

# 地誌資料のデジタル化および生成 AI 時代における 検索機能をめぐる諸問題に関する横断的研究 ～「はちのへ新聞」を題材として

埴原啓正＝根城隆幸

## 要旨

八戸市立図書館所蔵の明治期の郷土新聞である『はちのへ新聞』のマイクロフィルムをデジタル化し、持続可能なデジタルアーカイブ構築の在り方を検討するものである。マイクロフィルムは耐久性に優れる一方で、老朽化や閲覧性の低さが課題となっていた。そこで本稿では、スキニングや画像検査、トリミング等の工程を経て当該新聞の PDF 化を実施した。その過程で、デジタル化は閲覧の地域間格差を是正し、保存と活用の両立を可能にする一方で、旧字体への OCR 対応やマイクロフィルムの破棄と維持管理に関するコスト、広報活動の強化といった課題も明らかになった。

## はじめに

本研究は、全国的に過疎化や高齢化による地域社会の衰退が進み、それと同時に各地方で発行されていた歴史的に貴重な資料の保存が課題となっている現在の状況に鑑み、地方紙を題材として、当該歴史的資料を後世に伝えていくための保存方法（資料のデジタル化・アーカイブ化）および付随する諸問題について横断的に検討するものである。

具体的には、青森県八戸市立図書館において長期保存媒体として利用されてきたマイクロフィルムのデジタル化を取り上げ、その技術的手法、保存戦略および法的課題等を整理し、持続可能なデジタルアーカイブ構築の在り方を検討する。

マイクロフィルムは 20 世紀における貴重な記録保存媒体であり、高い耐久性を有する一方、閲覧性および検索性に制約がある。本研究では、スキニング、画像補正、OCR 処理といったデジタル化工程を分析し、保存フォーマットおよび運用モデルを検討する。

今年度は、明治 33 年創刊の『はちのへ新聞』マイクロフィルムをデジタル化した。本研究は、地域資料へのアクセス環境を改善し、閲覧の地域間格差の是正に寄与するものである。本稿ではその経緯を報告する。

## 1 事業開始

八戸市立図書館は、明治 7 (1874) 年設立の「八戸書籍縦覧所」を起源とする。これは、東京、京都に次ぐ全国で 3 番目に古い縦覧所とされる。その後、明治 13 年設立の「73ヶ町村立八戸町立書籍館」および明治 28 年設立の「八戸青年会付属図書館」が統合され、大正 2 年に八戸町立図書館となっている。昭和 4 年の八戸市政施行に伴い現名称となり、現在に至っている。



明治時代中期から昭和時代前期頃の八戸経済圏の政界は、大きく分けて憲政会系の土曜会と政友会系の公民会の二派閥で構成されていた。両派閥は独自に機関紙を発行しており、土曜会が発行していたものが『はちのへ新聞』である。

同図書館は、明治時代後半に発刊された『はちのへ新聞』および『奥南新報』を所蔵しているが、原資料の損傷が著しいため、昭和 61 年に作製されたマイクロフィルムにより保存を図ってきた。しかし、製作から約 40 年が経過し、マイクロフィルムおよび閲覧機器の老朽化が顕在化している。事業期間中にもマイクロリーダーの破損が発生し、作業を中断せざるを得ない状況が生じた。

以上の状況を踏まえ、郷土資料の保存と活用を両立させるため、文字検索機能の付与およびネットワーク環境下での利用を視野に入れたデジタル化を実施した。

## 2 マイクロフィルムの特徴

マイクロフィルムは、新聞、公文書、学術資料の保存媒体として広く利用されてきた。第二次世界大戦後、情報量の増大に対応する手段として各国で導入が進み、日本においても国立国会図書館や各自治体の公文書館等で採用されてきた。

紙媒体は劣化や保管スペースの問題を抱えるが、マイクロフィルムは縮小撮影による省スペース化が可能であり、適切な環境下では約 100 年の保存が可能とされる。しかし、利用時の物理的摩耗、専用機器への装着やコマ送り操作の必要性など、閲覧効率の低さが課題である。特に一連の紙面を通読する場合、多大な時間を要する点は利用上の大きな制約となっている。

デジタル技術の進展は、これらの課題を克服し、保存資料を再活性化させる可能性を有している。本研究は、マイクロフィルムのデジタル化の意義と実践的課題を多角的に検討するものである。

## 3 デジタルデータの作成

### (1) 著作権の確認

デジタル化に先立ち、著作権の存続状況を確認した。当該新聞の発行所は現存せず、著作権保護期間は満了していると判断された。法的問題が生じないことを確認した上で事業に着手した。

### (2) スキャニング単位

新聞資料の利用実態を踏まえ、1 ページを基本的な単位とした。見開きではなくページ単位とすることで、検索および活用の利便性を高めている。保存形式は汎用性と公開利用を考慮し、PDF 形式を採用した。作業に際しては傾き補正を実施し、不要な余白部分をトリミングした。

### (3) 色彩情報

原資料がモノクロ印刷であるため、モノクロ画像としてデータ化した。彩度の際限よりも諧調の明瞭性を重視し、判読性の確保に努めた。



#### 4 デジタル化作業の実際

マイクロフィルムのデジタル化作業は、マイクロフィルムが館内の閲覧室にあるキャビネットに保管されているため、作業スペースにて実施した。職員の協力のもと、指定した発行年のマイクロフィルムをスキャニングの作業場所へ運搬し、マイクロリーダーを用いたスキャニング準備を行う体制を整えた。

##### (1) スキャニング

各コマを逐次確認しながらスキャニングを実施した。必要に応じて傾き補正や不要部分の削除を行ったが、色調補正については実施を見送った。

##### (2) 画像の検査

全画像について目視検査を行い、傾き、欠損、汚損等を確認した。不備がある場合は再度スキャンを実施した。ただし、裏写りによる判読困難な箇所については、技術的制約から補正が困難な事例も見られた。

##### (3) トリミング

閲覧時の視認性を考慮して、資料外周の余白削除をできるだけ実施するようにした。

#### 5 デジタルデータの利活用

八戸市立図書館は、令和6(2024)年に150周年を迎えた。長年継承されてきた郷土資料を適切に保存・活用し、魅力ある図書館となるよう住民サービスの向上に努めている。国内で郷土資料のデジタル化は進展しているが、未着手の資料も依然として多い。

デジタル公開は物理的距離による制約を解消し、地域間格差の是正に寄与する。また、テキスト化(OCR処理)が進めば全文検索が可能になり、利活用の飛躍的な促進が期待される。提供にあたっては、目次データを整備し、利用者の検索利便性向上を図った。

6で後述するように、現行のOCR技術によるテキスト化では依然として旧字体の判読等に課題が残るため、生成AIを用いた検索機能に関する検討を引き続き行なっていく予定である。



#### 6 今後の課題 ～国立国会図書館・日本大学法学部図書館におけるインタビューを踏まえて～

本研究では、機会を得て、2025年9月27日に国立国会図書館および日本大学法学部図書館を訪問し、インタビュー調査を実施した。得られた知見に基づく今後の課題は次の通りである(公開を前提としないインタビューのため、詳細についての説明は省略する)。

まず、単なる郷土資料のデジタル化のみならず、テキストデータの生成を行うことが今後の課題である。当時の新聞には現在使用できる常用漢字にはない旧字・正字体が多く使用されており、現行のOCR技術のみでは完全なデータ化はなかなか至難である。現在開発が進んでいる古

文書解読AIや旧字対応OCR技術の導入により、検索機能が充実すれば、利用価値は一層高まるであろう。

また、デジタル化を終えたマイクロフィルムの取り扱いについても課題が残る。具体的には、マイクロフィルムの量によっては多額の処分費用を要し、デジタル化を終えたとしても保管コストが生じる可能性がある。それに加え、経時劣化したマイクロフィルムから発する酢酸臭（ビネガーシンドローム）への対策など、維持管理コストの検討が必要となる。

さらに、広報活動の強化も重要である。市公式ウェブサイト等を通じた公開により、幅広い世代への周知を図る必要がある。郷土資料への接触は地域住民の生涯学習を支え、幅広い世代の方々が郷土に興味関心を抱き、地域文化への理解と誇りの醸成に寄与するものである。

自治体の財政的制約はあるものの、文化資源の保存と継承は公共的使命である。今後も施設設備の維持と資料保存が継続的に推進されることが望まれる。

## むすびに

八戸学院大学地域経営学部は、地域社会との連携に重きを置いている。本研究は、大学と地域が相互に協力し、地域が保有している歴史的に貴重な学術・教育資源を有効に活用することを目指したものである。

今回、郷土の貴重資料をデジタルデータとして保存する基盤を構築した。作成した画像データは、限られた期間で作成したものであり、パソコンでの使用を考慮してディスプレイ上での判読を主眼とした制度であるが、技術革新に合わせたさらなる改善の余地を残している。

貴重資料のデジタル化は、資料保存の意義を持つものであり、遠隔地の利用者へも公開できることになる。貴重資料そのものは、物理的な経年劣化や破損汚損から内容を保護し、次世代へ継承する責務を果たすものである。また、遠隔地からのアクセスを可能にするなど、資料の公開性と保存性を両立させる大きな意義を持つ。

八戸市立図書館は、予算的・人為的リソース不足から地誌資料の長期保存に関する課題を解決できずにいた。本研究はそのような課題に対し、地誌資料をデジタル化・アーカイブ化することにより解決しようとするものである。本研究を通して、地域の歴史や文化を継承することが可能となり、地域課題の解決だけでなく、地域住民の生涯学習や本学学生の地域理解の促進、引いては歴史・文化理解を通じて八戸地域の発展に寄与する点で大きな意義がある。

今後、ネットワーク社会の進展に伴い、デジタル化された情報が新たな技術によってより有意義に活用されることを期待する。

図書館で所蔵している古文書等の「デジタルデータ」を作成、及び本研究にかかるインタビュー調査にご協力いただいた八戸市立図書館、国立国会図書館、日本大学法学部図書館の皆様には心から感謝申し上げます。

【付記】本稿は、令和7年度八戸学院大学特別研究費「地誌資料のデジタル化及び生成AI時代における検索機能をめぐる諸問題に関する横断的研究—『はちのへ』新聞を題材として—」の研究成果の一部である。

参考文献・資料

- (1) 国立国会図書館『資料デジタル化の手引き』2017年版  
([https://dl.ndl.go.jp/view/download/digidepo\\_10341525\\_po\\_digitalguide170428.pdf?contentNo=1&alternativeNo=](https://dl.ndl.go.jp/view/download/digidepo_10341525_po_digitalguide170428.pdf?contentNo=1&alternativeNo=))(2026年3月2日最終閲覧)。
- (2) 広報はちのへ令和6年11月号2頁  
(<https://www.city.hachinohe.aomori.jp/material/files/group/1/koho20241102.pdf>) (2026年3月2日最終閲覧)。
- (3) 滝尻侑貴「政党新聞」  
(<https://historia8.org/cultural-property/af3jo1ytw1/>) (2026年3月2日最終閲覧)。