

ハンセン病回復者の語り部の証言記録に対する 質問応答システム構築に向けたベクトル検索精度の検証

孝壽真治¹ 竹内孔一²¹ 岡山大学大学院環境生命自然科学研究科
s-kouju@s.okayama-u.ac.jp² 岡山大学大学院環境生命自然科学研究科
takeuc-k@okayama-u.ac.jp

概要

ハンセン病回復者の語り部との対話を通じた人権学習は、自身の考えや価値観を見つめ直す対話型体験学習として重要である。しかし、語り部の高齢化やハンセン病を取り巻く環境などの背景から、次世代の語り部への記憶の継承が難しい。本研究では、語り部の実体験や記憶を後世に継承するために、ハンセン病回復者の語り部との対話による対話型体験学習の効果を再現できる質問応答システムの構築を目指す。本論文においては、質問応答システム構築の事前実験として、証言記録のテキストにベクトル検索を適用した場合の精度を検証する。実験の結果、GLuCoSE-base-ja-v2が最も高い性能を発揮した。

1 はじめに

戦争、災害、公害、差別偏見などにおいて様々な教訓を後世に伝えるために語り部として活動している人々が存在する。本研究で対象としているハンセン病回復者の語り部もその一つであり、ハンセン病回復者との対話を通じた人権学習は、ハンセン病の差別・偏見における当事者の実体験や感情を伴った記憶（ストーリー）に触れ、その考えや価値観を知り、自身の考えや価値観を見つめ直す対話型体験学習として重要である。しかし、語り部活動では共通の課題として語り部自身の高齢化による継承問題がある。戦争や災害など多くの語り部が子や孫世代への語り部活動の継承を行なっているが、担い手の不足や当事者ではない者が記憶を継承することの難しさなどの課題は依然として残っている。ハンセン病回復者の語り部もその記憶の継承が急務であり、次世代の語り部育成が取り組まれているが、ハンセン病を取り巻く事情により戦争や災害の語り部以上に難しい状況である。ハンセン病回復者に対する偏見は現代においても残っており、語り部として活動す

る回復者は限られる。さらに、ハンセン病では患者の家族にも差別偏見が及んだという背景から、ハンセン病回復者は家族との繋がりを断ち、現在でも家族と連絡を取らない者も少なくない。加えて、ハンセン病患者にはかつて不当な断種・墮胎が行われていたという背景も合わさり、戦争や災害などの語り部のように当事者の家族が語り部活動を継承することが難しく、ボランティアや支援団体による継承を進めているが、戦争や災害の語り部と比べても継承が難しいと言える。本研究では、語り部の実体験や記憶を後世に継承するために、ハンセン病回復者の語り部との対話による対話型体験学習の効果を再現できる質問応答システムの構築を目指す。

近年では、LLMの発展によって質問応答に Retrieval Augmented Generation (RAG) を利用して LLM を用いる事例が増えている。本研究で取り扱う語り部の証言記録に対する質問応答システムも RAG を用いたアプローチが考えられるが、このような質問応答システムは RAG 内部のベクトル検索の精度による影響が大きい。そこで本論文においては、質問応答システム構築の事前実験として、証言記録のテキストにベクトル検索を適用した場合の精度を検証する。

実験の結果、GLuCoSE-base-ja-v2が最も高い性能を発揮したが、いずれのモデルもファインチューニングなしの精度は20%以下であり、ファインチューニングなどの手法を用いた検索精度の向上が課題であることが明らかとなった。

2 関連研究

本研究のように、特定の人物を再現するような質問応答システムの事例は複数存在するが、大きく分けて語り部による証言記録をベースとして利用した事例とそれ以外に分けられる。

証言記録を利用しない事例としては、オルセー美

術館の AI ゴッホ [1] などのように歴史上の人物を言語モデルで再現する事例が存在する。

語り部における先行研究では、ホロコーストの語り部における Dimensions in Testimony [2, 3, 4, 5] があるが、360 度カメラによる証言映像のリアリティによって語り部との対話空間という場の再現を重視しており、本研究とは趣旨が異なっている。また、日本でも質問の回答として適切な証言記録の検索を行う戦争の AI 語り部 [6] が存在する。しかしこれらの事例は、我々の研究で扱うハンセン病回復者の語り部を対象としたものではなく、ハンセン病事例に対してこのような質問応答システムを構築するという試みは我々の知る限りでは存在しない。

3 データ構築

本研究で使用する語り部の証言記録はインタビュー形式で取材し、動画として収録した。証言記録を言語モデルで利用するため、動画の音声を whisper¹⁾ を用いて文字起こしし、固有名詞などの誤りを人手で修正してテキスト化する。

人手での修正では、図のような文字起こし補助システムを作成して利用した。このシステムを利用して、whisper による音声認識の誤りを修正するとともに、発話に対して質問者と語り部のいずれかのラベルを付与する。

また、whisper による文字起こしデータは一人の話者の一つの発話内容 (質問・回答) が複数の文に分割されている場合があるため、同一話者の連続する発話は一つの文にまとめるという前処理を行った。最終的なデータセットは表 1 のようになる。

本論文では、一人の語り部に対して 1 回 1 時間程度のインタビューを実施し、実験では 2 回分の証言記録を利用している。上記のような、文字起こしと前処理を適用した後の質問・回答のペアは合計で 121 件となった。

4 実験

4.1 実験設定

121 件の質問回答ペアを用いて、ベクトル検索による精度を検証する。質問テキストと回答テキストをそれぞれベクトル化し、質問テキストと回答テキストのコサイン類似度を計算し、質問に対してペア

として適切な回答が選ばれるかを評価する。本実験ではモデルのファインチューニングは実施せず、事前学習済みの埋め込みモデルをそのまま利用した。また、“query:” や “passage:” などのプレフィックスを用いることを推奨しているモデルはプレフィックスを付与した場合と付与しない場合の 2 通りの実験を実施し、より精度が高い方の結果をモデルのスコアとして採用する。

4.2 実験結果

実験結果を表 2 に示す。Top-1 の精度は質問に対して回答を正しく検索できた場合、Top-3 の精度は類似度の上位 3 件の中に正解が含まれる場合とする。

Top-1 の精度では GLuCoSE-base-ja-v2 が 19.01% で最も高い精度を示した。Top-3 の精度では Sarashina-Embedding-v1-1B が 35.54% で最も高い精度を示した。一方で、どの埋め込みモデルも 20% 以下の精度と低く、Top-3 の場合も性能は低い。

表 2: 実験結果

モデル	精度 (top-1)	精度 (top-3)
Multilingual-E5-large	0.1736	0.3223
Sarashina-Embedding-v1-1B	0.1818	0.3554
RoSEtta-base-ja	0.1488	0.2397
GLuCoSE-base-ja-v2	0.1901	0.3058
cl-nagoya/sup-simcse-ja-large	0.1322	0.2066
simcse-ja-bert-base-clcmlp	0.1157	0.1818
Japanese-MixCSE-BERT-large	0.1157	0.2066
Japanese-DiffCSE-BERT-base	0.0826	0.1983
cl-nagoya/shiorih-large-pt	0.1488	0.3306
bclavie/fio-base-japanese-v0.1	0.1240	0.2314
JaColBERTv2.5	0.1240	0.2314

5 考察

本実験では、一般的な日本語埋め込みモデルや多言語埋め込みモデルを用いて、ファインチューニングなしの場合に、ハンセン病回復者の語り部による証言記録データセットにおけるベクトル検索の精度を検証した。RAG を用いた質問応答システムでは、Top-1 もしくは Top-k でヒットした参照テキストを LLM に与えられ、使用するデータによっては埋め込みモデルのファインチューニングなしで使用される場合もあるが、我々の証言記録データセットにおいては、ファインチューニングなしのモデルによる埋め込みは十分な性能を示さなかった。

ハンセン病回復者の語り部の証言記録に対する質

1) <https://huggingface.co/openai/whisper-large-v3>
2025/1/8 アクセス

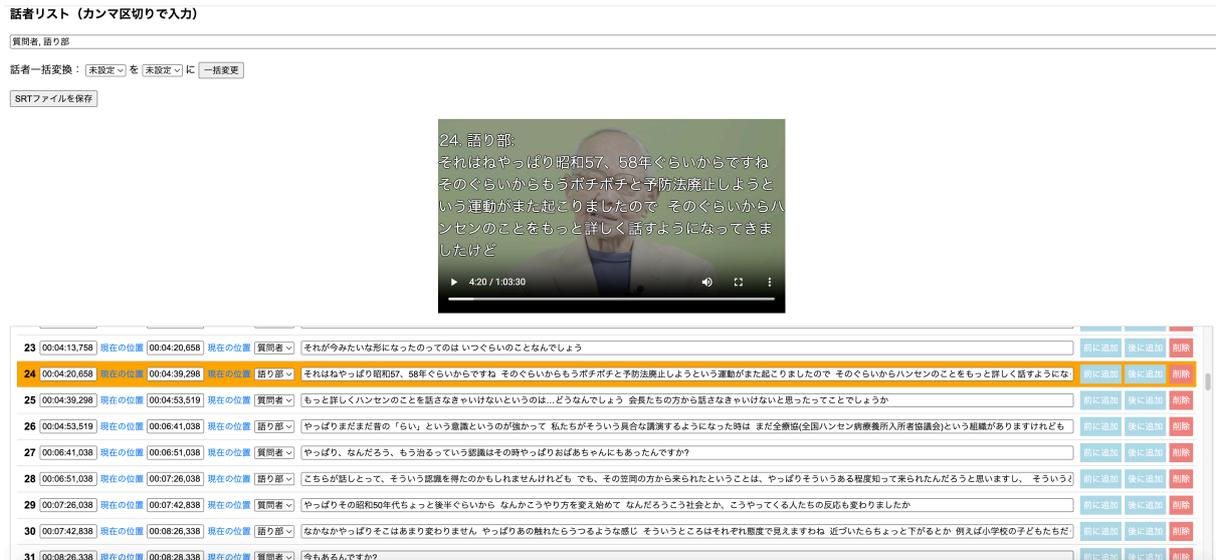


図 1: 文字起こし補助システム

表 1: 構築したデータセットの詳細

id	動画タグ	質問開始時間	質問終了時間	質問テキスト	回答開始時間	回答終了時間	回答テキスト
1	中尾会長_20241108	00:00:33,200	00:00:47,200	初めてこんな風に愛生園に来る人たちにお話をされ始めたのっていいのはいつぐらいのことなんですか?	00:00:50,200	00:01:01,899	初めて言ったら昭和30年代の終わり頃からかなそのくらいからですね
2	中尾会長_20241108	00:01:01,899	00:01:07,159	となるともう60年くらい前ですね	00:01:07,159	00:01:36,739	なりますねその頃は一般の方々よりも看護学校の生徒さんがよく見学を来られていてそんな人たちに一人だけじゃなくて何人もが話し相手になるという形でやりましたから集団で相手に話を伝えるという方法でやりましたから
3	中尾会長_20241108	00:01:36,739	00:01:46,640	看護の学生さんたちはどんな話が聞きたくて愛生園に来ていたのでしょうか	00:01:46,640	00:02:24,500	ハンセンのことはほとんどなくて大体普通の教科の中でハンセンが入ってくるとその頃からぼちぼち抜けていっていったようですね教科の中でもう済んだというような形になってきかけとったような時代ですからそれでも古いハンセンの話は私たちの方から仕掛けていくという形で持っていましたけど

問応答システム構築においては、ファインチューニングや他の手法を利用したの検索精度向上が今後の課題となる。

の検索精度向上を目指す。

6 おわりに

本論文では、ハンセン病回復者の語り部による証言記録データセットを構築し、質問応答システム構築に向けたベクトル検索の精度を検証した。実験の結果、Top-1の精度ではGLuCoSE-base-ja-v2が19.01%で最も高い精度を示した。Top-3の精度ではSarashina-Embedding-v1-1Bが35.54%で最も高い精度を示した。一方で、どの埋め込みモデルも20%以下の精度と低く、ファインチューニングなしのモデルでは十分な精度は得られなかった。今後の課題として、ファインチューニングや他の手法を利用して

謝辞

本研究の実施にあたって、長島愛生園入所者自治会 会長 中尾 伸治 様に語り部として多数の証言をいただきました。証言記録の作成にあたっては、長島愛生園関係者の皆様および、岡山大学 学術研究院 教育学域 国吉康雄記念・美術教育研究と地域創生講座 才土真司 特任准教授より多大なるご協力をいただきましたこと、深く感謝いたします。また、本論文の筆頭著者は博士後期課程学生であり、公益財団法人 大本育英会様から奨学金を賜り、本研究の遂行に専念することができました。心から感謝申し上げます。

参考文献

- [1] Musée d'Orsay. Digital technology- hello vincent., 2024. (<https://www.musee-orsay.fr/en/articles/digital-technology-hello-vincent-275618> 2025/1/8 アクセス).
- [2] USC Shoah Foundation. Dimensions in testimony., 2016. (<https://sfi.usc.edu/dit> 2025/1/8 アクセス).
- [3] David Traum, Andrew Jones, Kia Hays, Heather Maio, Oleg Alexander, Ron Artstein, Paul Debevec, Alesia Gainer, Kallirroi Georgila, Kathleen Haase, Karen Jungblut, Anton Leuski, Stephen Smith, and William Swartout. New dimensions in testimony: Digitally preserving a holocaust survivor's interactive storytelling. In Henrik Schoenau-Fog, Luis Emilio Bruni, Sandy Louchart, and Sarune Baceviciute, editors, **Interactive Storytelling**, pp. 269–281, Cham, 2015. Springer International Publishing.
- [4] トラウムデービッド. 双方向型歴史学習の支援のための対話システム技術の活用. 人工知能, Vol. 31, No. 6, pp. 806–812, 2016.
- [5] Ron Artstein, Anton Leuski, Heather Maio, Tomer Mor-Barak, Carla Gordon, and David R. Traum. How many utterances are needed to support time-offset interaction? In Ingrid Russell and William Eberle, editors, **FLAIRS**, pp. 144–149. AAAI Press, 2015.
- [6] 安田晴彦. まるでそこにいる人と話をしているような、新しい感動対話体験 AI 映像対話システム「talk with」. 人工知能学会研究会資料 言語・音声理解と対話処理研究会, Vol. 96, p. 30, 2022.