

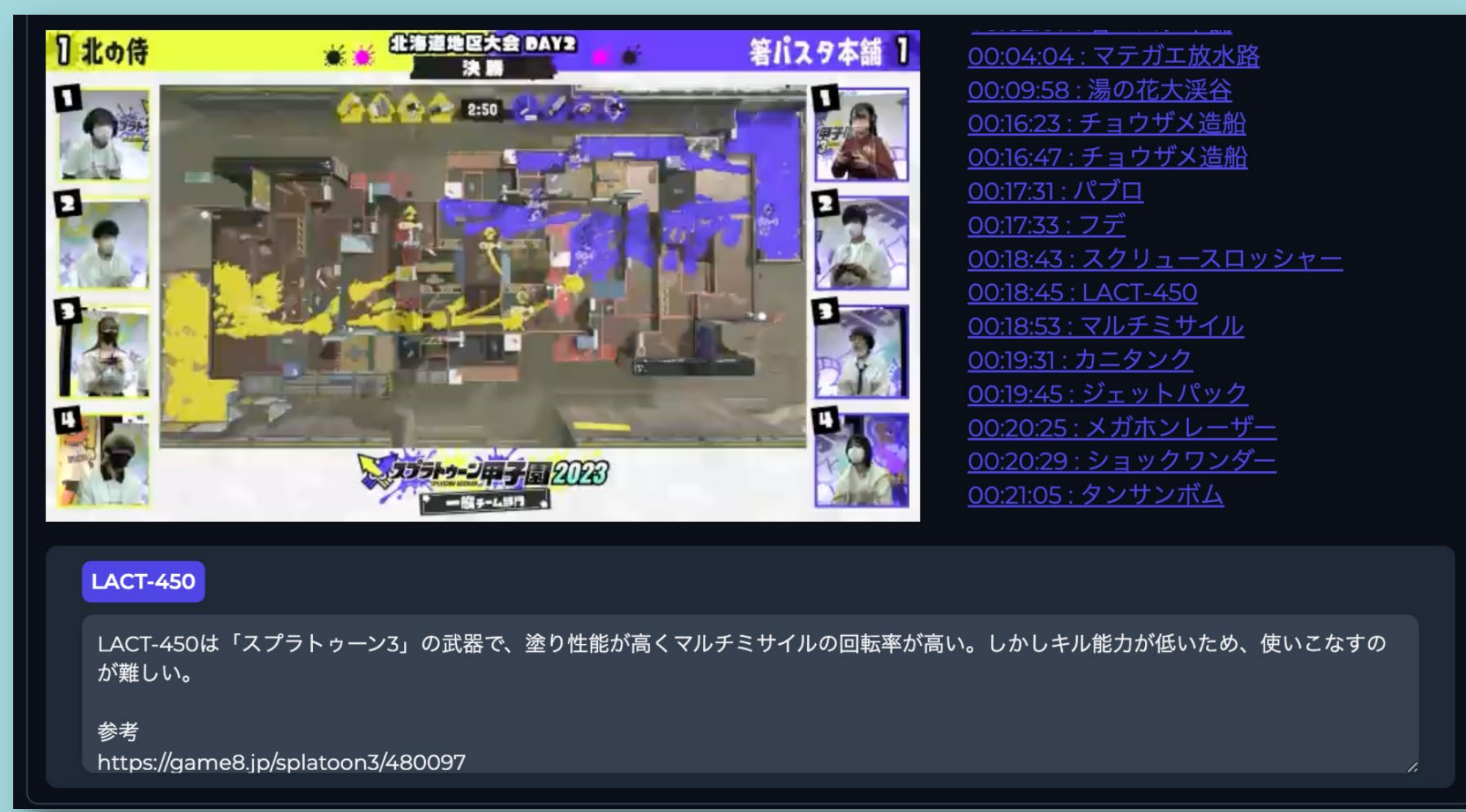
ニコニコ大百科データセットを更新し大量データ追加



- 10年ぶりのデータ更新を行いました。
- 2008年5月から2024年1月までの約51万投稿記事すべての記事ヘッダと記事本文データです。特にサブカルチャー文脈の用語に強い日本語のコーパスとして利用できます。
- 特にLLM関連の研究での利用を期待しています。
- IDRサイトから即日ダウンロード・利用が可能です。

ニコニコドメイン(動画・生放送)へのLLM応用デモ例

放送中に言及された用語の自動認識・自動解説生成



動画紹介記事の自動生成



階層型シーン分割(音声認識内容・視聴者コメント・画像認識内容等に基づく)



ニコニコ大百科 <https://dic.nicovideo.jp/>

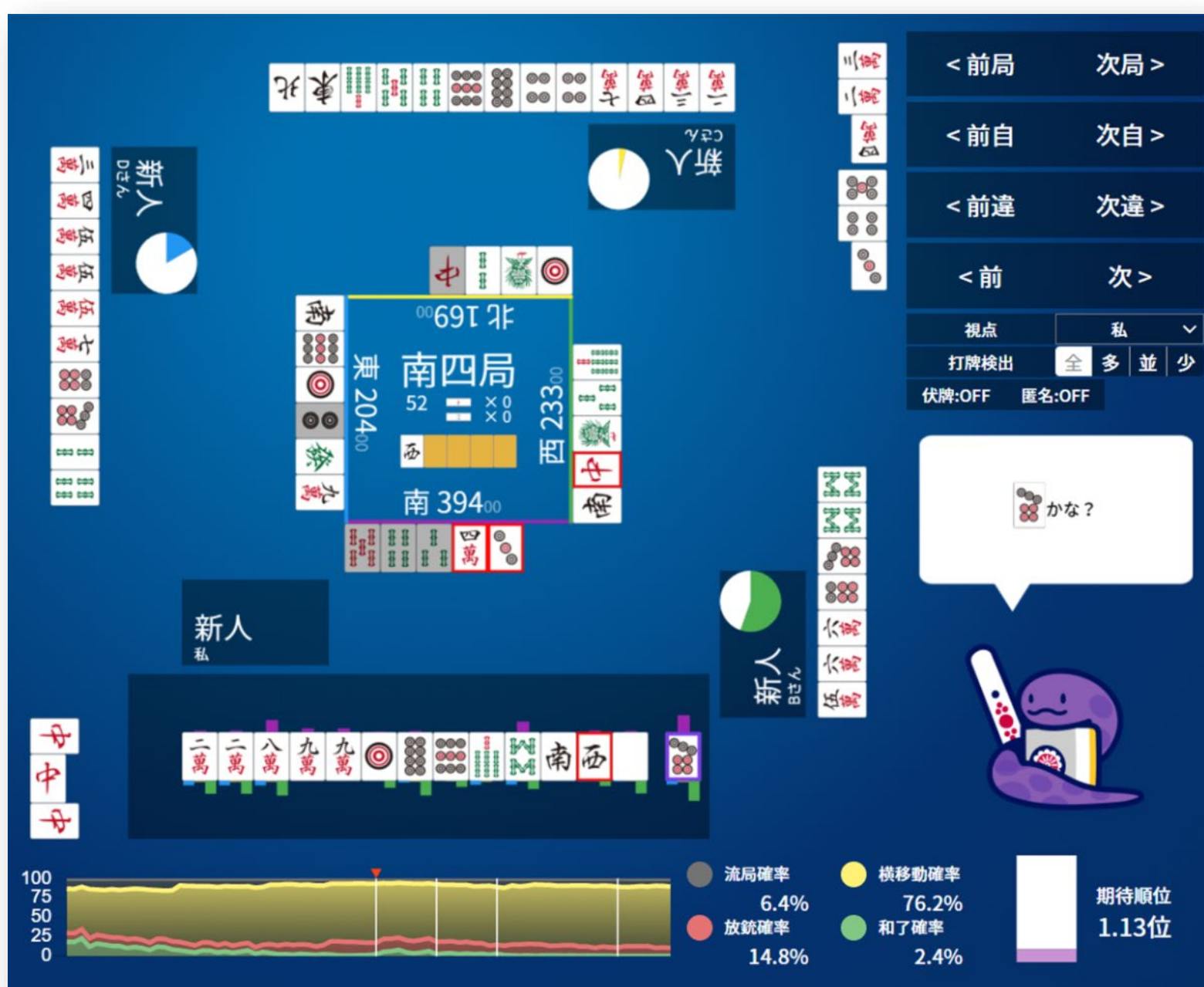
キャラクター声変換ソフト SEIREN VOICE



Deep Learningによる声変換を行うソフトウェア。

リアルタイム変換ではなくバッチ変換処理のため、変換品質が高いという特徴があります。

麻雀AI 牌譜解析「NAGA」



インターネット麻雀サイト「天鳳」にて十段を達成した麻雀AI「NAGA」による牌譜解析サービス。

自分の牌譜を解析し、麻雀の学習に役立てることが出来ます。

ニコニコにおけるレコメンデーション



ニコニコ動画や生放送のレコメンデーションシステムは、コンテンツベースや行動履歴ベース等の様々なアルゴリズムで実装したレコメンデーションエンジンを組み合わせたものになっています。最近、LLMで得られるコンテンツ特徴を利用するレコメンデーションエンジンの開発・導入も行っています。