

IIIFを介して グローバルとローカルを繋ぐ： 『みんなで翻刻』の今後の展開

2021年1月15日 NDLデジカフェ
橋本雄太（国立歴史民俗博物館）

自己紹介

- 橋本雄太
- 身分：国立歴史民俗博物館 助教
- 専門：人文情報学， 歴史情報学
- これまでに公開したシステム
 - みんなで翻刻
 - 日本の中世文書WEB
 - くずし字学習支援アプリKuLA
 - 軽量マークアップ言語Koji
 - Etc..
- Twitter: @yuta1984



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。

みんなで翻刻

- 概要

- 歴史資料の**市民参加型翻刻**プラットフォーム
- 翻刻・・・歴史学用語で史料を活字に起こすこと
- 京都大学古地震研究会を中心に開発・運営
- 2019年のリニューアルで IIIF に対応

- 翻刻の意義

- 歴史資料が**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
- 機械学習の教師データにも



閲覧

入力

ノート

編集履歴

天明二壬寅の今年以前三十餘年の

天明二壬寅の今年以前三十餘年の
あゝ自由を思ふ事とありて是れ
徳川の世と云ふ事一と今年七月
四原村火災の初書信事書
郷土情失せり此時信向るも一
原とあり也
去年安永十辛丑年我生年今之
川とて少藤那とて山村也是も
日道少とて山とて東陽回跡并
也



これまでの成果（2019年7月～）

運用期間	543日
翻刻が完了した史料	920/1,854点
翻刻が完了した画像	12,659/104,432枚
参加登録者数	1,315人
総入力文字数	699.8万字

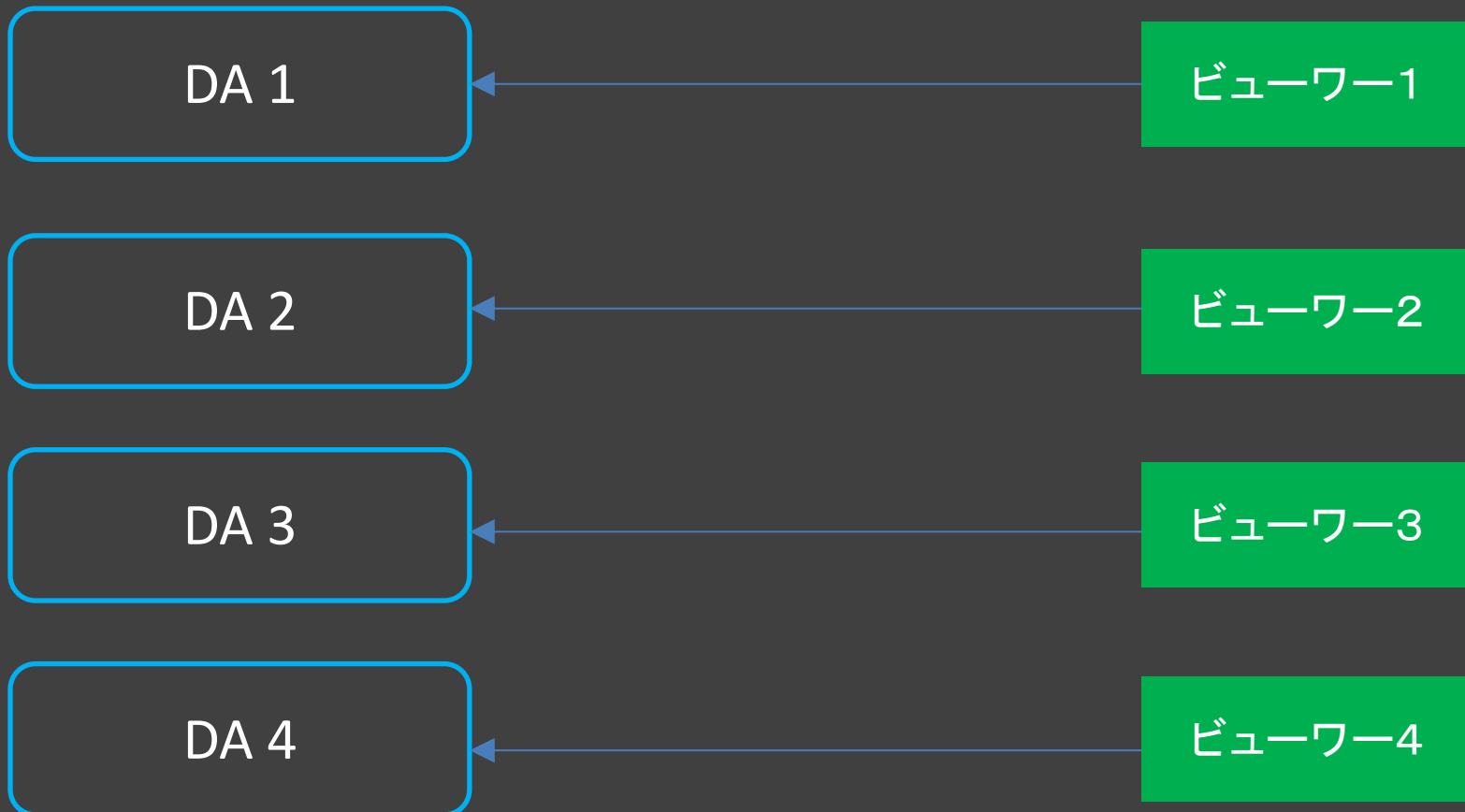
毎日1.3万字のペースで進行中！

IIF (International Image Interoperability Framework)

- 画像データへのアクセスを標準化し，DA間で相互運用性を確保するための仕組み
- 対応機関
 - NDL，国学研，日文研，歴博，…
 - 北海道大学，東京大学，京都大学，関西大学，九州大学，佐賀大学，琉球大学，…
 - 一部公共図書館や文書館も

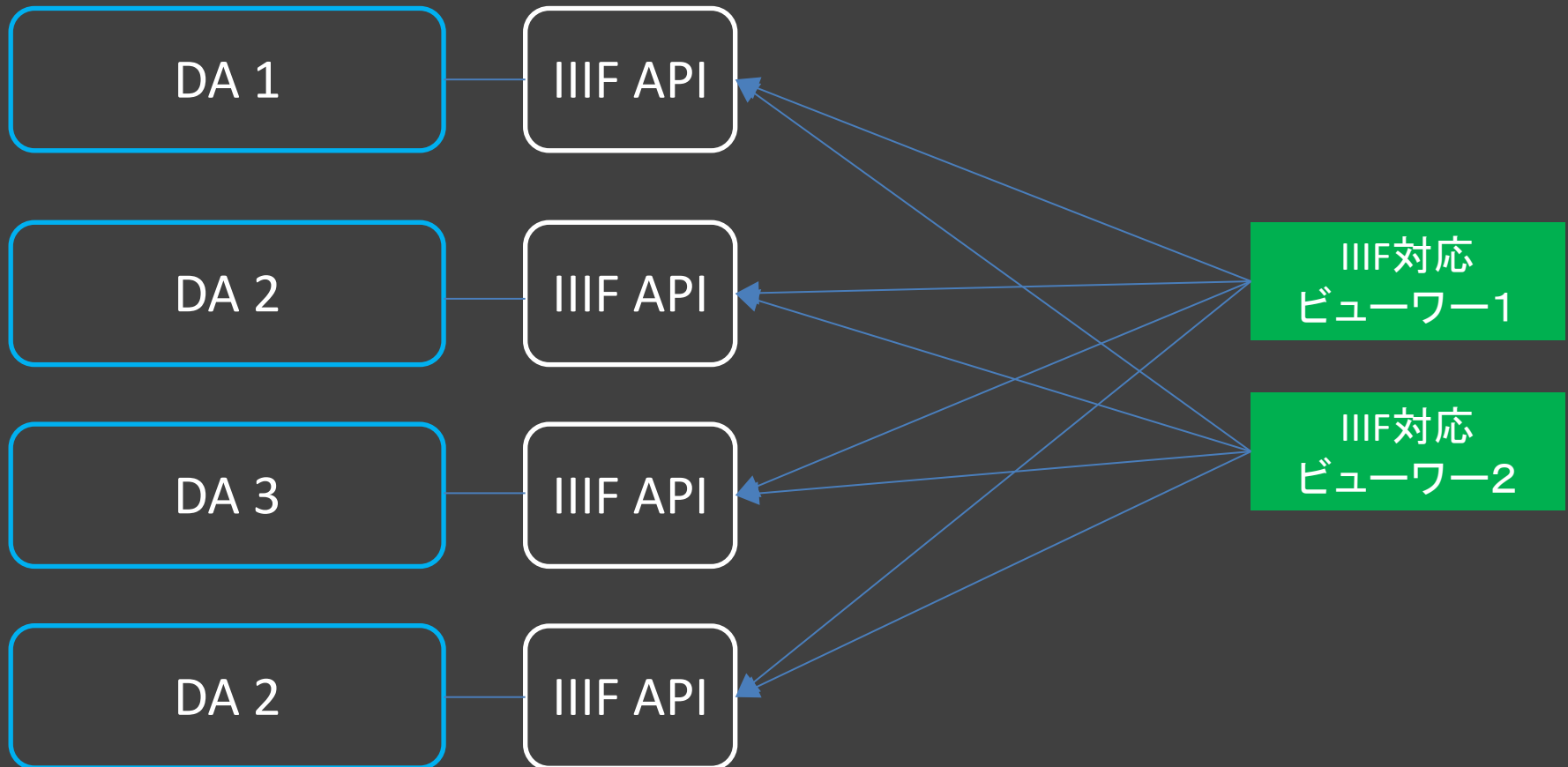
IIIFの相互運用性

これまでのデジタルアーカイブ



IIIFの相互運用性

IIIF 対応後のデジタルアーカイブ



プロジェクト一覧



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

41/96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

23/37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

19/36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

18/34点の翻刻が完了



Gallicaの日本資料を翻刻！

フランス国立図書館（Bibliothèque nationale de France, BnF）のデジタルライブラリーGallicaで公開されている日本資料の翻刻プロジェクトです。
 所蔵機関：フランス国立図書館（BnF）

19/296点の翻刻が完了



翻刻！東寺百合文書

京都学・歴彩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・歴彩館

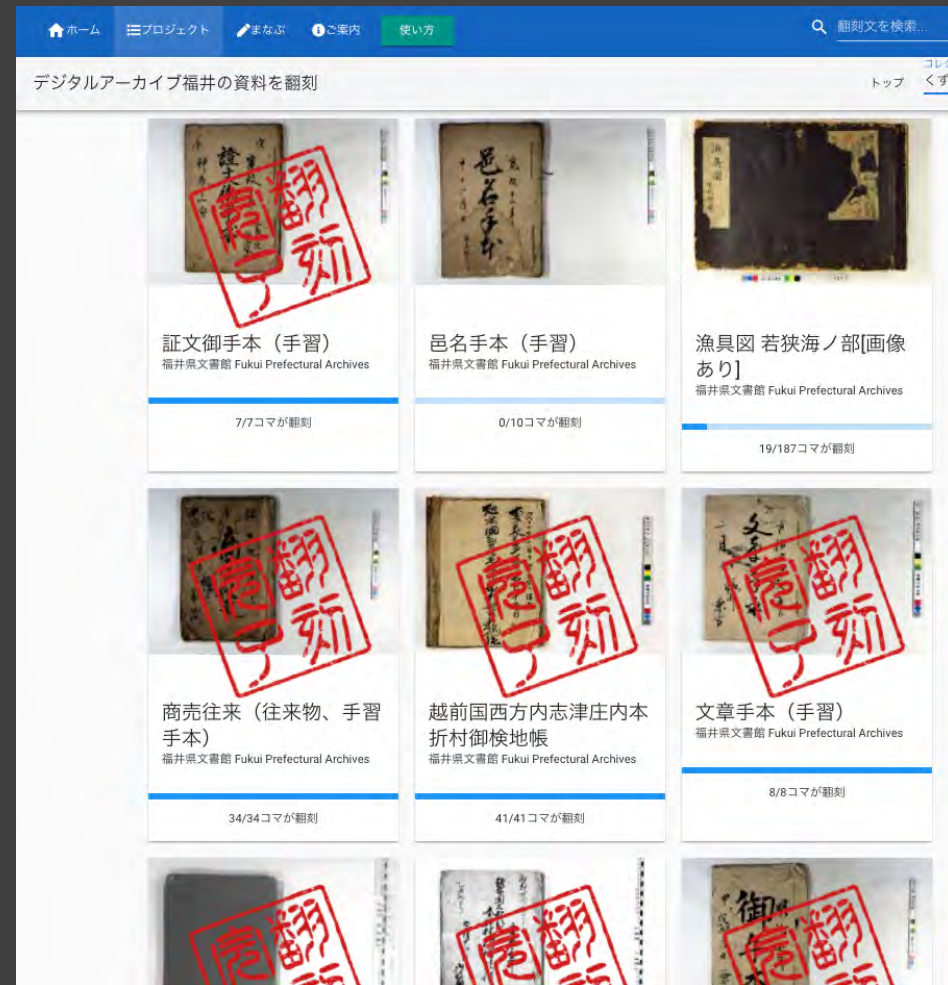
71/210点の翻刻が完了

IIIFを介して ローカル・グローバルと繋がる

- DAにおける「ローカル」
 - 地域の図書館や博物館，文書館
 - 地域で継承されてきた資料
 - 地域に住む人々
- DAにおける「グローバル」
 - 海外の機関が公開する資料
 - 海外の日本研究者，専門司書，愛好家

デジタルアーカイブ福井 翻刻プロジェクト

- デジタルアーカイブ福井
 - 福井県文書館，福井県立図書館らが運営するDA
 - 明治期の新聞画像や越前松平家資料群などを公開
- みんなで翻刻との関わり
 - システムがIIFに対応しているとTwitter上で判明
 - 県立文書館副館長の協力で翻刻プロジェクトが実現



“Tackling Pandemics in Early Modern Japan”

- ケンブリッジ大学 和本サマースクール
 - 日本文学研究者のL. モレッティ博士が毎年開催
 - 今年はCovid-19により開催できず
- Tackling Pandemics
 - みんなで翻刻のシステムを利用した完全オンラインの代替プロジェクト
 - 近世の疫病関係資料約20点を翻刻
 - 英国, 米国, ロシア, 中国, アルゼンチンなど10カ国以上から37名が参加

The screenshot shows a web application interface with a blue header containing navigation links: HOME, PROJECTS, LEARN, ABOUT, USER'S MANUAL, and a search bar. The main content area displays three user-submitted transcriptions of Japanese text from a historical document, each with a 'Review me' button and a 'Like' button.

User 1: Nancy Hamilton (23日前)
Thank you Joseph! I made the recommended changes, but left your comments in where it seemed like they'd be helpful. I agree it looks like there is a し at the end of かない but I wonder why. At first I thought it could be related to 汚 to the right but that doesn't make sense either. hmm.. Hope you are well, Nancy

Text: 麻疹必用: 一名痘疹年代記 - コマ18
麻疹/近隣(ちかく)へ来るといふ ならばよき香(かう)きやら)を焚(たき)家 并二疹(はしか)せぬ輩(|ともが...

User 2: Joseph Bills (24日前)
I'm afraid to say I have no idea either about the last characters on the first line, but the rest is complete :)

Text: 麻疹必用: 一名痘疹年代記 - コマ36
痘瘡中食して好(よき)物(もの) 一大(たい)根(こん) 一千(ほし) 大根一かんひやう一独活(うど) 一牛(こ...

User 3: Joseph Bills (24日前)
This one's done now - well done Anna!

Text: 麻疹必用: 一名痘疹年代記 - コマ33
一人をしかり嗔(けん) 嘩(くわ) 口(こう) 論(ろん) 等痘者に聞(きか) しむる事なかれ 一痘者に向(むか) ひ髪(かみ) ...

Gallica日本資料翻刻プロジェクト

- Gallica
 - フランス国立図書館の運営する電子図書館
 - 画像資料はIIFに対応
- Gallica における日本資料
 - 近世以前の日本語資料約300点が公開
 - 翻刻することで Gallica にフィードバックできるかも



おわりに

- IIIFが可能にしたこと
 - 海外と国内の公開資料を統一的に扱うことが可能になった
 - デジタルアーカイブ世界の普遍言語
- 課題
 - 画像資料IIIF化のハードルはいまだ高い
 - 特に地方の小規模館には予算・技術面で困難
 - NDLがIIIF化を支援してくれたり…？