



国立歴史民俗博物館  
National Museum of Japanese History

歴博の新たなデジタルデータの紹介と  
日本におけるデジタルアーカイブの動向  
New Online Research Resources at Rekihaku and Trends of  
"Digital Archives" in Japan

National Institutes for Humanities  
National Museum of Japanese History  
Makoto Goto,

人文

歴博

# Outline

## 1. Recent Trend of “Digital Archive” and Digital Humanities in Japan

最近の日本におけるデジタルアーカイブとDHの傾向について

## 2. Introduce New Resource of nihu and REKIHAKU(National Museum of Japanese History)

人間文化研究機構と歴博の新データリソースの紹介



大学共同利用機関法人  
人間文化研究機構



国立歴史民俗博物館  
National Museum of Japanese History

# Recent Trend of “Digital Archive” and Digital Humanities in Japan

最近の日本におけるデジタルアーカイブと  
DHの傾向

人文

歴史

# リソース公開 (New Resource)

## 画像系(Image Data)

- IIFによる画像公開が大学を中心に一般化
- IIF image publication is becoming more common, especially in universities.
- Cultural Japanの登場

## テキスト系 (Text Data)

- TEIデータとして
  - 渋沢栄一記念財団 <https://shibusawa-dlab.github.io/app1/>
  - デジタル源氏物語 <https://kouigenjimonogatari.github.io/>
  - 延喜式も（後述） Engi-shiki (later mention)

# リソース公開 (New Resource)

## データセットの構築 (Data Set)

- JDCat：人文学・社会科学総合データカタログ (Japan Data Catalog for the Humanities and Social Sciences)
- nihuBridge (later)



# Cultural Japan

- 特に日本の画像情報について発見するためのサイト  
Especially for discovering about Japanese image information
- IIFをはじめとする、多様な画像セット
  - Various image sets, including IIF
- Japan Searchと必要に応じて使い分けを  
Use Japan Search and other search tools as needed.

# IIIFの充実

- IIIFの機関リスト

[https://github.com/2SC1815J/open-in-iiif-viewer/wiki/Test-Cases-\(in-Japan\)](https://github.com/2SC1815J/open-in-iiif-viewer/wiki/Test-Cases-(in-Japan))

- 目的は異なるが…
- 結果的にリストになっている



大学共同利用機関法人  
人間文化研究機構



国立歴史民俗博物館  
National Museum of Japanese History

# データセットとしてのJDCat

- JSPS「人社系データインフラ構築事業」により作成
  - As a component of JSPS's Program for Constructing Data Infrastructure for the Humanities and Social Science
- 人文学は東京大学史料編纂所のデータが入っている
- 今後、データの充実へ？  
To more data in the future?



# TEIデータの作成も進みつつある Progress in TEI data construction

- ただし、大元となるテキストデータがまだ不足しているのは事実
- However, there is still a lack of text data.



# TEI構築の試み

- 書籍刊行
- 一般財団法人人文情報学研究所（監修）  
石田友梨／大向一輝／小風綾乃／永崎研  
宣／宮川 創／渡邊要一郎（編）



## テキスト化の充実 Toward Enhancing Text Data

- NDLの重要な動き （個人送信も重要だが）  
Important NDL Movements
- 「次世代デジタルライブラリー」 Next Digital Library  
<https://lab.ndl.go.jp/dl/>
  - N-Gram ビューアも
- くずし字については、miwoの充実（凸版も）
- 「みんなで翻刻」も（later）

# 日本におけるDH研究

- テキスト解析全盛の時代へ
  - AIくずし字研究は落ち着きつつある（データとしては十分ではない）
  - データ基盤構築も引き続き議論されている
- 
- Text analysis is improving
  - AI Cursive Text Research is Settling Down (Not Enough Data)
  - Data infrastructure construction continues to be discussed



## 全体のトレンドとして As an overall trend

- 「人が見る」から「機械も見る」へ  
From Human Readable To Machine Readable (too)
- データベースからデータセットへ  
From Database To Data Sets
  - 機械がどれだけ使うことができるか、ということを意図した方向
- 研究方法の量の変化から質の変化へ  
From quantity change to quality change in research methods
- 一つ一つを見つけるという行為に加えて、「いかに大きく捉えてみることができるか」という流れ（ただし、これは資料を細かく見る人間の営為が前提）





人間文化研究機構と歴博の新リソース  
New resources of nihu and rekihaku

人  
文

歴  
博

nihu(National Institutes for Humanities)

人文

歷博

## 第4期における人間文化研究機構 nihuBridgeの全体像

機構・機関の情報・研究資源を発見するためのポータルページ  
将来的には、日本におけるDHの基本的なデータや手法を網羅的に発見できることを目指し、デジタルヒューマニティーズに関わる情報のワンストップを機構で実現



大学共同利用機関法人  
人間文化研究機構

リンクによる発見

機構内の  
既存情報資源

国際リンク集

研究者DB  
(必要に応じ見直し)

機構リポジトリ

人文系サイエンス  
マップ  
(評価から移動)

機構研究資源

機関データ  
機関のデータ一部受入  
文理や性質を問わないデータ  
受け入れ  
(機関の希望により)

基盤データ  
歴史地名・人名辞典など  
分野を横断した基盤データの  
構築と充実  
機関・機構を超えた研究に裨益

研究メソッド  
人文研究に基づく優れたデジ  
タル化手法に関する情報・  
ツールなど  
(例：解析ツール・多様なメ  
ディアデータ蓄積と共有手  
法)

一部移行

これまでも各機関で  
nihuINTの検索対象と  
なっていた  
研究資源

(ポータルの活用により直接  
機関の資源へのアクセスを増  
やし、機関へのリーチを高め  
る)



検索用データをシンプルな  
形式で統合検索システムへ  
(ジャパンサーチへもここから)

各機関にある大型情報基盤

基本的には機関の事業だが、機構全体  
や他機関の研究・事業とも必要に応じ  
て連携も。  
ポータルからアクセス可能にする。



khirin



新古典籍DB



言語資源・  
データセット



フォーラム型  
情報ミュージアム

外部連携  
(国内外の諸機関)

外部連携  
(国内外の諸機関)

国際的なDH研究  
・国際標準への準拠  
・国外のデータや手法との連携

国際展開

検索による発見

現行機能を縮小  
シンプルな検索機能

検索システム

情報事業全体を俯瞰

Japan Searchなど



国立歴史民俗博物館  
of Japanese History

## 大きな傾向を受けて In response to these trends

- nihuBridgeの構築
  - データセット展開サービスを実施
  - データベースを見通しよくするためのポータル
  - 機構の資源をここから全て見るができるように
- Construction of “nihuBridge”
  - Implementation of dataset deployment services
  - Constructing Portal
  - To be able to see all the resources of the nihu from this

歴博  
(National Museum of Japanese  
History)

人文

歴博

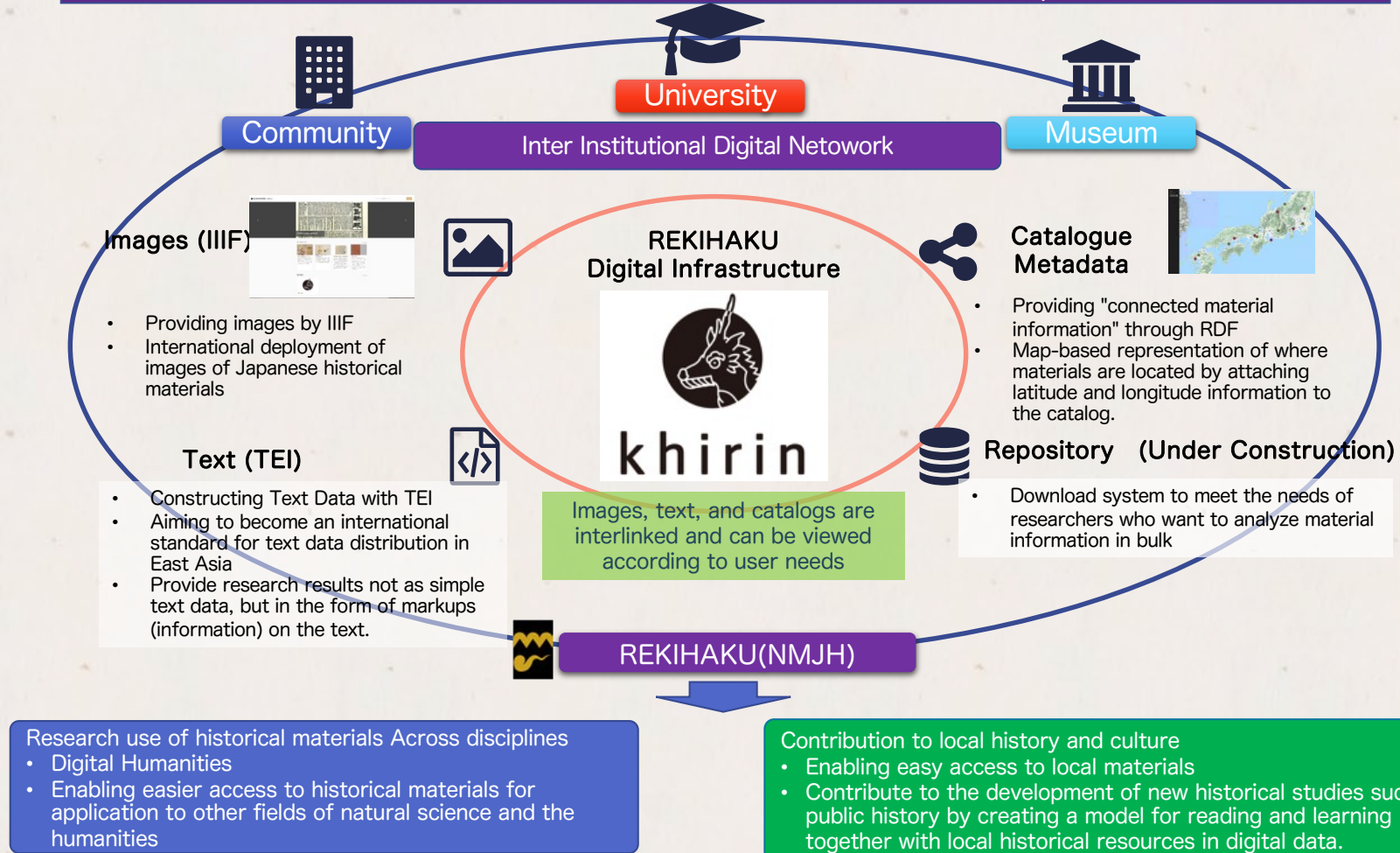


# khirinの充実

- khirin-tの完成
- khirin-a及びIdへの資料情報の追加
- khirin-cおよびkhirin-rとポータルの準備など
- Completion of khirin-t
- Addition of material information to khirin-a and Id
- Preparation of khirin-c and khirin-r and portal, etc.

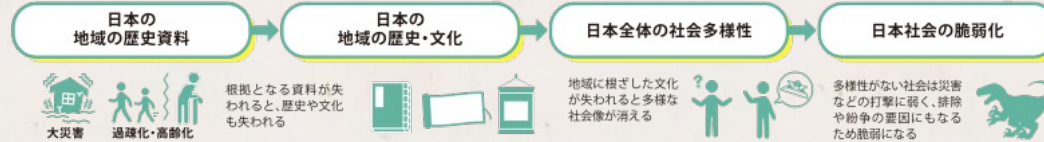


# Historical material data infrastructure "khirin" for universities, municipalities and museums

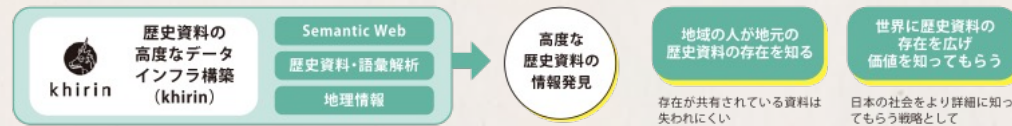


# データ基盤と総合的な知に基づいた日本歴史文化知の構築

## 01 社会課題: 地域歴史資料の消失による 地域社会の多様性の消失



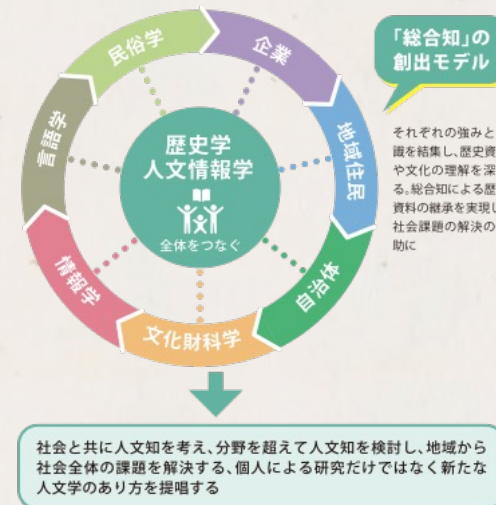
## 02 社会課題の解決提案 デジタル・ヒューマニティーズ の活用



## 03 社会課題の解決手段 歴史資料のデータネットワーク基盤"khirin"の 構築とデジタル・パブリック・ヒストリーの展開



## 04 社会課題の解決イメージ 理解を深め、多くのステーク ホルダーと歴史資料を理解する







khirin C

Knowledgebase of  
Historical Resources in Institutes

Refine Search

Institutions



Area



Resource Type



Enter keyword



☐ Search only materials with images

#### Introduction of database

The collection lists the historical documents that had been pass...

The archive of Kuki family in Sanda han

History of Fukusaki-cho

Items labeled as "archeology" from Rekimin cards

The collection lists the historical documents that had been pass...

Items labeled as "history" from Rekimin cards

common\_entity

graph\_info

A collection of individuals or objects.

agents

Collection of Goto clan's historical doc...

The archive of Kuki family in Sanda han

History of Fukusaki-cho

"Rekimin" cards · archaeology · for pub...

Collection of Machino clan's historical ...

"Rekimin" cards · history · for public u...

common\_entity

graph\_info

Class

agents

日本語



About khirin



# “History of Historical Materials”

- Not only where the material is currently
- But also where it has been in the past
- Khirin also considered how to describe such complex metadata.
- Meta-meta-data



# クラウドソーシングプログラム

- 「みんなで翻刻」
    - クラウドソーシングという手法→Wikipedia
    - Crowdsourced transcription platform for historical Japanese documents
    - 歴博・橋本雄太・京都大学古地震研究会による開発
    - Developed and run by Kyoto University and Rekihaku
    - 古文書を「読める」人がWeb上で古文書の翻刻を行っていく
    - わからないところは空白にしていると、他の読める人が読んでくれたりする→様々な工夫と教育効果も
    - 研究者以外の「くずし字好き」な人が多く参加
    - すでに数千万文字のデータが入れている
- Tens of millions of characters of data have already been put in.

## 「みんなで翻刻」の入力画面

大坂辺東海道本しらべ大地震津波図
13 / 1  
全部完了
編集開始

| 閲覧                                                                                                                                                                                                                                                                                                                       | 入力 | 編集履歴                                                                                                                                                                                                                                                                                                                                                                             | 凡例                                                                                                                                                                                                                                                                                                              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <div style="background-color: #007bff; color: white; padding: 2px; margin-bottom: 5px;">閲覧中</div> <div> <div>一 住吉とらうふ六部</div> <div>一 一なんば新地みぞのかは</div> <div>一 一龍屋町</div> <div>一 一福嶋五百らかん</div> <div>一 一京町堀三丁目</div> <div>一 一あわだ戸や町近辺</div> <div>一 一御池通五丁目四軒ばかりくすれ</div> <div>一 一北久太郎町井池北</div> <div>一 一京二三軒</div> </div> |    | <div> <div>一 一天王寺清水藤屋</div> <div>一 一堀町さのやはし</div> <div>一 一汐津ばし近辺</div> <div>一 一京町堀三丁目</div> <div>一 一あわだ戸や町近辺</div> <div>一 一御池通五丁目四軒ばかりくすれ</div> <div>一 一北久太郎町井池北</div> <div>一 一京二三軒</div> </div> <div style="margin-top: 10px;"> <div>一 一福嶋五百らかん</div> <div>一 一京町堀三丁目</div> <div>一 一あわだ戸や町近辺</div> <div>一 一御池通五丁目四軒ばかりくすれ</div> <div>一 一北久太郎町井池北</div> <div>一 一京二三軒</div> </div> | <div> <div>一 一京町堀三丁目</div> <div>一 一あわだ戸や町近辺</div> <div>一 一御池通五丁目四軒ばかりくすれ</div> <div>一 一北久太郎町井池北</div> <div>一 一京二三軒</div> </div> <div style="margin-top: 10px;"> <div>一 一福嶋五百らかん</div> <div>一 一京町堀三丁目</div> <div>一 一あわだ戸や町近辺</div> <div>一 一御池通五丁目四軒ばかりくすれ</div> <div>一 一北久太郎町井池北</div> <div>一 一京二三軒</div> </div> |

一 一福嶋五百らかん

一 一京町堀三丁目

一 一あわだ戸や町近辺

一 一御池通五丁目四軒ばかりくすれ

一 一北久太郎町井池北

一 一京二三軒

一 一福嶋五百らかん

一 一京町堀三丁目

一 一あわだ戸や町近辺

一 一御池通五丁目四軒ばかりくすれ

一 一北久太郎町井池北

一 一京二三軒

## 「みんなで翻刻」の特徴

- 研究者ではない人が多くのデータを蓄積する  
Non-researchers accumulate a lot of data
  - 古文書に親しむ機会
  - わからないところを他の人が読むというWeb上での「教育」
  - 精度は90%強→すぐに研究には使えない部分もあるが、あとは研究者がフォローすれば良い
- 検索で何かを見つけるには画像だと難しいがテキストだと楽
- Finding something in a search is difficult with images, but easy with text
- AIによるくずし字認識補助機能を持たせる
- Have AI assist in recognition of crumpled characters



道中嶋大合せん

1 / 1

着手済み

編集開始

通知

ユーザーアイコン

閲覧

入力

編集履歴

凡例

みちなかしま  
がっ  
 道中嶋大合せん  
だんげい  
 はいの死人芸

（このページには画像中央資料のみ、右の翻刻だけでよいと思われます）  
（画像左資料は同じページ中ほどにページがあるが、以下人力分はこのまま置いておきます）

まづの力ばなし  
 このる地震のうはさ遠国まできこ信州のなまづ  
 御当地へ下り地震の見舞は申さねども江戸のなまづと  
 こんいをむすびしばらくこにとりうしてありけるにあるひ  
 あまりとんななればなまづどじひげとひけとをむすび  
 あひそうはうの糸りにかけひげ首引をなさんこてたがひに  
 ちからをあらそひしが女房のあまづこれを見つけ  
 きまづをけしてはしりよりコレゝそんなわるいことをなぜ  
 しなさるのじやどちらにけがいあつてもわるいよしな  
 されととめるなゝイヤゝならぬイヤサなまづの  
 ひげのきれるぐらあはいといはせぬはゝそりやまた  
 なぜへイヤサ伊賀越の平作はなゝどうしたへ  
 沼津ではらさへ切ツたはへ  
 なまづのま夫やきもちばなし  
 はくひツ引をしけるにいつまで  
 しかゝでしなくしやうぶもつかざりしがとうゝ

（このページには画像中央資料のみ、右の翻刻だけでよいと思われます）  
（画像左資料は同じページ中ほどにページがあるが、以下人力分はこのまま置いておきます）

まづの力ばなし  
 このる地震のうはさ遠国まできこ信州のなまづ  
 御当地へ下り地震の見舞は申さねども江戸のなまづと  
 こんいをむすびしばらくこにとりうしてありけるにあるひ  
 あまりとんななればなまづどじひげとひけとをむすび  
 あひそうはうの糸りにかけひげ首引をなさんこてたがひに  
 ちからをあらそひしが女房のあまづこれを見つけ  
 きまづをけしてはしりよりコレゝそんなわるいことをなぜ  
 しなさるのじやどちらにけがいあつてもわるいよしな  
 されととめるなゝイヤゝならぬイヤサなまづの  
 ひげのきれるぐらあはいといはせぬはゝそりやまた  
 なぜへイヤサ伊賀越の平作はなゝどうしたへ  
 沼津ではらさへ切ツたはへ  
 なまづのま夫やきもちばなし  
 はくひツ引をしけるにいつまで  
 しかゝでしなくしやうぶもつかざりしがとうゝ

# 「みんなで古写真」の例

概要 人物 場所 時期 写真 参考資料 議論 作業履歴

1 「人物」「場所」タブなどに切り替えて、写真に情報を追加してみましょう！

伝記資料の当該箇所へジャンプ

|             |                                                                                                                                                                                                                                   |
|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 画像番号        | 別10-275                                                                                                                                                                                                                           |
| 画像URI       | <a href="https://eiichi.shibuyama.or.jp/denkishiryō-bekkan/spi/images/ul/1744/440_2352_1161_1710/max/0/default.jpg">https://eiichi.shibuyama.or.jp/denkishiryō-bekkan/spi/images/ul/1744/440_2352_1161_1710/max/0/default.jpg</a> |
| 「伝記資料」画像リンク | <a href="https://eiichi.shibuyama.or.jp/denkishiryō-bekkan/spi/images/ul/1744/full/max/0/default.jpg">https://eiichi.shibuyama.or.jp/denkishiryō-bekkan/spi/images/ul/1744/full/max/0/default.jpg</a>                             |
| 画像名         | 栄一胸像                                                                                                                                                                                                                              |
| 画像説明        | 中期                                                                                                                                                                                                                                |
| 掲載ページ       | 83                                                                                                                                                                                                                                |
| キャプション      | [P083] 大正十五年栄一の肖像除幕式が行われたが、席上栄一は往時を追想して初期の建物につき、「屋上は日本風、殊に古代建築を加味したらどうかと云うので、当時政府が独逸より聘したエンデー・ボックマン組合中の技術家を想々京都奈良方面に旅行させて、古社寺の建築法を研究せしめた」と語っている。大正十一年この建物は火災で焼失した。                                                                |
| 年月日         |                                                                                                                                                                                                                                   |



275 栄一胸像



大学共同利用機関法人  
人間文化研究機構



国立歴史民俗博物館  
National Museum of Japanese History



# 歴博の今後の方向性

- 地域歴史資料データの充実
- TEIデータをどのように作るか
- データセットの効果的な提供
- Digital Public Historyの展開
- Enhancement of Local Historical Resources Data
- How to create TEI data
- Provide effective datasets
- Developing Digital Public History



大学共同利用機関法人  
人間文化研究機構



国立歴史民俗博物館  
National Museum of Japanese History

人文

Thank you for your attention

應博