

報告書

1 今週の進捗

- anime-offline-database の調査
- RAG の実装

2 anime-offline-database の概要

anime-offline-database は、複数のアニメ情報サイトからメタデータを集約し、相互参照を可能にすることを目的としたデータベースである。本稿執筆時点 (2025 年 29 週目) の情報に基づくと、その特徴は以下のよう
に整理できる。

- **目的:** MyAnimeList.net, AniDB.net, Kitsu.app をはじめとする 11 以上の主要なアニメ情報サイトからメタデータを集約し、作品情報の網羅性を高める。
- **データ規模:** 約 39,000 件以上のアニメ作品エントリを収録している。
- **ファイル形式:** 用途に応じて複数の形式で提供されている。主要なファイルは以下の通りである。
 - anime-offline-database.json: 人間が可読な形式で全データを含む。
 - anime-offline-database-minified.json: JSON を minify してファイルサイズを削減した形式。
 - anime-offline-database.jsonl: 1 行に 1 アニメデータを格納した JSON Lines 形式。大規模データのストリーム処理に適する。
 - *.zst: Zstandard アルゴリズムで圧縮されたファイル。
- **ライセンス:** Open Data Commons Open Database License (ODbL) v1.0 であり、自由な利用と再配布が可能である。
- **データ構造:** 各アニメエントリは、構造化された JSON オブジェクトとして定義される。主要なフィールドは以下の通りである。
 - sources: 各情報サイトへの URL リスト。
 - title: 作品の公式タイトル。
 - type: TV, MOVIE, OVA などの種別。
 - episodes: エピソード数。
 - status: FINISHED, ONGOING などの放送状態。
 - animeSeason: 放送が開始された年と季節 (SPRING, SUMMER, FALL, WINTER)。
 - studios, producers: 制作スタジオと製作会社のリスト。
 - tags: 作品のジャンル, テーマ, 特徴を示すタグのリスト。
 - relatedAnime: 続編やスピンオフなど、関連作品への URL リスト。
 - synonyms: 各国語でのタイトルや別名。

2.1 データ構造の具体例

各アニメエントリは、構造化された JSON オブジェクトとして定義される。具体的な構造を理解するため、作品「DEATH NOTE」のデータ例を以下に示す (リスト 1)。

Listing 1: 「DEATH NOTE」のデータ構造例 (一部抜粋)

```
1 {  
2   "sources": [  
3     "https://anidb.net/anime/4563",  
4     "https://anilist.co/anime/1535",  
5     "https://myanimelist.net/anime/1535",  
6     ...  
7   ],  
8   "title": "Death Note",  
9   "type": "TV",  
10  "episodes": 37,  
11  "status": "FINISHED",  
12  "animeSeason": {  
13    "season": "FALL",  
14    "year": 2006  
15  },  
16  "studios": [  
17    "madhouse"  
18  ],  
19  "relatedAnime": [  
20    "https://anidb.net/anime/8146",  
21    "https://myanimelist.net/anime/2994",  
22    ...  
23  ],  
24  "tags": [  
25    "anti-hero",  
26    "battle of wits",  
27    "crime",  
28    "detective",  
29    "fantasy",  
30    "mind games",  
31    "mystery",  
32    "police",  
33    "psychological",  
34    "shounen",  
35    "supernatural",  
36    "thriller",  
37    ...  
38  ]  
39 }
```

リスト 1 に示すように、作品の基本情報 (title, type, episodes など) に加え、複数の情報源へのリンク (sources)、放送時期 (animeSeason)、制作スタジオ (studios)、関連作品群 (relatedAnime)、そして作品の特徴を表す多数のタグ (tags) が構造的に格納されている。このデータセットは、このような「事実に基づいたメタデータ」を網羅的かつ構造的に提供する点で非常に価値が高いと言える。

3 RAG の実装

現在, 苦戦中...

4 今後の課題

- RAG の実装
- 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