

デジタルアーカイブにおける聞き書き活用支援システムの構築

寺嶋 一将† 植竹 俊文† 竹野 健夫†

岩手県立大学 ソフトウェア情報学部†

1. はじめに

近年、地域の歴史や文化を後世へと継承させようという意識が全国で高まり、デジタルアーカイブの活用が拡大している¹⁾。先行研究²⁾でも、地域住民に対して取材調査を行い、そこで得られた歴史や文化に関する口述記録を聞き書きとして編集・蓄積し、公開する研究が行われてきた。

本研究では蓄積された聞き書きの活用支援を目的として、聞き書きの整理手法及び新たな提示手法を提案する。具体的には聞き書きに記されている時代や地域といった情報の整理を支援し、整理された聞き書きを地図や年表を活用しつつ提示していくシステムを構築し検証を行った。

2. 聞き書きの概要

聞き書きとは古老等から話を聞き、その内容について語り口調を活かしつつ文章に書き起こす手法、または書き起された文章を指す。本研究では岩手県花巻市の住民から得た聞き書きを扱う。

2.1 聞き書きの収集と公開

本研究では花巻市の郷土史研究団体である「ふるさと遺産研究所」と提携している。この団体は「花巻物語辞典」³⁾の構築を目指して郷土文化の収集等の活動を行っている。聞き書きの多くはこの団体によって収集されたものである。聞き書きは文章を話題ごとに小見出しとしてまとめ、話者の略歴と合わせて図1の様に公開されている。

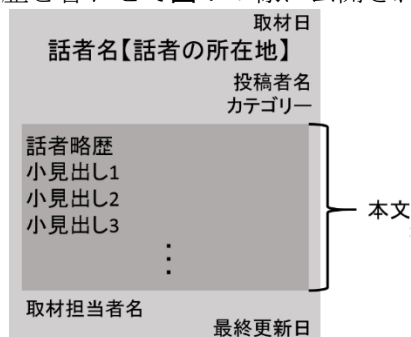


図1 聞き書きの構成

Development of Utilization Support System for Dictation in Digital Archive

†Kazumasa Terashima, Toshifumi Uetake, Takeo Takeno
Faculty of Software and Information Science,
Iwate Prefectural University

2.2 聞き書きの現状と課題

2017年1月6日時点で公開されている聞き書きは89件で、今後も増加が見込まれる。聞き書きの増加により閲覧者がアーカイブ全体を把握することが困難となり、十分に活用されない聞き書きが出てくる恐れがある。そこで聞き書きの提示手法を改める必要があるが、以下の2点が問題となる。

(1) 内容の整理手法が定まっていない

聞き書きには多様な情報が含まれており、内容を分析・整理することで聞き書きの活用の幅を広げることができる。しかし、具体的な整理手法が定まっておらず、現在は文章として提示されるのみである。

(2) 内容が閲覧者に伝わりにくい

聞き書きで示されている情報を読み取るには、文章を熟読し、その内容を分析する必要がある。しかし、多量の聞き書きを一件ずつ読み込んでいくことは困難であり、内容が十分に伝わらない恐れがある。

3. システムの概要

3.1 システムの構成

本研究は花巻物語辞典で公開している聞き書きを扱っており、以下の概念図(図2)の枠線内が本システムの範囲となる。

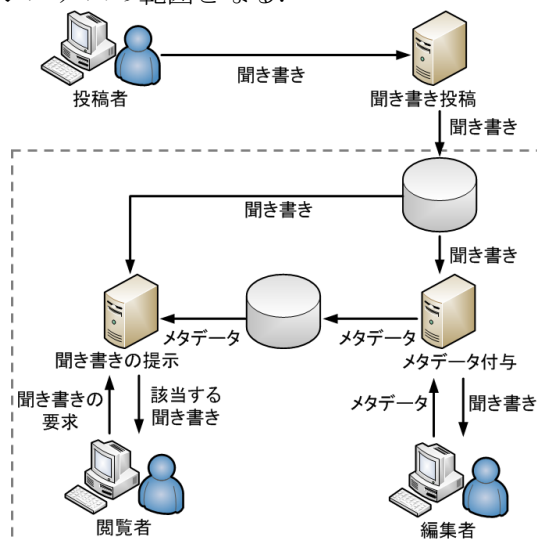


図2 システム概念図

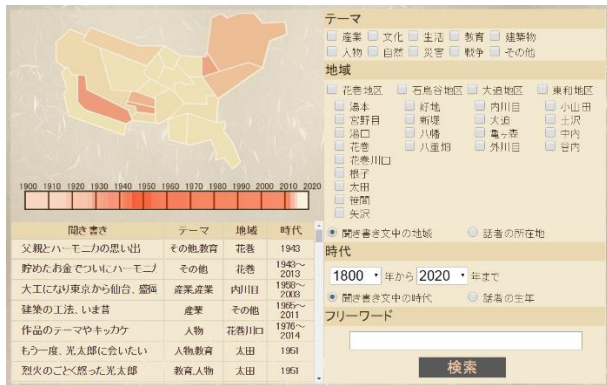


図3 聞き書きの提示画面

3.2 システムの機能

システムの主な機能は以下の2つに分けることができる。

(1) 聞き書きが示す内容の整理補助

まずは話者の名前や生年などといった情報を整理する。その後、編集者が聞き書きを小見出しごとに熟読・分析することで得られた情報(メタデータ)をデータベース上に格納していく。

(2) メタデータをもとに聞き書きを提示

メタデータをもとに聞き書きを提示する仕組みを構築した。閲覧者の目的に沿った聞き書きの絞り込みを支援し、地図や年表と共に閲覧者に対して聞き書きを提示する。聞き書きは小見出しごとに分割された状態で一覧として表示(図3)される。該当する小見出しを選択すると詳細な情報と花巻物語辞典へのリンクが示される。

4. 聞き書きの整理手法

話者と取材に関する情報(図4の左側の枠線で囲まれた範囲)を整理した後に、小見出しごとに、どのような内容が記されているかを、テーマ、時代、地域の3つの観点から整理する。また、これらのメタデータを付与する際に参照した聞き書き文中の語句に関しても記録し、後にメタデータを付与する際の基準とする。

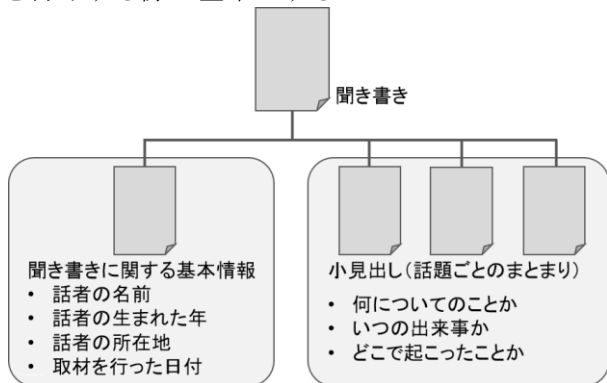


図4 聞き書きでまとめる情報

5. システムの評価

評価に関しては編集者側と閲覧者側の2つの側面から行った。

(1) 編集者側の評価

ふるさと遺産研究所の職員1名が、提案システムを利用して5件の小見出しが存在する聞き書きを1件整理した。その結果、小見出しの内容を整理するうえで、テーマを選定する作業で迷うことがあったという意見が得られた。

また、テーマとして用意した10項目のメタデータ(産業、文化、生活、教育、建築物、人物、自然、災害、戦争、その他)に関して、メタデータの数は適正だが、「その他」が付与された小見出しから新たなテーマを設定できるのではないかという指摘を受けた。そこで、評価者自ら聞き書き14件(小見出しに分けると90件)を整理したが「その他」が付与された小見出しに記された内容の関連性から新たなテーマを導き出すことはできなかった。

(2) 閲覧者側の評価

岩手県在住の学生5名が、花巻市出身の童話作家である「宮沢賢治」についての情報をもつ小見出しを抽出する作業を行った。既存システムと提案システムを用いて抽出にかかる時間を比較した。その結果、5名全員が提案システムの利用で抽出時間を短縮できていた。抽出の様子を観察したところ、無関係な小見出しを読み込むという作業にかかる時間が減少していた。これにより、該当する小見出しの抽出を効率的に行うことができていた。

6. おわりに

本研究では、地域アーカイブに蓄積された聞き書きの活用促進として、整理・分類による聞き書きの活用促進を図った。今後は複数の聞き書きの関連付け及び聞き書き以外の史料等のシステムによる統合的な管理を目指す。そこから、花巻市に関する歴史資料の統合的な管理システムの構築に取り組む。

参考文献

- 総務省関東総合通信局情報通信連携推進課: 地域住民参加型デジタルアーカイブの推進に関する調査検討会報告書(2010).
- 小田島瑞希, 竹野健夫, 植竹俊文, 菅原光政: 地域コミュニティにおけるデジタルアーカイブシステム, 情報文化学会講演予稿集 20, pp. 38-41 (2012).
- ふるさと遺産研究所: 花巻物語辞典, <https://hana-isan.com/Home> (参照 2017-01-06).