

報告書

1 今週の進捗

- Anime guide PDF について
- アニメ用語集の HTML を用いた SQL RAG の実装

2 Anime guide PDF について

Anime guide PDF は, 表面的な視覚情報とは別に, 内部に透明テキストが埋め込まれた特殊な形式であることが確認された. この構造により, 標準的な PDF のテキスト抽出処理が期待通りに機能せず, データの前処理において重大な課題となる. 図 1 にその画像を示す.

特に, この PDF を RAG の外部知識ソースとして直接利用した場合, 抽出されるコンテキストに大量のノイズが混入する. 内部の透明テキストには, OCR (Optical Character Recognition) の誤認識結果や, 本来の文脈とは無関係なレイアウト情報などが含まれることが多い.



図 1: 透明テキスト付き PDF の図.

3 アニメ用語集の HTML を用いた SQL RAG の実装

本研究では, アニメ制作の専門用語を解説した HTML ファイルを構造化データに変換し, それを外部知識源として利用する RAG (Retrieval-Augmented Generation) システムを実装した. 具体的には, HTML から用語と説明のペアを抽出し, PostgreSQL データベースに格納した. このデータベースに対して, キーワード検索とベクトル検索の 2 種類のリトリバル手法を適用し, 応答生成の品質を比較検証した.

3.1 データセットの構築

知識源として、アニメ『SHIROBAKO』公式サイト用語集ページを HTML 形式で取得した。Python のライブラリである BeautifulSoup を用い、HTML 構造を解析することで用語、読み、説明のペアを抽出した。抽出したデータは、Supabase 上に構築した PostgreSQL データベース内の 'anim_glossary' テーブルに格納した。構築したデータセットの一部を Table 1 に示す。

3.2 使用モデル

本システムでは、応答生成のための大規模言語モデル (LLM) と、テキストをベクトル化するためのエンベディングモデルをそれぞれ選定した。

- **LLM:** ELYZA, Inc. が開発した `elyza/Llama-3-ELYZA-JP-8B` を使用した。このモデルは、80 億 (8 billion) のパラメータを持ち、日本語の指示追従性に優れている。
- **エンベディングモデル:** OpenAI が提供する `text-embedding-3-small` を使用した。このモデルは、入力テキストを 1536 次元のベクトルに変換する。

3.3 実験手法と結果

構築したデータベースに対し、2 種類のリトリバル手法を用いて性能を比較した。

3.3.1 キーワード検索

SQL の ILIKE 句を用い、ユーザーの質問文字列が 'term' または 'description' カラムに部分一致（大文字小文字を区別しない）するレコードを検索した。

3.3.2 ベクトル検索

PostgreSQL の拡張機能である `pgvector` を利用した。データベース内の各用語の説明文を事前にエンベディングモデルでベクトル化し、`embedding` カラムに保存した。ユーザーの質問も同様にベクトル化し、コサイン類似度を用いて意味的に類似したレコードを検索した。実験では以下の 2 パターンのパラメータで検証を行った。

- **パターン A:** `match_threshold = 0.5, match_count = 3`
- **パターン B:** `match_threshold = 0.2, match_count = 5`

3.3.3 検索結果の比較

4 種類の質問に対する各手法の検索結果を比較する。

質問 1: 「アフレコ」

質問 2: 「アフレコについて教えて。」

質問 3: 「音監と各話作監ってなにが違うの？」

質問 4: 「A パートと B パートってなにが違うの？」

質問 1: アフレコ

----- キーワード検索の回答 -----

アフレコとは、アフター・レコーディングの略で、映像に合わせて音声を収録することを指します。アニメの場合は、完成版の映像ではなく、アフレコ用の映像を使用することが多くあります。

[参照した情報]

- 用語: アフレコ 説明: アフター・レコーディングの略。映像に合わせて音声を収録する事。アニメの場合は完成版の映像ではなく、アフレコ用の映像を使用することが多い。逆に、先に収録した音声にあわせて映像を制作する手法を「プレスコ」(プレスコアリング)という。
- 用語: 音戻し 説明: アフレコした音声、ダビングしたSE音などを映像と合わせること。
- 用語: 音響制作 説明: 音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ

----- パターン A の回答 -----

アフレコとは、アニメや映画などで声優が声を当てる録音作業のことです。アフレコの際には、原画や絵コンテを基に、声優がセリフや効果音を収録します。アフレコは通常、スタジオで行われます。

[参照した情報]

----- パターン B の回答 -----

アフレコとは、アフター・レコーディングの略で、映像に合わせて音声を収録することを指します。アニメの場合は、完成版の映像ではなく、アフレコ用の映像を使用することが多くあります。

[参照した情報]

- 用語: 音戻し 説明: アフレコした音声、ダビングしたSE音などを映像と合わせること。
- 用語: アフレコ 説明: アフター・レコーディングの略。映像に合わせて音声を収録する事。アニメの場合は完成版の映像ではなく、アフレコ用の映像を使用することが多い。逆に、先に収録した音声にあわせて映像を制作する手法を「プレスコ」(プレスコアリング)という。
- 用語: M 説明: 劇伴(演劇やアニメで使用する伴奏音楽)楽曲のこと。
- 用語: 絵コンテを切っている 説明: 絵コンテを描いている事を表す。
- 用語: 音響制作 説明: 音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ

質問 2: アフレコについて教えて。

----- キーワード検索の回答 -----

アフレコとは、アニメや映画、テレビドラマなどの音声収録方法の一つで、既に完成した映像に後から声を当てることを指します。声優が演技をし、収録した音声を映像に合わせて編集することで、キャラクターに命を吹き込む作業です。アフレコには、セリフや効果音、BGMなどを加える作業も含まれます。アフレコは、主にスタジオで行われ、声優と音響監督、エンジニアらが参加します。

[参照した情報]

----- パターン A の回答 -----

アフレコとは、アニメや映画などの音声収録の工程で、既に完成した映像に声を当てる作業のことです。声優が台本に沿ってセリフを読み上げることで、キャラクターに命を吹き込みます。アフレコには、通常の収録と同様に、演出家や監督が指示を出し、声優が演技を行います。アフレコの際には、声優は映像を見ながら収録することが多く、表情や動きを確認しながら演技をします。アフレコは、作品の世界観や雰囲気をもより深く表現するために重要な工程です。

[参照した情報]

----- パターン B の回答 -----

アフレコとは、アフター・レコーディングの略で、映像に合わせて音声を収録することを指します。アニメの場合、完成版の映像ではなく、アフレコ用の映像を使用することが多くあります。逆に、先に収録した音声にあわせて映像を制作する手法を「プレスコ」と呼びます。

[参照した情報]

- 用語：音戻し 説明：アフレコした音声、ダビングしたSE音などを映像と合わせること。
- 用語：アフレコ 説明：アフター・レコーディングの略。映像に合わせて音声を収録する事。アニメの場合は完成版の映像ではなく、アフレコ用の映像を使用することが多い。逆に、先に収録した音声にあわせて映像を制作する手法を「プレスコ」（プレスコアリング）という。
- 用語：絵コンテを切っている 説明：絵コンテを描いている事を表す。
- 用語：M 説明：劇伴（演劇やアニメで使用する伴奏音楽）楽曲のこと。
- 用語：UP 説明：カメラを被写体に寄せる事。

質問 3：音監と各話作監ってなにが違うの？

----- キーワード検索の回答 -----

----- パターン A の回答 -----

「音監」は音響監督の略称であり、作品における音響関連の責任者です。一方、「各話作監」は各話の作画監督の略称です。両者は異なる役割を担っています。音響監督は音響面での演出や指揮を担当しますが、作画監督は絵に関する部分の統括や指揮を担当します。したがって、音監と各話作監は全く別の仕事をしている人たちです。

[参照した情報]

- 用語：音監 説明：音響監督の略称。
- 用語：音響監督 説明：作品における音響関連の責任者。声優のキャスティングや、劇中で使用する音楽の発注、ダビングの際には、BGMや効果音をどのように使うかなどの音響を演出し、声優への演技指導なども行う。
- 用語：音響制作 説明：音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ

----- パターン B の回答 -----

音監（音響監督）と各話作監は異なる役割を持つ人物です。

音響監督は、作品の音響関連の責任者として、声優のキャスティングや音楽の発注、ダビングの指揮、声優へ

の演技指導などを担います。

一方、各話作監は、1話数を担当する作画監督のことで、作画面での絵コンテに基づいて、原画マンや動画マンに指示を与えます。作画のクオリティや統一感を保つために重要な役割を果たします。

大きな違いは、音監が音響面を主に担当するのに対し、各話作監は作画面を主に担当する点です。両者は異なる部署や役割を担いながら、協力してアニメーション作品を作り上げています。

[参照した情報]

- 用語：音監 説明：音響監督の略称。
- 用語：音響監督 説明：作品における音響関連の責任者。声優のキャスティングや、劇中で使用する音楽の発注、ダビングの際には、BGMや効果音をどのように使うかなどの音響を演出し、声優への演技指導なども行う。
- 用語：音響制作 説明：音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ
- 用語：各話作監 説明：1話数を担当する作画監督のこと。
- 用語：演出 説明：絵コンテを元に、各セクションに演技や映像イメージ等を指示するポジション。制作工程の全体を通して打ち合わせとチェックに立ち会う。

質問 4: A パートと B パートってなにが違うの？

----- キーワード検索の回答 -----

----- パターン A の回答 -----

A パートと B パートは、主にテレビアニメや漫画などで使われる用語です。一般的には、物語を2つ以上の部分に分けた場合、各部分を「A パート」や「B パート」と呼びます。

大きな違いとしては、A パートは通常、メインの主人公や主要キャラクターが中心となる展開を描くことが多く、物語の本筋や核心部分を担います。一方、B パートはサブ的な役割や補助的な要素を含むことが多く、物語全体のバランスや緩急を調整するために使用されます。

ただし、明確な定義やルールはなく、作品や作者によって解釈や使い方が異なるため、A パートと B パートの区別は曖昧な面もあります。

[参照した情報]

----- パターン B の回答 -----

A パートと B パートは、CM 前の前半部分と後半部分を指します。具体的には、A パートはCM 前の前半部分で、B パートはCM 後の後半部分を指します。ただし、明確な区別や定義はなく、制作現場や作品によって異なる場合があります。

[参照した情報]

- 用語：A パート 説明：CM 前の前半部分
- 用語：回収 説明：アニメ業界では、スタッフが一つの場所に集まって作業をしていない場合が多く、制作

がスタッフの作業場所まで素材の集配業務を行っている

- 用語: 音響制作 説明: 音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ
 - 用語: 演出 説明: 絵コンテを元に、各セクションに演技や映像イメージ等を指示するポジション。制作工程の全体を通して打ち合わせとチェックに立ち会う。
 - 用語: 入れ切り 説明: 作業用の素材を全て渡し切ったことを指す
-

3.4 考察

実験結果から、ベクトル検索はキーワード検索に比べて、ユーザーの曖昧な質問や意図を汲み取った検索が可能であり、RAG システムの全体的な性能を大幅に向上させることが確認された。特に、質問に直接的なキーワードが含まれない場合において、その優位性は顕著であった。また、ベクトル検索のパラメータである `match_threshold` と `match_count` はトレードオフの関係にある。閾値を下げ、取得件数を増やす (パターン B) ことで、より多くの潜在的な関連情報を取得できる一方、ノイズの増加を招くリスクもある。したがって、アプリケーションの目的に応じてこれらのパラメータを適切に調整することが重要である。

4 今後の課題

- アニメ用語集の HTML を用いた GPTs との比較
- アニメ制作に関するデータの収集

参考文献

表 1: 構築したアニメ用語集データセットのサンプル

用語	説明
UP	カメラを被写体に寄せる事。
アフレコ	アフター・レコーディングの略。映像に合わせて音声を収録する事。アニメの場合は完成版の映像ではなく、アフレコ用の映像を使用することが多い。逆に、先に収録した音声にあわせて映像を制作する手法を「プレスコ」（プレスコアリング）という。
色付き	編集や音響素材に使用するムービーに色がついている事。完成映像に近い状態。色がついていない線画の場合は、「線撮」または「白い」などと言う。線画の状態では「色付き」に比べて情報量が少なく、精度の高い編集や音響作業は望めない。制作状況の悪化で、昨今のアニメーション制作においては、色付きで音響作業が行えない場合が多々ある
色指定・検査	色彩設計がシーン毎に作成した、キャラクターをどんな色で塗るかを記載した”カラーモデル”と呼ばれるものがあり、色指定は、1カット毎にどのカラーモデルを使用して...
入れ切り	作業用の素材を全て渡し切ったことを指す
A パート	CM 前の前半部分
絵コンテ	シナリオを元に、カット割り、画面の構図、キャラクターの芝居、セリフや秒数などをイラストとともに記入したもの。アニメーションの映像の設計図と言える
絵コンテ撮	コンテの絵素材を使って撮影すること。ほぼ動きのない映像となる。
絵コンテを切っている	絵コンテを描いている事を表す。
演出	絵コンテを元に、各セクションに演技や映像イメージ等を指示するポジション。制作工程の全体を通して打ち合わせとチェックに立ち会う。
M	劇伴（演劇やアニメで使用される伴奏音楽）楽曲のこと。
オールコンテ撮	すべて絵コンテ撮影した素材を繋げたムービー
重いカット	カロリーが重たい＝作業が大変なカットという意味
音戻し	アフレコした音声、ダビングした SE 音などを映像と合わせる事。
音監	音響監督の略称。
音響制作	音響スタッフ、声優、制作スタッフへの手配や連絡、音響スタジオの手配、アフレコ台本の発注などを行うセクション、またはその業務を行うスタッフ
音響監督	作品における音響関連の責任者。声優のキャスティングや、劇中で使用する音楽の発注、ダビングの際には、BGM や効果音をどのように使うかなどの音響を演出し、声優への演技指導なども行う。
回収	アニメ業界では、スタッフが一つの場所に集まって作業をしていない場合が多く、制作がスタッフの作業場所まで素材の集配業務を行っている
各話作監	1 話数を担当する作画監督のこと。
カット袋	カット単位でレイアウトや原画、動画、タイムシートなどの素材をいれる袋。厚手の紙で出来ており素材保護をすると同時に、カット内容の確認や作業伝票としての役割も担う。