

歴史地震研究における異分野連携 とシチズンサイエンス

INTERDISCIPLINARY COLLABORATION AND
CITIZEN SCIENCE IN HISTORICAL SEISMOLOGY

国立歴史民俗博物館
橋本雄太

Yuta Hashimoto
National Museum of Japanese History

Japan Open Science Summit 2018
D1 人文学研究のデジタル化とオープン化

はじめに

- 自己紹介
 - 国立歴史民俗博物館 助教
 - 専門：人文情報学， 歴史情報学
- 本発表では
 - 人文学領域における**シチズンサイエンス**の事例として， 京都大学古地震研究会が運営する『みんなで翻刻』プロジェクトを紹介

京都大学古地震研究会



- 京都大学所属の地震学者が主宰する，歴史地震史料の研究グループ
- 地震学，気象学，日本史，地理学，科学史，情報学，図書館司書など多分野の参加
- 週一回の活動で，毎年10冊程度の史料を解読



解読対象の例（『元禄地震記』）



くずし字で筆記されており、訓練を受けない限り現代人には解読困難

地震史料の解読が重要な理由

- 日本で計器を利用した地震の観測が始まったのは、明治20年代のこと
- それ以前に発生した地震については、基本的に文献史料を使用して発生状況を調べる必要がある
- 地質学的な調査では百年単位での誤差が生じる
- しかし地震史料の数は膨大
- AIによる自動解読もまだ道半ば



みんなで翻刻

【地震史料】

古文書・古記録を解読して、
歴史災害研究に参加しよう！

参加する

現在はPCでのアクセスのみ対応しています。

いいね！ ツイート G+ B! ブックマーク LINEで見る

京都大学
KYOTO UNIVERSITY

古地震研究会

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

「みんなで翻刻 地震史料」

- 概要

- 歴史災害史料の市民参加型翻刻プロジェクト
- 2017年1月10日に Web サイトを正式公開
- 翻刻…歴史学の専門用語で、古文書に書かれた文字を活字に起こすこと
- 翻刻されたテキストは CC BY-SA で公開

- プロジェクトの目的

1. 災害史料の大規模テキストデータベース構築
2. 翻刻作業を通じた市民の防災意識向上
3. 機械学習の教師データ構築

翻刻の対象

- 東大地震研図書室「石本文庫」
 - 東京大学地震研2代目所長・石本巳四雄（1893-1940）が収集した災害史料コレクション
 - このうち、現在地震研でデジタル公開されている史料114点（画像数3,193枚）の全文翻刻が当初の目標
→公開5ヶ月で完了。現在は地震研の他の史料を翻刻中



『信州浅間山焼』



『地震年代記』



『元禄地震記』

市民参加を実現するために

- 古文書解読の学習サービスとして設計されている
 1. くずし字解読の学習コンテンツ提供
 2. 参加者間のコラボレーションによる「学び合い」
- くずし字学習の延長として、自然に翻刻作業に参加してもらうことが目的
 - 「やりがい搾取」にならないように
 - 参加者が明確なメリットを得る形で参加してもらう

用例画像を利用した確認テスト

テスト：漢字 その2

■の部分は何て読むかな？1文字で答えてね



去 ■

回答を入力

ひ 来 へ

閉じる 回答

【用例】



【無】

テストを

史料一覧画面

みんなで翻刻【地震史料】

はじめに

- ホーム
- ごあいさつ
- ブログ・SNS・ML

地震史料を翻刻する

- 翻刻ガイドライン
- 史料一覧 (STAGE1)
- 史料一覧 (STAGE0)
- 史料一覧 (STAGE3)
- 史料一覧 (STAGE2)
- 史料一覧 (STAGE2)
- 史料一覧 (STAGE1)
- お気に入り・作業履歴
- フォーラム

くずし字解説を学ぶ

- まなぶ
- つながる


ライセンス/クレジット

- お問い合わせ

キーワードで検索


翻刻完了した史料も表示する

安政二卯十月二日大地震附類焼...




1/13コマの翻刻が完了
東京大学地震研究所図書室所蔵
[完成江戸地図](#) [見る](#)

附録 地震劇風災害豫防法圖説...




4/30コマの翻刻が完了
東京大学地震研究所図書室所蔵
(1867)
[完成江戸地図](#) [見る](#)

天學略名目羽翼 全




1/1106コマの翻刻が完了
東京大学地震研究所図書室所蔵

大旋風




1/13コマの翻刻が完了
東京大学地震研究所図書室所蔵
(1862)
[完成江戸地図](#) [見る](#)

[首引をする地震蛇と鹿島大明神]



1/13コマの翻刻が完了
東京大学地震研究所図書室所蔵
[完成江戸地図](#) [見る](#)

[伊勢の馬を引く鹿島大明神]



1/13コマの翻刻が完了
東京大学地震研究所図書室所蔵
[完成江戸地図](#) [見る](#)

翻刻文エディター

地震年代記 ★ 9 /32コマ 翻刻完了

編集開始 保護

閲覧 入力 ノート 編集履歴 凡例

如し夜分も地震して声あり○同二年六月十五日地震大に震ふ以上三代実録
 ○同三年七月朔日二日夜地震六日夜も地震同月晦日申の時地震大に
 震動し数刻を経てやまず天皇は仁寿殿を出させたまひ紫宸
 殿にみなみのおには

北國年令言
 西の堅壇長八留刻表宮城京師の戸舎類損する事
 察奏園しそいそ地震の徴は兵賊飢痕と慎む
 其夜戌刻より子刻迄地震又震動し八日は辰刻より
 丑刻に至る迄に四度より九日の夜も二度十日は五度十一日はしばしば震動し
 十二日は子一刻大にゆり寅四刻小くゆる十三日十四日十七日十八日もゆる
 十九日戌の時天に声ある事二度地震又震動す廿一日も戌一刻空に声
 ありて丑刻に地震廿二日辰の時大に震二たび動きてやむ廿三日廿四日
 廿五日廿九日も地震する同五年正月六日酉の時地震する十一日十二日
 十四日十六日も九月十九日廿一日地震する以上三代実録
 光孝天皇 仁和
 元年四月十四日大風雨して地震声あり同十六日霜ふり曇て地震す
 扶桑略記同十二月廿日巳の時天の東南に声ありて高樓の墮落が

元北の堅壇長八留刻表宮城京師の戸舎類損する事甚多し七日陰陽
 寮奏聞していはく地震の徴は兵賊飢疫を慎むべしと
 其夜戌刻より子刻迄地震又震動し八日は辰刻より
 丑刻に至る迄に四度より九日の夜も二度十日は五度十一日はしばしば震動し
 十二日は子一刻大にゆり寅四刻小くゆる十三日十四日十七日十八日もゆる
 十九日戌の時天に声ある事二度地震又震動す廿一日も戌一刻空に声
 ありて丑刻に地震廿二日辰の時大に震二たび動きてやむ廿三日廿四日
 廿五日廿九日も地震する同五年正月六日酉の時地震する十一日十二日
 十四日十六日も九月十九日廿一日地震する以上三代実録
 光孝天皇 仁和
 元年四月十四日大風雨して地震声あり同十六日霜ふり曇て地震す
 扶桑略記同十二月廿日巳の時天の東南に声ありて高樓の墮落が


元北の堅壇長八留刻表宮城京師の戸舎類損する事甚多し七日陰陽
 寮奏聞していはく地震の徴は兵賊飢疫を慎むべしと
 其夜戌刻より子刻迄地震又震動し八日は辰刻より
 丑刻に至る迄に四度より九日の夜も二度十日は五度十一日はしばしば震動し
 十二日は子一刻大にゆり寅四刻小くゆる十三日十四日十七日十八日もゆる
 十九日戌の時天に声ある事二度地震又震動す廿一日も戌一刻空に声
 ありて丑刻に地震廿二日辰の時大に震二たび動きてやむ廿三日廿四日
 廿五日廿九日も地震する同五年正月六日酉の時地震する十一日十二日
 十四日十六日も九月十九日廿一日地震する以上三代実録
 光孝天皇 仁和
 元年四月十四日大風雨して地震声あり同十六日霜ふり曇て地震す
 扶桑略記同十二月廿日巳の時天の東南に声ありて高樓の墮落が

参加者間の翻刻文添削

- 翻刻文の保存時に、
翻刻文の添削を他の参加者に
依頼することができる
- 相互添削の目的
 - 参加者の能力向上
 - 翻刻文の品質向上

編集を終了

この内容で保存していいかな？
すべての文字が翻刻できていたら「翻刻完了」にチェックをつけてね。読めない箇所が残っていても大丈夫！
編集コメントを記入しておく、後の人の助けになるかも！



添削おねがいします

翻刻完了

タイムラインで共有

添削希望（他の参加者に翻刻文の添削をお願いします）

キャンセル 保存

翻刻作業のタイムライン共有

 **みんなで翻刻【地震史料】**

はじめに

- 🏠 ホーム
- 📄 ごあいさつ
- 👍 ブログ・SNS・ML

地震史料を翻刻する

- 📖 翻刻ガイドライン
- 🔍 史料一覧
- ★ お気に入り・作業履歴
- 💬 フォーラム

くずし字解読を学ぶ

- 🎓 まなぶ
- 🔗 つながる

👥 ライセンス/クレジット

📧 お問い合わせ

 田原輝男さんが318文字翻刻しました 追加希望
1時間前

 **浅間大焼騒動記 中一コマ26**
も動ればて夫より種月然へ押寄大勢ふミ入れれ共人巻人も見へず唯本堂計ぞおわしける然共委...

👁 あっばれ!

 ramayakatさんが424文字翻刻しました 追加希望
2時間前

 **江戸大地震井ニ出火場所細見録一コマ3**
一新吉原江戸町京町角町揚屋町伏見丁仲の丁すべてくるはのこらずゆりつぶれ(?:てみへ出...

👁 あっばれ!

 小林じゃさんが679文字翻刻しました 追加希望
2時間前

翻刻文添削時の通知









現在の進捗 (日次更新)

総入力文字数	3061194文字 (+12760文字)
翻刻完了画像	4551/4636枚 (98%, +13枚)
翻刻完了史料	263/280点 (+0点)
参加登録者数	3428人

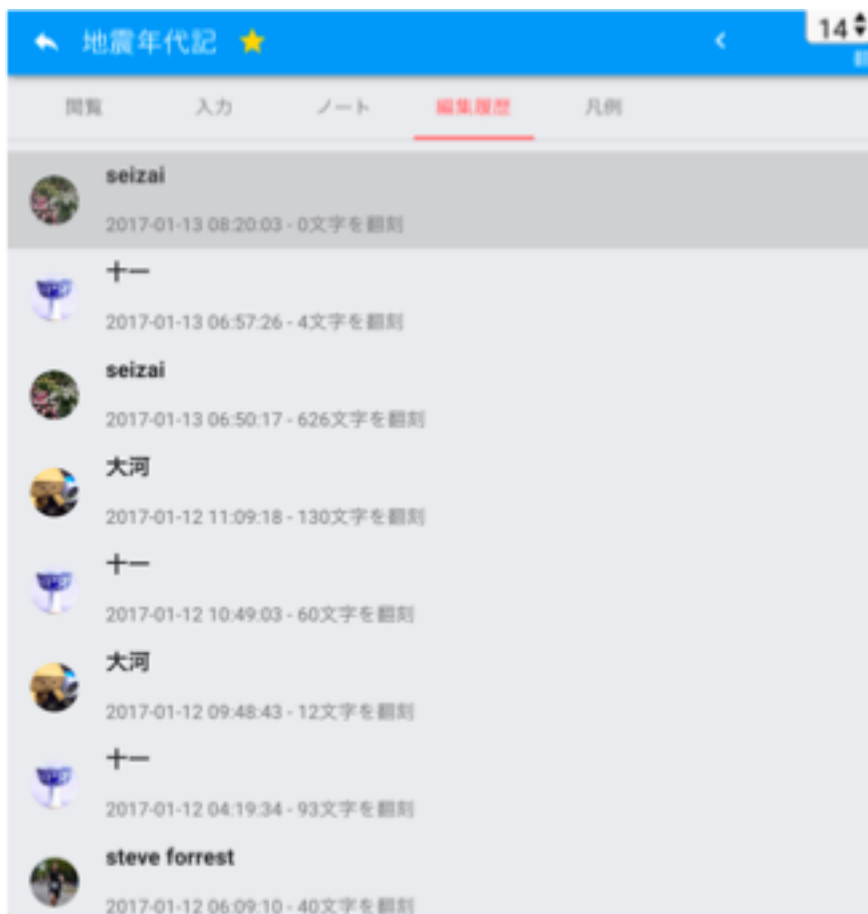
翻刻文字数ランキング

-  Takano Kiyoshi
343211文字
-  田原輝男
322530文字



-  Hideo KIHARAさんも【善光寺地震被害...・コマ7を翻刻しました
14時間前
-  たかたに やすひこさんがあなたの翻刻に「あっぱれ！」と言っています
14時間前
-  レベルが19に上がりました
14時間前
-  レベルが19に上がりました
14時間前
-  小林じゃさんも伊豆国修善寺温泉...・コマ2を翻刻しました
17時間前

添削結果の表示



他の参加者から添削を受けた際の修正箇所が一目で確認できる
 緑色 = 追記箇所, 赤色 = 削除箇所

公開後の成果

- 2017年1月のシステム公開後，1日最大4.5万文字が入力
- 公開5ヶ月後には石本文庫114点の全文翻刻が完了
→ 地震研の所蔵するデジタル化済み和古書全点を投入



プロジェクトの進捗（6/18時点）

- 総入力文字数…501万文字
- 翻刻が完了した画像…6588/6966枚（94%）
- 翻刻が完了した史料…452/472点
 - 残り20点で東大地震研所蔵（デジタル化済み）和古書全点の翻刻が完了
- 参加登録者数…4301名
 - このうち330名が実際の翻刻に参加

成果物の品質について

- 中世史研究者に依頼し、翻刻文10万文字を検証した結果：

表 3.2 検証作業を通じて見つかった要修正箇所の数（資料種別・タイプ別）

資料種別	合計文字数	誤読箇所		未読箇所		表記ゆれ箇所		合計	
		実数	比率 (%)	実数	比率 (%)	実数	比率 (%)	実数	比率 (%)
木版本	78,774	400	0.5	285	0.4	395	0.5	1,080	1.3
筆写本	21,901	210	1.0	105	0.5	132	0.6	447	2.0
合計	100,675	610	0.6	390	0.4	527	0.5	1,527	1.5

- 100文字あたり平均1.5文字のエラー
- 学術出版される史料集には及ばないものの、
内容把握や全文検索には十分な品質

まとめ

- 『みんなで翻刻』の成果
 - 人文学の基礎的作業である歴史資料の翻刻に、市民参加のプロセスを導入することができた
 - 地震史料という特殊ジャンルではあるが…
- 今後の展望
 - 災害史料以外の一般の歴史資料への展開
 - 翻刻テキストの地震研究への活用
 - 機械学習技術の導入