

デジタルアーカイブの持続可能性条件の研究 ～自然科学分野のデジタルアーカイブに関する通時的分析～

Research on regional revitalization through digital archives of regional materials

深瀬香緒里*1 久世 均*2

1990年代、先進国を中心に国家レベルで情報基盤構想が注目され、インターネット環境の開発・整備がすすめられた。その結果、所蔵資料をデジタル化してインターネット上で提供する図書館・博物館・美術館等の文化施設が急増した。しかし、陸前高田震災アーカイブ NAVIのように、東日本大震災アーカイブの中にはすでに閉鎖され、持続可能性の課題が明らかになってきた。

本研究では、デジタルアーカイブの持続可能性を高めるために必要な条件のひとつは利活用にあるとの仮説を立て自然科学分野のデジタルアーカイブに関する通時的分析をしたので報告する。

<キーワード> デジタルアーカイブ, 持続可能性, 自然科学分野, TRMM 衛星

1. はじめに

1990年代、先進国を中心に国家レベルで情報基盤構想が注目され、インターネット環境の開発・整備がすすめられた。その結果、所蔵資料をデジタル化してインターネット上で提供する図書館・博物館・美術館等の文化施設が急増した。2000年代に入ると、日本においても、アジア歴史資料センター、文化遺産オンライン等のデジタルコンテンツが提供された。

2011年の東日本大震災以降、民間企業、研究機関、地方自治体など様々な組織により、被害写真などの震災記録を収集した震災アーカイブが作られた。翌年には、総務省の「東日本大震災アーカイブ」基盤構築プロジェクトが開始し、『震災関連デジタルアーカイブ構築・運用のためのガイドライン』（総務省 2013）が整備された。震災アーカイブを一元的に検索できる国立国会図書館東日本大震災アーカイブ（ひなぎく）が公開されるなど、デジタルアーカイブを取り巻く環境は大きく進展した。

2014年に大規模学術フロンティア促進事業に採択された国文学研究資料館（2020）の「日本語の歴史的典籍の国際共同研究ネットワーク構築計画（略称：歴史的典籍NW事業）」は、唯一の人文社会科学分野プロジェクトである。10年かけて日本語の歴史的典籍をデータベース化しているが、収集範囲は人文科学分野に留まらなく、医学・数学等の自然科学分野も対象としており、分離融合の研究が期待されている。その後、2020年8月にはジャパンサーチ正式版が公開されるなど、日本におけるデジタルアーカイブは第二次黄金期に入っているとされる

しかし、陸前高田震災アーカイブ NAVIのように、東日本大震災アーカイブの中にはすでに閉鎖され、2000年代前半の第一次黄金期に発現した持続可能性の課題は残ったままといえる。

2. デジタルアーカイブの持続可能性

本研究の目的は、デジタルアーカイブの

*1 FUKASE, Karina 岐阜女子大学大学院 *2 KUZE, Hitoshi 岐阜女子大学

持続可能性を高める条件について考察することである。

事例研究を行う目的は、先行する自然科学分野の成功事例を分析することで、デジタルアーカイブの持続可能性を高めるために必要な条件を一般化することである。

まず、本研究では、デジタルアーカイブの持続可能性を高めるために必要な条件の1つは、利活用にあると仮説を立てて調査を開始した。

(1) デジタルアーカイブの定義

本研究におけるデジタルアーカイブの定義は、『我が国におけるデジタルアーカイブ推進の方向性』(2017)に基づく。「様々なデジタル情報資源を収集・保存・提供する仕組みの総体」であり、「デジタルアーカイブで扱うデジタル情報資源は、「デジタルコンテンツ」だけでなく、コンテンツの縮小版や部分表示である「サムネイル/プレビュー」のほか、(アナログ媒体の資料・作品を含む)コンテンツの内容や所在に関する情報を記述した「メタデータ」も対象」となる。

デジタルアーカイブは、1994年頃に月尾嘉男氏が提起した言葉とされる。一般向けのデジタル事典(日経BP社2001)では、用語を「デジタル・アーカイブ」としており、「收藏品や資料などをデジタル的に保管し、資料間の関連などを保持した形で管理されたものを指す。これは文字、画像情報、3次元情報、音、化学式、動画など各種の情報が含まれたマルチメディア・データベース」「1996年4月に設立されたデジタルアーカイブ推進協議会が、「デジタルアーカイブ構想」として進めているプロジェクト」と2つの解説がある。マルチメディアとともにデジタル用語として普及していたことが伺える。

武邑(2003)は、デジタル技術を駆使したマルチメディア・コンテンツとして

CD-ROM『レオナルド・ダ・ヴィンチ-真のルネサンス人の神秘的草稿』を取り上げ、閲覧者を未曾有の経験に参加させると評した。また、デジタル資源をデジタルアーカイブとして構築することで、資産価値が拡大する可能性があるとは指摘している。

(2) デジタルアーカイブの対象

総務省のガイドライン(2012)では、「図書・出版物、公文書、美術品・博物品・歴史資料等公共的な知的資産をデジタル化し、インターネット上で電子情報として共有・利用できる仕組み」と定義している。アーカイブについては、「古文書・公文書・記録文書の集合、もしくはその保管所、図書館を指す言葉」としている。また、「有形・無形の文化財をデジタル情報として記録し、劣化なく永久保存するとともに、ネットワークなどを用いて提供すること。最初からデジタル情報として生産された文化財も対象」(日本図書館情報学会用語辞典編集委員会2013)とあり、デジタルアーカイブの対象を文化財や公共的な知的資産に限定している。

しかし井上(2020)は、「有形・無形の文化・科学・産業資源等をデジタル化により保存し、検索により利用者の求めに応じ多様な情報を持ったメディアで継続的に提供し、人々の意思決定や創造的活動に活用するシステム」と、科学・産業資源等も対象としている。

全国の図書館、博物館、公文書館などの文化施設で資料のデジタル化が行われ、インターネットで提供されている。近年では、国立国会図書館のジャパンサーチをはじめとして、少しずつではあるが利活用を意識した、オープンデータによる提供も増えてきている。

(3) デジタルアーカイブの種類

デジタルアーカイブには様々な種類があるが標準的な分類基準がない。デジタルア

ーカイブを構成するデジタル・コレクションの内容による分類と、規模と運営主体・方法による分類が考えられる（柳・中川 2017）。

杉本（2008）は、デジタルアーカイブを「収集されたデジタル資源を管理するための総合的なサービス」と定義し、「文化遺産アーカイブ、学術論文のリポジトリ、政府や企業の記録アーカイブ」などをあげている。

（4）持続可能性の定義

持続可能性とは、「環境の側面から持続可能であると同時に、経済・社会の側面についても健全で持続可能であること」（環境省 2019）である。

持続可能な開発の最も代表的な定義は、環境と開発に関する世界委員会（ブルントラント委員会）が 1987 年に公表した報告書『Our Common Future（邦題：我ら共有の未来）』（World Commission on Environment and Development 1987）で提唱された、「将来の世代のニーズを満たす能力を損なうことなく、今日の世代のニーズを満たすような開発」である。

ここでは、持続可能な発展の概念をもとにし、より広い領域や分野を対象としている。

したがって、デジタルアーカイブの持続可能性には、システムと文化の 2 つの意味を持つことが多い。

（5）デジタルアーカイブの持続可能性

本研究におけるデジタルアーカイブの持続可能性とは、保存の側面から持続可能であると同時に、経済・社会の側面についても持続可能であることである。

海外では、デジタルライブラリの持続可能性についての研究があるが、Fyffe and Warner（2005）などの技術的持続可能性の研究が多い。

2. で指摘したように、デジタルアーカ

イブには持続可能性の問題がある。地域映像アーカイブの持続性について、Web 調査とアンケート調査を行った川上他（2011）は、多くのデジタルアーカイブが消滅していることを指摘した。さらに、①更新しない団体はアクセス解析をしていない、②設置母体が自治体の場合が消滅しやすいとしており、デジタルアーカイブの管理運用や組織の面で問題があることが伺える。

権（2005）は、文化財のデジタルアーカイブが持続性を持つために必要な 3 つの要件をあげている。

- ① コンテンツの標準化
- ② コンテンツの市場性を持つこと
- ③ 総合的な管理・運営

まず、②については、「コンテンツが、それを求める人々にとって価値があり、かつ容易に入手（あるいはアクセス）できる状態にあること」を指すが、教育分野に限定している。③については、杉本（2008）は、規模の小さい組織が保管する情報資源を保存するためには、①コスト削減、②効率的な構築のための協力体制が必要であると示した。

永崎他（2015）は、「オープンデータ化、すなわち再配布可能なライセンスによるコンテンツの公開」と「再配布を可能にすることによるデータ共有の継続という方法」が必要であり、「何を以て持続しているとみなし得るのか」というコンセンサスの形成を提案している。また、事業のライフサイクルに焦点を当てている（永崎 2019）。

後藤（2008）は、人文科学におけるあらたなデジタル利用の像という課題のため、保存と活用を軸に分析を行い、知的完全性を保証しうる標準を作成する必要性を指摘した。

知的完全性を保証しうる標準を作成する必要性、「デジタルシステム」更新の際の透明性と遡及可能性、「デジタルシステム」の

表現の工夫とメッセージの自覚化を両立させるヒントは、自然科学分野のデジタルアーカイブにあるのではないだろうか。

このように、デジタルアーカイブの持続可能性の問題について、様々な視点から言及されているが、十分な検討はされていないと言える。

3. オープンデータ

ここでは、デジタルアーカイブと密接に関係する、オープンデータについて概観する。

(1) オープンデータの定義

自然科学分野のデジタルアーカイブについて考察するためにオープンデータを概観する。

2013年6月、英国ロッキンガム・アーンで開催された主要8か国首脳会議において「G8オープンデータ憲章」(外務省, 2013)が合意された。日本でも2012年7月に「電子行政オープンデータ戦略」がIT総合戦略本部(2012)により公表され、オープンデータの推進が重要な政策課題とされたことから、国や地方自治体での取り組みが進んでいる。

2016年12月に官民データ活用推進基本法(2016)が公布・施行され、第11条において、国、地方公共団体が保有するデータについて、国民がインターネット等を通じて容易に利用できるような措置を講じることが義務付けられた。これを受け、2017年5月にオープンデータ基本指針(IT総合戦略本部, 2017)が公表された。国、地方公共団体、事業者が公共データの公開及び活用に取り組む上での基本方針をまとめたもので、オープンデータの意義、定義、基本的ルール等を示している。ここでは、オープンデータを「①営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの、②機械判読に適したもの、③無償で利用できるもの」と定義している。

るもの」と定義している。

庄司(2015)は、オープンデータを「自由に使えるデータ」とし、オープンデータ活用を政府活動・社会生活・経済活動にさまざまな効果を生み出すこととしている。イギリスのOpen Knowledge International(OKI)が公開するオープンデータガイド「The Open Data Handbook」(OKI, 2015)では、Open Definition(OKI, 2014)の定義に従うデータをオープンデータとしている。そこでは、オープンデータを「自由に使える再活用もでき、かつ誰でも再配布できる」データとしている。Open Definition 2.1では、データがオープンデータとなるための要件として11の要件を設定しており、大きく「オープンライセンス」「アクセス」「オープンフォーマット」の3つの原則にまとめることができる。

本稿では、オープンデータについて、Open Knowledge Internationalの定義に従うこととする。

オープンデータには、科学技術分野のオープンデータも含まれる。高谷・吉村(2015)は、科学技術分野のオープンデータについて、「オープンサイエンス=オープンアクセス+オープンデータ」と整理している。オープンサイエンスに関する研修プログラムを提供するFOSTERプロジェクトでは、オープンサイエンスに関連する用語と概念を分類し、Open Science Taxonomyを提供している。

2007年のOECD(経済開発協力機構)による「公的資金による研究データへのアクセスに関するOECD原則とガイドライン」、2013年のG8科学大臣会合における研究データのオープン化を確約する共同声明への調印を契機に、データ公開とその義務化が推進されている。国内では、2015年に内閣府「国際動向を踏まえたオープンサイエンスに関する検討会」報告書が公開された。

また、2017年11月には国立情報学研究所が gacco で無料オンライン講座「オープンサイエンス時代の研究データ管理」を開講するなど様々な試みが始まっている。

以上のように、自然科学分野ではデータの利活用が進んでいる。本研究は、データの利用促進を図ることを目的として、自然科学分野の研究データについて国内事例を用いて分析し、文化財を中心としたデジタルアーカイブの利活用促進の可能性について言及する。

Christine L. Borgman(2015)は、「データの概念は実際の業務に深く根付いており、暗黙知として定義される」とし、自然科学、社会科学、人文科学の3分野における事例分析からデータの多様性を示している。データの価値はその利用にあるとしている。

(2) オープンデータの時代背景

自然科学系のオープンデータの歴史は古い。1957～1958年の国際地球観測年(International Geophysical Year: IGY)を契機に設立されたWorld Data Centerは、地球科学分野を中心とした観測データを収集、保存、配布する国際的なデータセンターの組織である。日本も当初から参加している。

4. 事例研究-自然科学分野

本研究における事例研究の目的は、自然科学分野の成功事例を分析することで、デジタルアーカイブの持続可能性を高めるために必要な条件を一般化することである。

調査方法は、事例資料を分析する物語分析による。物語分析とは、単独・理論事例を対象にしてダイナミクスを解明する分析手法である(田村 2016)。

単独事例としての熱帯降雨観測衛星(Tropical Rainfall Measuring Mission; TRMM, 以下 TRMM 衛星)において、「データ

の利活用とデジタルアーカイブの持続可能性との間にどのような関係性があるのか」をリサーチクエスチョンとして設定する。

TRMM 衛星の事例研究から理論仮設を生成する際に、物語分析を研究手法として用いた。そして、TRMM ミッションと利用推進事業を物語として再構成することにした。その際、TRMM 衛星の運用終了と後継ミッション(全球降水観測計画: GPM)への引継ぎを TRMM 衛星の成功と捉えて、物語の終点に位置付けた。また、TRMM 衛星が長期間にわたって観測を継続したことから本事例を複雑物語としての斬新物語として位置付け、過程追跡を行った。

過程追跡とは、「特定の単独事例の従属変数の結果を生み出す因果過程の諸段階を、歴史的なコンテキストにおいて識別する手順」である(田村 2016)。事例の歴史年表としての出来事年代記を作成し、出来事構造を明らかにした。事例データは、TRMM ホームページ(1)のアーカイブや宇宙航空研究開発機構(以下、JAXA という)のプレスリリース、図書・雑誌等から収集した。

5. TRMM 衛星の概要と研究の展開

(1) TRMM 衛星の概要

TRMM 衛星は、NASA が開発した衛星に、JAXA(当時 宇宙開発事業団: NASDA)が情報通信研究機構(当時 通信総合研究所)の協力を得て開発した降雨レーダ(PR)を搭載した日米共同プロジェクトである。十年以上の歳月をかけたプロジェクトは、1997年11月にH-IIロケットにより種子島宇宙センターから打上げられた。設計寿命の3年を大幅に超えて稼働した TRMM 衛星は、2015年4月に運17年以上にわたる運用を終了し、6月16日に大気圏に再突入した(岡本ほか 2016)。

気象・気候研究のために、熱帯・亜熱帯

の降雨を観測することを目的とした TRMM 衛星の主要ミッションは、次の 3 つに集約される (JAXA EOC 2006)。

- ①熱帯域の降雨量を定量的に把握することにより、地球のエネルギー・水循環を解明する。
- ②熱帯域の降雨量の様々な時間-空間スケール変動の実態と大気大循環に与える影響のメカニズムを明らかにするとともに、これらを再現・予報するための数値モデルの検証・開発を行う。
- ③宇宙からの降雨量観測の手法を確立する。

(2) TRMM 衛星の観測装置

TRMM 衛星は 5 つの観測装置を持つ。降雨レーダ (PR), TRMM マイクロ波観測装置 (TMI), 可視赤外観測装置 (VIRS), 雲及び地球放射エネルギー観測装置 (CERES), 雷観測装置 (LIS) である。

PR, TMI, VIRS は、熱帯や亜熱帯地域における降雨観測 (降雨の鉛直構造や輝度温度等) のためのセンサである。TMI は傾斜軌道に搭載される初の受動型マイクロ波放射計であり、PR は世界で初めての衛星搭載用降雨レーダで、降水に関するデータセットを提供する。

降雨観測機器の他に、CERES と LIS が TRMM 衛星に搭載されている。これらの機器は NASA の地球観測システム (EOS) プログラムのための機器であり、CERES は地球の放射エネルギーを測定し、LIS はグローバルな雷の分布を測定するために利用される。

(3) 地球観測衛星データ

衛星データとは、人工衛星に搭載された観測装置によって観測されたデータのことである。離れた場所から非接触で、対象物を計測する技術をリモートセンシングといい、物質の電磁波 (光, 電波) に対する特性を利用している。観測装置の種類は、大きく分けると光学センサとマイクロ波センサを用いたレーダに分けられ、地表からの

光の反射や放射、電波の散乱を画像情報として取得する。

地球観測衛星データには、その処理段階に応じたレベルがある。衛星から受信したデータは伝送用に加工されているため、元の Raw データに戻すことをレベル 0 処理という。このレベル 0 データは一般には提供されない。レベル 1 処理は基本的な処理で、レベル 0 データからシーン単位で切り出して補正したデータである。レベル 2 および 3 は高次処理と言われ、レベル 1 データを用途にあわせてアルゴリズム処理を行う。多くの研究者と報道機関等はレベル 2 または 3 のデータを利用する。TRMM 衛星プロダクトのレベル定義はレベル 0 から 3 までであるが、レベル 4 まで設置する場合もある。

分析方法の向上などによってアルゴリズムの改訂が行われ、リリースされる。このアルゴリズムのバージョン管理は、観測データの正確性・信頼性を獲得する上で重要である。

(4) TRMM 衛星のデータ

TRMM 衛星の観測データは、CCSDS 準拠の packets 方式によりフォーマットされた後、米国の追跡・データ中継衛星 (TDRS: Tracking and Data Relay Satellite) を経由して NASA の地上局で受信する。宇宙データシステム諮問委員会 (CCSDS: Consultative Committee for Space Data System) は、1982 年に各国の宇宙機関により設立された宇宙データ通信システムに関わる国際標準化検討委員会で、宇宙データ通信システムの定義・規格化している。全データの前処理、PR, TMI, VIRS の高次処理等は、NASA ゴダード宇宙飛行センター (略: GSFC: Goddard Space Flight Center, 以下 GSFC という) で行われ、観測後 48 時間以内にレベル 0 データに処理される。

CERES と LIS については、前処理済みのデータ (レベル 0) が、それぞれ、NASA の

ラングレー研究センター (LaRC : Langley Research Center), マーシャル宇宙飛行センター (MSFC : Marshall Space Flight Center) に送られて処理している。

日本へは, GSFC より PR のレベル 0 データが送られ, JAXA EOC (地球観測センター) で高次処理される。EOC や GSFC において高次処理されたデータ (レベル 1~3) は日米のほか, 気象, 気候, 水文等の分野の科学者を含むユーザに配布している。

JAXA EORC (当時 地球観測データ解析研究センター) においては, TRMM 衛星データを用いた解析研究を行うとともに, 研究用データセットを作成し研究者に提供している。

1999 年秋より, NASA が準リアルタイムデータの配信を開始した。JAXA EOC では, 2000 年 11 月より, PR 2A25-R1, PR 2A25-R2, TMI 1B11 の準リアルタイムデータを TSDIS よりオンラインで取得し, 気象庁に伝送している。

(5) 提供プロダクト

日本が提供するプロダクトの種類は 4 つあり, 2006 年時点での提供プロダクトは以下のとおりである (NASDA EORC 2003)。

①標準プロダクト

JAXA や NASA において作成される基本プロダクト

②切り出しプロダクト

標準プロダクトのデータから経度 10 度おきに, または特定の領域 (日本域・東南アジア域) を切り出したプロダクト

③研究プロダクト

研究を目的として作成されるプロダクト

④準リアルタイムプロダクト

天気予報などに用いることを目的としたプロダクト。配信されるまでの時間は早い, データは多少簡略化されて

おり, 位置情報にも数キロのズレがある。

(6) TRMM 衛星の特長と成果

岡本ほか(2016)は, TRMM 衛星の特長として, ①気候変動に大きな影響を及ぼす熱帯・亜熱帯降雨の観測や降雨に伴う潜熱加熱の算出に特化したミッションであること, ②日米共同ミッション, ③最長の部類に属する 17 年間にわたる観測, ④日本のロケットによる初の外国製衛星の打ち上げ, ⑤日本が開発した世界初の衛星搭載降雨レーダ, 日米等のサイエンスチームによる降雨レーダアルゴリズムの開発などをあげている。

寺門(2015)は, TRMM 衛星の成果について, 衛星による降水の観測推定精度が著しく向上したこと, 「降水システム気候学」ともいふべき新しい研究分野が確立したこと, 雨についての標準データとなったことをあげている。

TRMM 衛星は, 当初の目的である科学的な研究だけでなく, 実利用に対しても成果を残している。降水分布を準リアルタイムで提供することで, 災害監視, 漁業, 農業といった社会と密接に関わる分野への利用が進んだ。

(7) 提供プロセス

第 1 期は 1997 年 11 月の観測開始からとした。データとツールの提供および PR ブラウズ画像を公開した。第 2 期は, 1999 年 3 月の PR 準リアルタイム画像の配信を開始, 12 月には TMI と VIRS の海面水温等を公開した第 3 期は 2000 年 4 月の月平均降雨データページの公開から始まり, 台風データベース・台風速報などの複数センサのデータと画像を提供した。第 4 期は, 台風の 3 次元動画の提供開始にあわせて, 他の衛星の台風データベースと統合し, 利便性が拡大した。第 5 期は, 複数衛星データによって準リアルタイム配信される世界の雨分布速報 (GSMap) が開始。第 6 期は, 2014 年 2

月に GPM が打ち上げられ、TRMM から GPM へのデータの連続性を確保した。

6. おわりに

本研究では、デジタルアーカイブの持続可能性条件は利活用にあるのではないかと考え、自然科学分野の成功事例を調査した。

特筆すべきは 17 年間にわたる長期運用である。長期運用に至った理由としては、2001 年の当初ミッション達成後に、運用延長を目的とした高度変更を行っていたことが大きい。次に、運用終了が決定した時には、ハリケーンの進路予測にとって必要なデータとなっていたことである。

また、GPM 計画が早くから立ち上がっており、TRMM 衛星と主衛星との運用期間を重ねることができたことも大きい。同一地域を同時に観測したデータの比較が可能になり、TRMM 運用終了後もデータの連続性を確保できたことで、データを継続的に取得することが可能になった。

TRMM 衛星は、サイエンスチームの研究成果を提示し、利用を促した。台風の経路情報などの付加価値を付与して提供する、同様のデータベースを統合するなどコンテンツを発展させた。

初画像から運用終了までの 17 年間にわたって観測画像に解説を付与して提供し続けたこと、ホームページの情報がアーカイブされ、1999 年からのデータが残っていることも TRMM 衛星チームの努力が伺える。

参考文献

- (1) Borgman C. L. (2016) Big data, little data, no data : scholarship in the networked world. The MIT Press.
- (2) Chowdhury, G. (2014) 'Sustainability of digital libraries: a conceptual model and a research framework', International Journal

(3) 井上透 (2016) 'デジタルアーカイブの総合的な開発デジタルアーカイブの総合的な開発～開発計画、開発実践、評価の観～開発計画、開発実践、評価の観点から～点から～', 岐阜女子大学デジタルアーカイブ研究所年報岐阜女子大学デジタルアーカイブ研究所年報 2015. 2015.

(5) 宇宙開発事業団宇宙開発事業団 and and 通信総合研究所通信総合研究所 (2002) 宇宙から見た宇宙から見た雨雨 : : 熱帯降雨観測衛星熱帯降雨観測衛星 44 年間年間の軌跡軌跡. Japan Advance Plan Co., Inc. Japan Advance Plan Co., Inc.

(6) 宇宙開発戦略本部 宇宙開発戦略専門調査会 宇宙開発戦略本部 宇宙開発戦略本部 宇宙開発戦略専門調査会 議事次第 議事次第, , 首相官邸首相官邸. Available at: Available at:

<https://www.kantei.go.jp/jp/singi/utyuu/seimon/dai13/gijisidai.html>

<https://www.kantei.go.jp/jp/singi/utyuu/seimon/dai13/gijisidai.html> (Accessed: 18 January 2021). (Accessed: 18 January 2021).

(7) 宇宙航空研究開発機構 (2018) GPM(2018) GPM データ利用ハンドブックデータ利用ハンドブック. . 第 3.13.1 版 版. Available at: Available