

報告書

1 今週の進捗

- Wikipedia を用いた RAG システムの比較

2 Wikipedia の RAG システムの比較

Wikipedia の情報を外部知識として利用する 3 つの異なるアプローチを比較し, その特性を評価する. 比較対象は, RAG を利用しないベースライン, API ベースの RAG, そして特定カテゴリの知識に特化したデータベース (DB) ベースの RAG である.

2.1 比較対象システム

2.1.1 システム A: RAG なし (ベースライン)

本システムは, RAG 機構を一切利用せず, LLM が持つ内部知識のみで応答を生成する. これは, モデルの基本的な能力とハルシネーションのリスクを評価するための基準となる.

表 1: システム A の主要パラメータ

項目	値 / モデル名
使用モデル	elyza/Llama-3-ELYZA-JP-8B (事前学習済みモデル)
max_new_tokens	512
temperature	0.2
repetition_penalty	1.1

2.1.2 システム B: API ベース RAG

本システムは, LangChain の WikipediaRetriever を利用する. ユーザーの質問を検索クエリとして, Wikipedia API をリアルタイムで呼び出し, 取得した記事をコンテキストとして LLM に与える. 事前のデータ準備が不要である一方, 検索はキーワードベースに限定される.

表 2: システム B の主要パラメータ

項目	値 / モデル名
使用モデル	elyza/Llama-3-ELYZA-JP-8B (事前学習済みモデル)
Retriever	WikipediaRetriever(lang="ja", top.k=3, doc.content_chars_max=512)
max_new_tokens	512
temperature	0.2
repetition_penalty	1.1

2.1.3 システム C: カテゴリ特化 DB-RAG

本システムは、Wikipedia の「アニメーション」カテゴリに関連する記事を事前にスクレイピングし、独自のベクトルデータベース (ChromaDB) を構築する。検索は、このローカル DB に対してセマンティック検索で行われる。ドメインに特化した高精度な検索が期待できるが、事前準備と知識の更新にコストを要する。今回はサーバの容量の都合上、最大 20 記事しか利用していない。

表 3: システム C の主要パラメータ

項目	値 / モデル名
使用モデル	elyza/Llama-3-ELYZA-JP-8B (事前学習済みモデル)
Embedding モデル	intfloat/multilingual-e5-large-instruct
Vector DB	ChromaDB
Retriever	vectorstore.as_retriever(search_kwargs={"k": 3})
max_new_tokens	512
temperature	0.2
repetition_penalty	1.1

2.2 実験設定

各システムの特性を比較するため、以下の 3 つの異なる性質を持つ質問を用意した。

質問 1 (専門用語): アニメ制作における「撮影」とは具体的にどのような工程か。

質問 2 (抽象的・分析的): アクション作品に共通する作画上の特徴は何か。

質問 3 (ハルシネーション誘発): アニメの歴史における幻の技法「虹彩透過法」について説明せよ。

----- システム A のプロンプト -----

あなたはアニメ制作に詳しい専門家です。ユーザーの質問に親切に答えてください。【質問】

----- システム B, C のプロンプト -----

あなたはコンテキストを参考に回答するアニメ制作に詳しいアシスタントです。
コンテキストに情報がない場合は、「内部情報のみを使います。」と答えてから、内部情報を使って回答してください。

【コンテキスト情報】

【質問】

2.3 実験結果

2.3.1 質問 1 (専門用語) への応答

システム A (RAG なし) の応答

- 生成時間: 126.01 sec.
- 参照情報: なし (内部知識のみ).

----- システム A の回答 -----

アニメ制作の「撮影」は、実写映像と異なり、2Dや3Dの動画をカメラで「撮る」作業ではありません。代わりに、原画や背景美術、特殊効果などを組み合わせて、完成した絵コンテ通りの動きや演出を実現する工程を指します。

アニメの撮影工程は、主に以下の手順で行われます：

1. ****原画の作成****：原作者や監督が描いた原画をもとに、各シーンの絵コンテを作成。
2. ****レイアウトの作成****：絵コンテに基づいて、各シーンの構図やキャラクターの配置を決める。
3. ****背景美術の作成****：建物、自然環境、道具など、各シーンに必要な背景を描く。
4. ****特殊効果の作成****：火花、水、煙、光など、各シーンに必要な特殊効果を描く。
5. ****撮影用の素材準備****：上記の原画、レイアウト、背景美術、特殊効果を、撮影用の素材としてデジタルデータ化。
6. ****撮影****：撮影用ソフトウェアを使用して、素材を組み合わせ、カットごとの画像を生成。
7. ****編集****：複数のカットを繋ぎ合わせ、ストーリーに沿って整理し、テンポやリズムを調整。
8. ****音声の追加****：音響効果、音楽、セリフを加え、完成したアニメ作品を仕上げる。

このように、アニメの「撮影」は、実写とは異なる独自のプロセスで行われ、絵コンテや原画を元に、各シーンを組み立てて動画を完成させる工程です。

システム B (API ベース RAG) の応答

- **生成時間**: 117.5 sec.
- **参照情報**: アニメ監督, テレビアニメ, アニメーション

----- システム B の回答 -----

内部情報のみを使います。

アニメ制作における「撮影」は、実写映画やドラマの「撮影」と同様に、カメラを使用して映像を収録する工程を指します。ただし、実写映画やドラマの「撮影」が、被写体を実際に撮影することを目的として行うのに対し、アニメ制作の「撮影」は、既に描かれた絵やCGをカメラで捉え、映像として収録する工程を指します。具体的には、背景やキャラクターの絵をスキャンやデジタル化し、カメラを使用して映像を収録することで、実写のように見える映像を生成します。

システム C (カテゴリ特化 DB-RAG) の応答

- **生成時間**: 34.2 sec.
- **参照情報**: アニメーション, アニメーション映画

----- システム C の回答 -----

内部情報のみを使います。

アニメ制作における「撮影」とは、完成した原画や設定資料をもとに、実際の映像を撮影する工程を指します。この工程では、カメラやレンズを使用して、原画や設定資料を忠実に再現した映像を撮影します。撮影方法は、実写撮影と同様で、フレームレートやシャッター速度などを調整して、滑らかに動く映像を撮影します。

2.3.2 質問 2 (抽象的・分析的) への応答

システム A (RAG なし) の応答

- 生成時間: 129.64 sec.
- 参照情報: なし (内部知識のみ).

----- システム A の回答 -----

アクション作品に共通する作画上の特徴として、以下のような点が挙げられます。

1. 効果線や爆発エフェクトなどの派手な演出: アクションシーンでは、動きや衝撃を強調するために効果線や爆発エフェクトなどを使用し、視覚的にインパクトを与えることが重要です。
2. スピード感と迫力のあるカメラワーク: アクションシーンでは、速いカット割りや大胆なパンやズームを用いてスピード感や迫力を表現します。
3. 動きの激しいキャラクターの描写: アクション作品では、キャラクターの動きが激しくなるため、作画もそれに応じて激しく、ダイナミックに描く必要があります。
4. 血しぶきや破壊された背景などのグロテスクな要素: アクション作品には、暴力や破壊を伴うことが多いため、血しぶきや破壊された背景など、グロテスクな要素が含まれることがあります。
5. 緻密で詳細な背景設定: アクション作品では、世界観や舞台となる場所の設定が重要であり、緻密で詳細な背景設定が求められます。
6. カラーのコントラストを利用した色彩設計: アクション作品では、明暗や暖色・寒色を活用して、緊張感や熱気を表現することができます。
7. 適度なデフォルメとリアリティのバランス: アクション作品では、デフォルメとリアリティのバランスを取ることで、見栄え良くかつ説得力のあるアクションシーンを描くことができます。

これらの特徴を組み合わせることで、アクション作品は、視聴者を引き付けるエキサイティングな映像を創造することができます。

システム B (API ベース RAG) の応答

- 生成時間: 157.9 sec.
- 参照情報: 東映アニメーション, アニメーション映画, 庵野秀明

----- システム B の回答 -----

内部情報のみを使います。

東映アニメーションは日本のアニメ制作会社であり、現存する日本のアニメ制作会社としては最も歴史が古い。同社は1947年1月に東映の前身である東横映画に開発部が設置されたことを始まりとしている。東映動画は教育映画事業から生み出され、常時120班ほどの上映班が全国を巡回し、「十六ミリ映写隊」等の活動を行っていた。終戦直後には、講堂や映写機材も多くが戦災で不足していたことから、1940年代後半を中心に映画館への引率観覧が積極的に行われ、それは「映画教室」として全国的に波及していった。

このような背景から、東映アニメーションの作画上の特徴は、教育映画や「十六ミリ映写隊」の影響を受けている可能性があります。具体的には、以下のような特徴が挙げられます。

教育的な要素を含む作品が多い: 東映アニメーションは教育映画事業から生み出されたため、教育的な要素を含む作品が多くあります。

地域密着型の活動：東映動画は常時 120 班ほどの上映班が全国を巡回し、「十六ミリ映写隊」等の活動を行っていたため、地域密着型の活動を重視している可能性があります。

映画館での上映を意識した作画：東映アニメーションは映画館での上映を意識した作画を心掛けている可能性があります。例えば、映画館で上映するための高品質な映像や音響設備に対応した作画を目指すことで、作品のクオリティを向上させることができます。

以上の特徴は、東映アニメーションの作画上の特徴として挙げられます。ただし、内部情報のみを使用した回答のため、実際の作画上の特徴とは異なる可能性があります。

システム C (カテゴリ特化 DB-RAG) の応答

- 生成時間: 50.6 sec.
- 参照情報: Portal: アニメ, アニメーション

----- システム C の回答 -----
内部情報のみを使います。

アクション作品に共通する作画上の特徴としては、以下のようなものが挙げられます。

動きの激しい場面では、線や塗りが荒くなったり、省略されたりすることがあります。

アクションシーンでは、背景や小物などを簡略化や省略することで、主人公などの主要なキャラクターに視線を集中させることができます。

動きの速いアクションシーンでは、コマ送りのように、短時間で多くのコマを使用することがあります。

アクション作品では、力強さやスピード感を出すために、線や塗りに勢いや流れを与えることが重要です。

作画の工程では、下絵や中間色を省略することもあります。

これらの特徴は、アクション作品において、動きやスピード感を重視し、視聴者の興味を引くために用いられる作画上の技法です。

2.3.3 質問 3 (ハルシネーション誘発) への応答

システム A (RAG なし) の応答

- 生成時間: 64.14 sec.
- 参照情報: なし (内部知識のみ).

----- システム A の回答 -----

「虹彩透過法」という技法は、実際には存在しないアニメの技法です。虹彩とは、眼球の中心にある部分で、光を集める役割を果たす器官です。虹彩自体が持つ色や構造を利用して絵画やイラストレーションを描くという手法はありますが、「虹彩透過法」は架空の技法です。

アニメの歴史においても、虹彩透過法は知られていないし、使用された記録もありません。アニメの技法は常に進化しており、新しい表現方法や技術が開発されていますが、虹彩透過法はその中には含まれていません。

したがって、具体的な情報や解説することができませんが、もし、架空の世界やフィクションの中で虹彩透過法が登場するのであれば、創作物として捉える必要があります。

システム B (API ベース RAG) の応答

- 生成時間: 124.7 sec.
- 参照情報: 僕のヒーローアカデミア

----- システム B の回答 -----
内部情報のみを使います。

内部情報では、提供されたコンテキストに「虹彩透過法」という単語は見つかりませんでした。したがって、回答は以下の通りです。

「虹彩透過法」は、残念ながら私の内部情報には存在しない用語のため、説明することができません。

システム C (カテゴリ特化 DB-RAG) の応答

- 生成時間: 37.8 sec.
- 参照情報: アニメーション, アニメーション映画

----- システム C の回答 -----
内部情報のみを使います。

残念ながら、内部情報には「虹彩透過法」に関する記述が見つかりませんでした。したがって、回答できません。

3 今後の課題

- RAG の論文の調査
- アニメ制作に関するデータの収集
- PlanetScale を用いた RAG の導入

参考文献