

つくしプロジェクト

日本の古典文化を融合したマルチモーダル基盤モデルのためのデータインフラストラクチャの構築

日本古典籍と生成AIを融合し、過去の文化と現代の人々との距離を縮めるAIを開発！

生成AIチャット: 古典籍からのテキストを基に、現代日本語で対話を行います



- **IIF Tsukushi Viewer:** 生成AIチャットをサポートする画像ビューアで、関連サービスと連携して古典籍のテキストを現代日本語で探索します
- **Tsukushi Content Service:** IIF識別子を使いテキストデータを登録・検索するウェブサービスで、AIくずし字OCRとテキスト埋め込みモデルを利用しています
- **Tsukushi Chat Service:** ユーザーのプロンプトに基づき大規模言語モデルに問い合わせして回答を提供するサービスです。Claude3.5、GPT-4oなど、10種類のLLMから選択できます
- **つくしサーチ:** 日本古典籍データセットを対象としたテキスト検索サービスです

AI自動テキスト化／翻訳実験



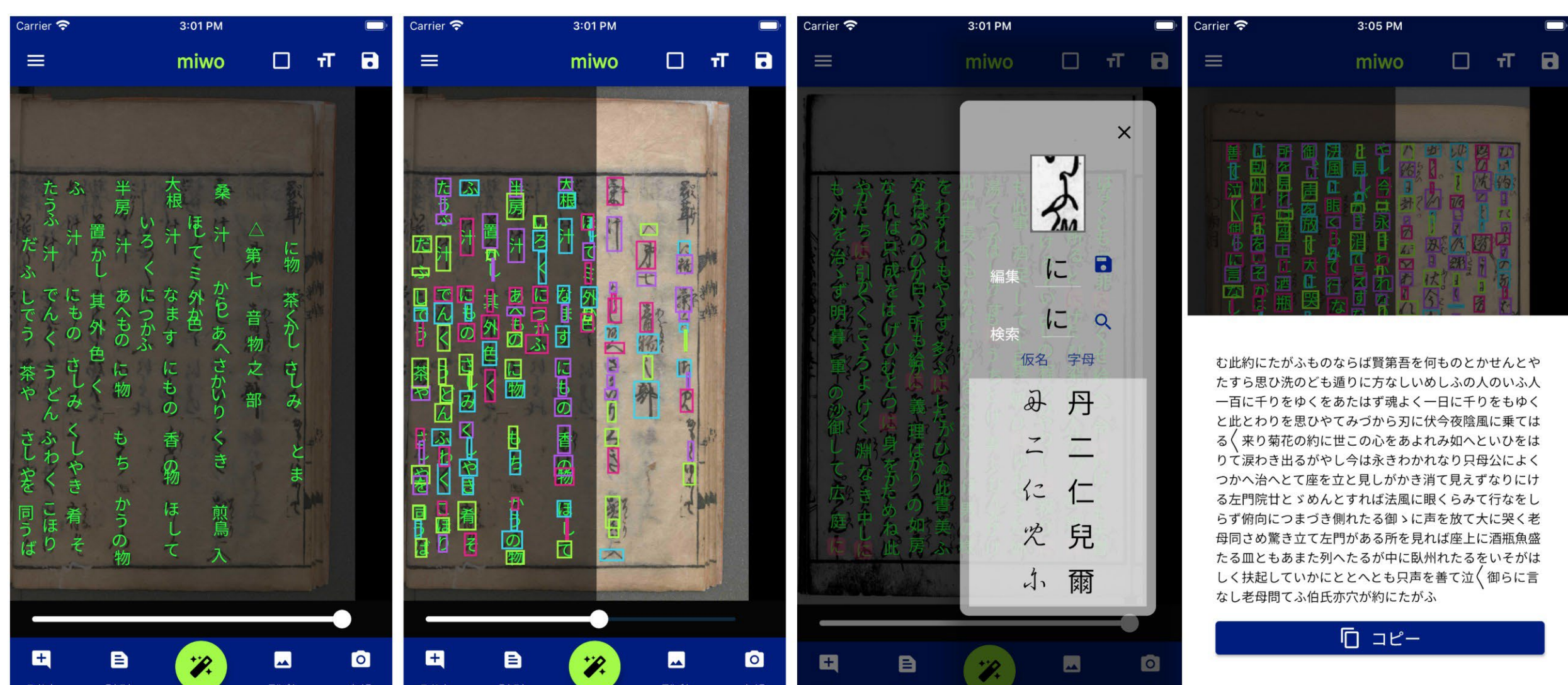
日本古典籍の挿絵のカラー化



Sakana AIが開発した画像生成モデルEvo-Nishikieを活用して日本古典籍の挿絵を対象としたカラー化（単色摺の墨摺絵から多色摺の錦絵への変換）に取り組めます。



くずし字認識アプリ



SCAN ME



「つくし」とは

「つくし」は、『源氏物語』第14帖「みをつくし」にちなんだ名前です。「みを」アプリがくずし字資料の海を旅するツールとなる一方、「つくし」はくずし字資料を調べ「つくす」ためのツールとなることを目指しています

SCAN ME



北本 朝展, 本間 淳, カラーヌワット タリン, "IIF Tsukushi Viewer: 日本古典籍へのアクセス性を向上させる生成AIチャット機能", 情報処理学会技術報告, 2024年7月

国文研プロジェクト型共同研究「日本の古典文化を融合したマルチモーダル基盤モデルのためのデータインフラストラクチャの構築」
ROIS戦略的研究プロジェクト「大規模言語モデルのための歴史的日本語テキストデータセットの開発」