタル資料館の歴史資料に市民が様々な形でアノテーションを行えるシステムである。アノテーションの方法としては、目録の項目への追記、目録の記述へのコメント、画像中の被写体(矩形領域)へのコメント、画像中の被写体と他の画像またはその被写体との関連付け、函館市史年表編およびはこだて人物誌 RDF データセットとの関連付け、関連画像の投稿などを検討している[7]。

この仕組みで収集したアノテーションより「市民編纂 RDF データセット」(仮称)を生成する。このデータセットの情報は、学芸員により編集された目録のような学術的価値や信頼性は担保されないが、市民文化活動や観光において利用価値がある[8]。

歴史資料の観光における活用法としては、観光中にスマートフォン等を用いて観光スポットに関する情報を入手するための観光アプリでの利用を想定している。観光スポットの基本情報は、3.2で示した絵葉書資料からの場所(観光スポット)の抽出に用いた観光情報サイトのコンテンツから作成した観光スポット RDF データセットを用いる。このデータセットについては、公立はこ

だて未来大学のプロジェクト型演習 (PBL) において開発した Web API を用いて作成している。また、このデータセットを用いて、電子マップ上に現在位置と観光スポットを表示し、近隣の観光スポットの情報を入手できる iPhone アプリについても、同演習で開発し一般公開している[10]。このアプリをベースとして、市民編纂 RDF データセットを含む 4 種類の歴史資料 RDF データセットから、場所 (観光スポット) をキーにしてデータを取得して表示するように拡張する。

4. あとがき

本取り組みでは、デジタルアーカイブ CMS の構築・運用と連携して、地域の様々な歴史資料を LOD として公開し、市民文化活動や観光などの分野で活用することを目指した活動を行っている。

本報告では、デジタルアーカイブ CMS により運用されているデジタル資料館の成り立ちと概要、資料の公開までの業務フローについて説明し、業務フローおよび資料の閲覧における課題について述べた。そして、閲覧時のファインダビリティの問題を解決する方法として、LOD による他の歴史資料との関連付けによるアプローチについて説明した。特に、日付、人物、場所に着目した歴史資料からのリソース抽出と、デジタル資料館の絵葉書資料の RDF データセット作成方法について述べた。さらに、それを函館市史年表編 RDF データセットと組み合わせる歴史的エピソードで関連付けて閲覧する Web アプリケーションを紹介した。最後に、作成した RDF データセットを利用してデジタル資料館の資料を活用するいくつかの試みと構想について説明した。

現時点では、デジタル資料館の歴史資料の RDF データ化については絵葉書資料に留まっているが、今後、写真やポスターなど観光資源として重要な資料に範囲を広げていく予定である。一方、関連付ける函館市史年表編については、現在、昭和期に限定して RDF データセット化しているが、写真資料が多数残存し、絵葉書ブームや函館市史において重要な役割を果たした「函館四天王」の活躍時期にも重なる明治・大正年代の年表を追加することにより、デジタル資料館の歴史資料との関係強化を図る。さらに、より効果的なデータ関係に向けて、非構造化テキストから日付、人物、場所以外の様々な意味的関連性を抽出することについても検討する。

参考文献

- 1) デジタル資料館, 函館市図書館所蔵デジタルアーカイブ, 入手先 [\(http://archives.c.fun.ac.jp/\)](http://archives.c.fun.ac.jp/) (参照 2014-11-06).
- 2) 出口貴也, 中原裕成, 高橋正輝, 奥野拓, 川嶋稔夫: 地域の記録と市民の記憶を共有するディ

ジタルアーカイブ CMS, 情報処理学会研究報告, Vol. 2011-DD84 (2011).

3) 『函館市史』デジタル版, 入手先 [〈http://www.lib-hkd.jp/hensan/hakodateshish/shishi_index.htm〉](http://www.lib-hkd.jp/hensan/hakodateshish/shishi_index.htm) (参照 2014-11-06).

4) はこだて人物誌, 入手先 [〈http://www.lib-hkd.jp/hensan/jimbutsu_ver1.0/〉](http://www.lib-hkd.jp/hensan/jimbutsu_ver1.0/) (参照 2014-11-06).

5) 山田亜美, 高橋正輝, 奥野拓: 函館地域史データの観光資源化に向けた地域史 LOD の構築-人物誌を対象とした文章からの歴史情報抽出と観光情報との関連付け-, 情報処理北海道シンポジウム 2013 講演論文集 (2013) .

6) 函館市公式観光情報サイト「はこぶら」, 入手先 [〈http://www.hakobura.jp/〉](http://www.hakobura.jp/) (参照 2014-11-06).

7) 高橋正輝, 奥野拓: 観光情報サービスでの活用に向けた地域写真アーカイブにおける建造物の関連付け手法, 観光情報学会第 5 回研究発表会講演論文集, pp. 1-5 (2012)

8) 高橋正輝, 奥野拓, 川嶋稔夫: 函館の歴史資料を用いた地域写真アーカイブの編纂, 情報処理学会研究報告, Vol.2013-DD88 No.9 (2013) .

9) 函館観光情報の有効活用のための Web API とアプリの開発, 公立はこだて未来大学プロジェクト学習成果報告書, No. 15, [〈http://www.fun.ac.jp/sisp/report/report_2012.html〉](http://www.fun.ac.jp/sisp/report/report_2012.html) (参照 2014-11-06).

10) FUTURE UNIVERSITY HAKODATE, はこだて Map+ ~まちあるきで発見!おすすめ観光コース~, 入手先 [〈https://itunes.apple.com/jp/app/id744881889〉](https://itunes.apple.com/jp/app/id744881889) (参照 2014-11-06).