



#みんなで翻刻

みんなで翻刻
MINNA DE HONKOKU

歴史資料を読み解く

歴史資料を解読するために

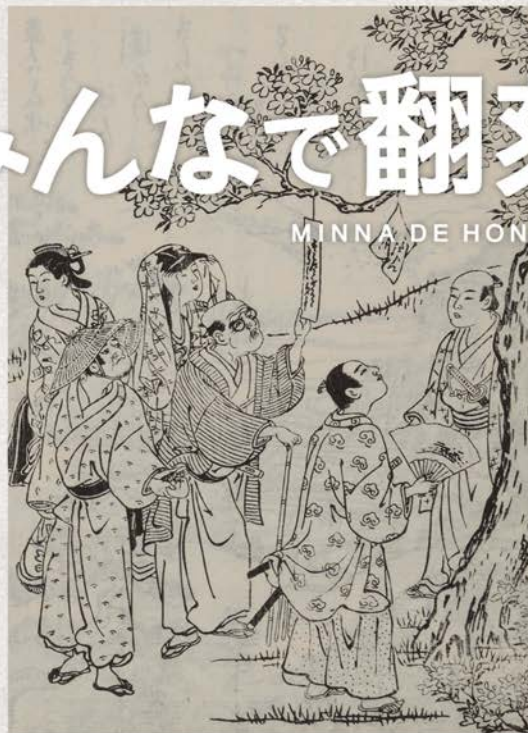
お問い合わせ

JOIN 参加する



みんなで翻刻

MINNA DE HONKOKU



仲間と、AIと協力して、
古文書解読に参加しよう！

JOIN

参加する

※現在新バージョンの試験運用中です。
安定稼働するまで旧バージョンも併せてご利用ください。

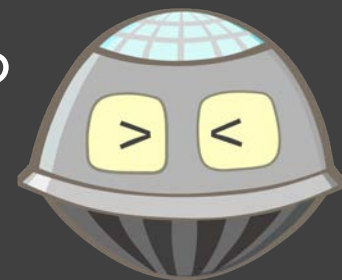
URL: <https://honkoku.org/>

みんなで翻刻

- 歴史資料の大規模データ化を目指す、
ユーザー参加型プロジェクト
- 京都大学古地震研究会が開発・運営
- Webサイト
 - <https://honkoku.org/>
 - Twitter (@CloudHonkoku)や Facebook ページもあります

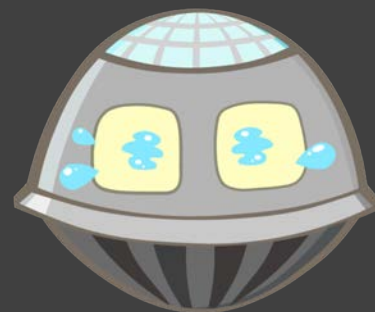
「翻刻」とは

- 古文書などの歴史資料に書かれた文章を、活字に起こし出版・再公開すること
- 「翻刻」することで、歴史資料は**データとして活用しやすくなる！**
 - 専門家以外にも読みやすくなる
 - 全文検索が可能になる
 - 自然言語処理技術で解析が可能になる
 - AIを訓練する教師データにも



「くずし字」とは

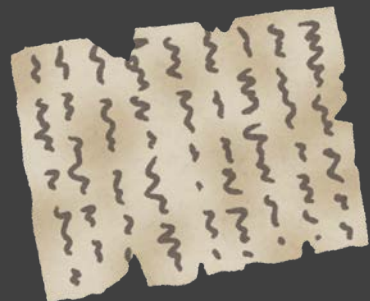
- 江戸時代以前に使われた、字画が省略された書体で書かれた文字の総称
 - 文章の書き易さに重点が置かれている
- 現在は使われない文字や書体で書かれているため、現代人には**とても読みにくい！**



「みんなで翻刻」がめざすもの

江戸時代以前の地震・災害史料
(東京大学地震研究所所蔵)

テキストデータ
(CC BY-SAで公開)



翻刻



活用

活用



地震・防災研究



オンラインで参加するユーザー



機械学習ベースの
文字認識研究

これまでの作業進捗（2017年1月～）

祝！ 東京大学地震研究所の所蔵史料約500点が
「**翻刻完了**」に！

- 翻刻が完了した史料 500/508点
- 翻刻が完了した画像 8,411/8,925枚 (94%)
- 参加登録者数 5,346人
- 総入力文字数 625万字
- 翻刻の正確性 98.5% (10万字をサンプリング検証)

「みんなで翻刻」の特徴

くずし字解読を勉強しながら参加できる！

1. くずし字学習支援アプリ「KuLA」と完全連携
2. 参加者同士で教え合う「学び合い」システム

くずし字学習支援アプリ 「KuLA」

- 大阪大学で開発された
くずし字の学習支援アプリ
- 特徴
 - 3000枚の用例画像を収録
 - 方丈記や刀剣書で読解訓練
- ダウンロード
 - App Store or Google Playで
「くずし字」と検索！



翻刻文の相互添削サービス

共通タイムライン

翻刻



フィード
バック



共有



添削結果の差分表示



添削

みんなで翻刻の新マスコット 「そらまる」

どこか遠い空間からきた女の子と、
女の子が乗っていた乗りもの。
記憶を失ってどこから来たのか
わからなくなっていました。



《女の子》

赤い髪がとっても
チャーミング！
本に興味津々。

《そらまる》

こう見えて
意外と高性能。
古文書の文字を
認識する機能を
開発中。

自分たちによく似た姿の
ものが古典籍に描かれていると
聞いて、くずし字が読めるように
なりたいと思っている。



みんなで翻刻の新マスコット 「なまぞう」



地震を起こしているという
疑惑を晴らすために
地震研究に協力している
なまぞう（推定1000歳）。
地震を鎮めるという要石を
自らかぶっている。

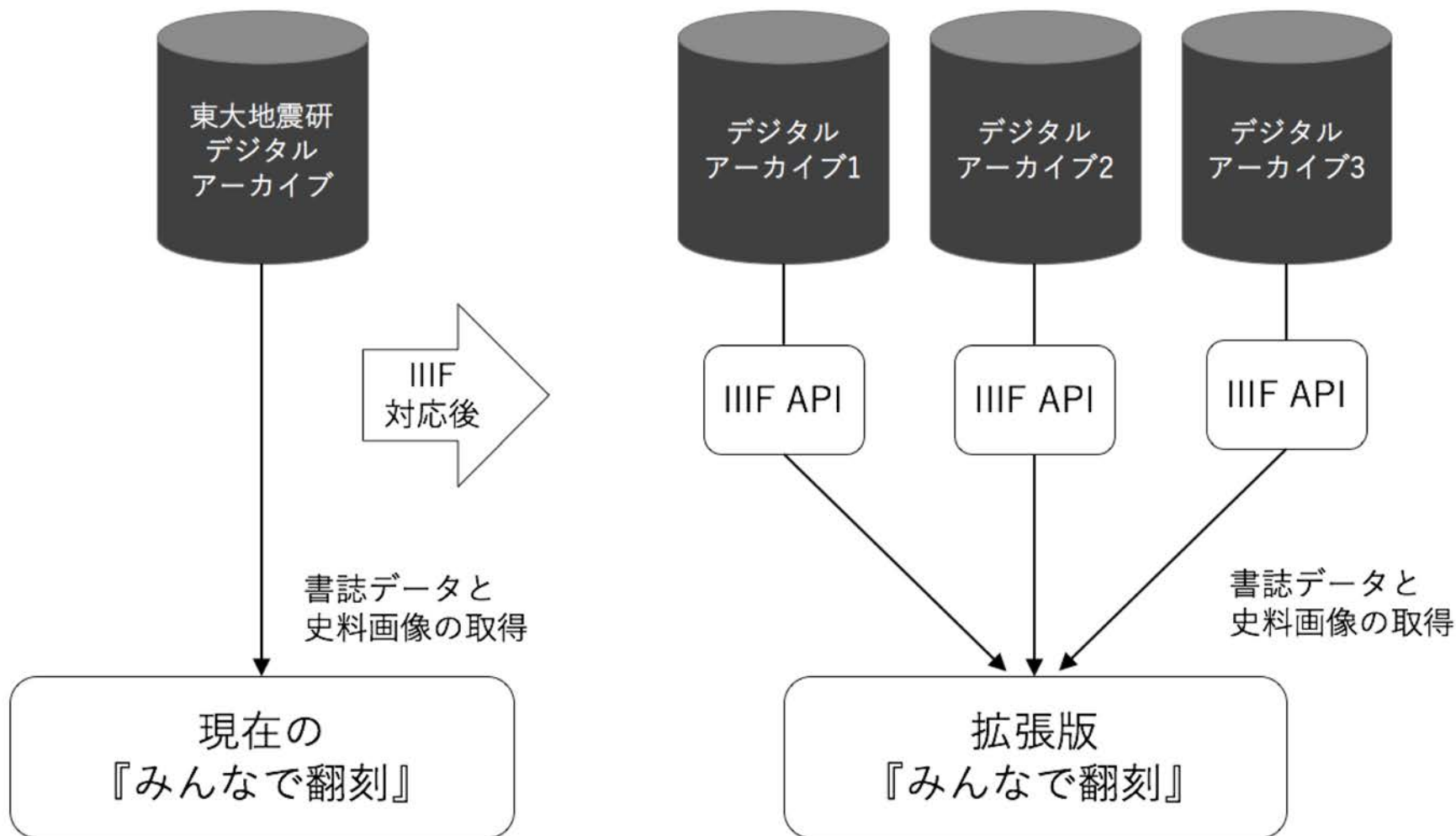
「みんなで翻刻」新バージョン

新バージョンの特徴：

どんな史料も「みんなで翻刻」できるように！

AIが「くずし字」を読んでもくれる！

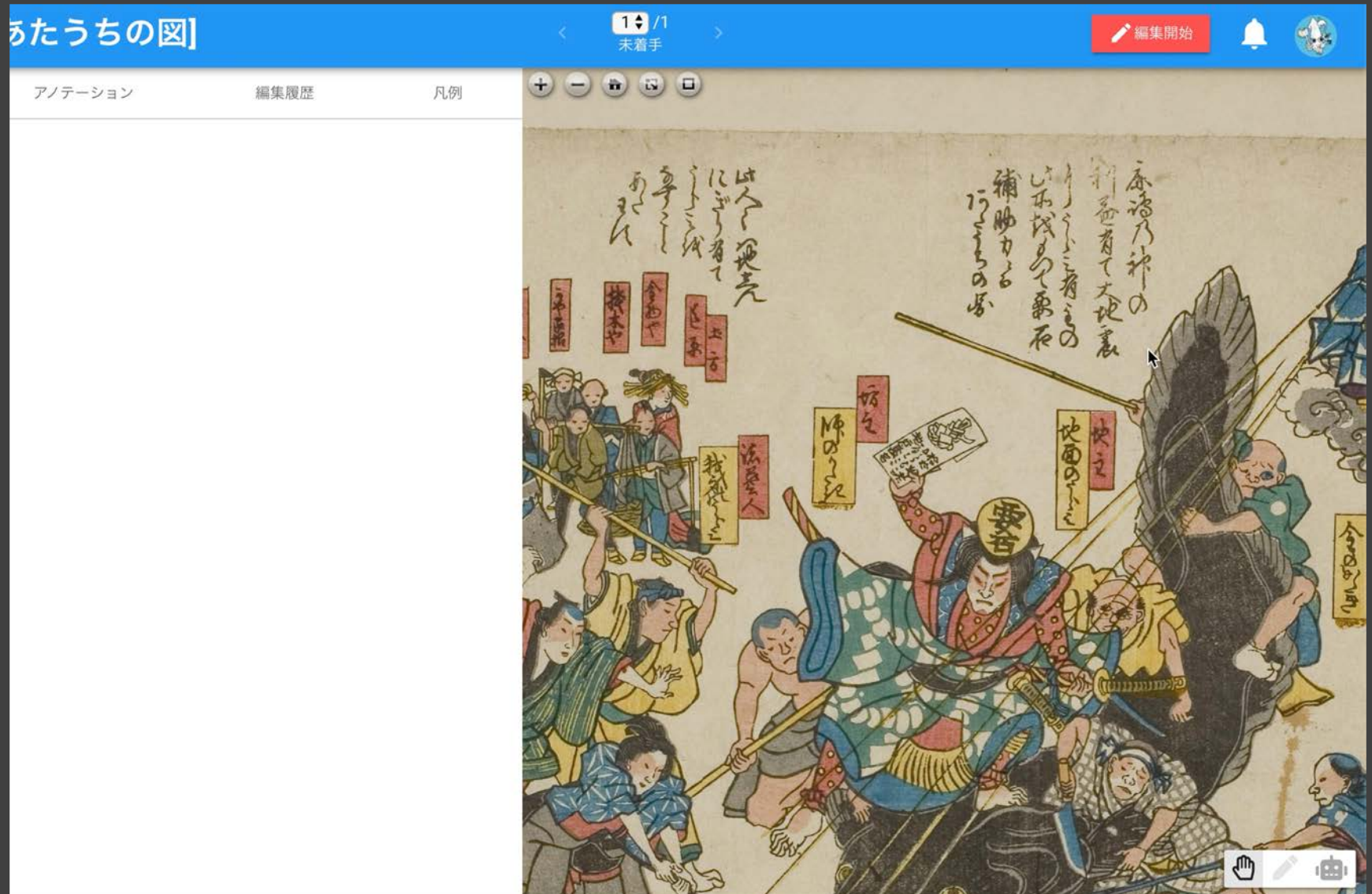
デジタルアーカイブの 標準規格「IIIF」に対応！



MLA諸機関との連携

- 京都府 京都学・歴彩館
- 福井県公文書館・福井県立図書館
- 茨城大学図書館
- 東京大学総合図書館
- 東京大学地震研究所図書室
- 京都大学図書館
- 関西大学 KU-ORCAS
- 国立歴史民俗博物館
- まだまだ増加予定！

くずし字認識 AI を搭載！



新バージョンの成果（2019年7月～）

翻刻が完了した史料	833/1,384点
翻刻が完了した画像	8,872/37,647
参加登録者数	1173人（60人up!）
総入力文字数	514万字（100万字up!）

現在進行中の翻刻プロジェクト

[ホーム](#)[プロジェクト](#)[まなぶ](#)[ご案内](#)[使い方](#)[A](#) [☆](#)

日本の仏典を翻刻

SAT大蔵経テキストデータベース研究会による、日本の仏典の翻刻プロジェクト
所蔵機関：SAT大蔵経テキストデータベース研究会等

32/96点の翻刻が完了



Code4Lib JAPAN × みんなで翻刻

Code4Lib JAPAN 2020カンファレンスで登録した資料です。
所蔵機関：国立国会図書館など

14/37点の翻刻が完了



デジタルアーカイブ福井の資料を翻刻

デジタルアーカイブ福井で公開している地域資料を翻刻するプロジェクトです。
所蔵機関：福井県文書館

12/24点の翻刻が完了



疫病関連資料を翻刻！

近世の疫病に関連する資料を翻刻するプロジェクトです。
所蔵機関：京都大学、国立歴史民俗博物館、他、各地のデジタルアーカイブ

14/34点の翻刻が完了



翻刻！東寺百合文書

京都学・歴史館が所蔵するユネスコ世界記憶遺産『東寺百合文書』の翻刻プロジェクトです。
所蔵機関：京都府京都学・歴史館

71/210点の翻刻が完了



茨城大学図書館所蔵資料を翻刻

茨城大学図書館が所蔵する近世資料の翻刻プロジェクト
所蔵機関：茨城大学図書館

10/15点の翻刻が完了



翻刻！石本コレクション

地震研究所の第2代所長である石本巳四雄（1893-1940）が収集した、主に災害の際に出版されたかわら版や錦絵、新聞付録などの資料群です。
所蔵機関：東京大学総合図書館

404/578点の翻刻が完了



翻刻deかるたLOD

古典籍に含まれる「かるた」を翻刻するプロジェクトです。翻刻したデータをもとにLOD（Linked Open Data）を作成し、公開します。
所蔵機関：国立国会図書館、国文学研究資料館、京都大学ほか

101/105点の翻刻が完了



関西大学の多彩な東アジア研究資料を翻刻！

関西大学アジア・オープン・リサーチセンター（KU-ORCAS）では、関西大学が誇る豊富で多彩な東アジア関係資料を“オープン”に使えるように、さまざまなデジタルアーカイブのプロジェクトに取り組んでいます。その中から、戦前の東アジアの映画関係資料と江戸時代の中国語関係資料を提供しております。



翻刻！草双紙の世界

国会図書館デジタルコレクションなど、各地のアーカイブで公開されている草双紙を読んでみましょう。
所蔵機関：国会図書館デジタルコレクションなど

48/73点の翻刻が完了

翻刻をもとにした電子書籍が刊行！

- 『化物七段目』（珍獣の館文庫）
 - 翻加者の @chinjuh（蘭藍沐）さんが刊行
 - 現代語訳と着色済みの挿絵も収録！



ニコニコ生放送 『みんなで翻刻してみた』

ドワンゴとの共同企画. 2017年3月から月1回放映



第1回放送はリアルタイムで1.2万人が視聴

「ニコニコ超会議」への出展



2017年

2018年

まとめ：みんなで翻刻とは

- MLA機関が所蔵する歴史資料の大規模テキスト化をめざす市民参加型プラットフォーム
- くずし字解読の教育サービスとしてデザイン
 1. 学習アプリとの連携
 2. ソーシャル添削システム
 3. AI による参加者サポート
- 3年半で6,000人超の参加により1,300点（1,100万文字）の翻刻を実現