

人文学・社会科学データインフラストラクチャー強化事業フォーラム  
「データ共有・利活用促進のための研究基盤」

# 東京大学史料編纂所における取組

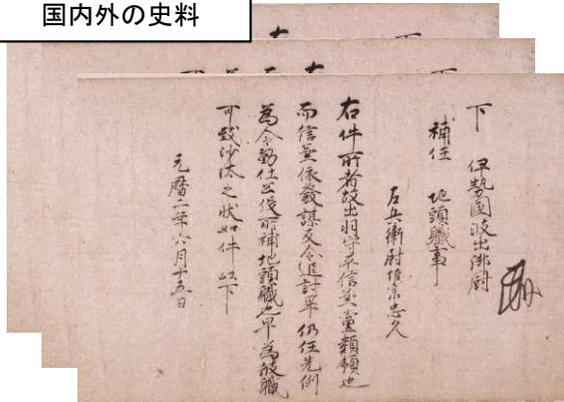
山田太造



# 史料編纂所における研究活動

## 史料採訪

国内外の史料



## 収集・調査

メタデータ

- 文書名: 源頼朝下文
- 所蔵先: 東京大学史料編纂所
- 和暦: 元暦2年6月15日
- ...

デジタル画像

- JPEG 9184x7239 8-bit sRGB

## 編纂

翻刻

索引

再編成

網文

下 伊勢国波出御厨  
 補任 地頭職事  
 左兵衛尉惟宗忠久  
 右件所者故出羽守平信兼

オブジェクト識別・同定

- 史料
- 人名
- 地名
- ...

## 出版

平信兼  
類ノ所領黨

頼朝伊勢忠久  
出羽守伊勢忠久  
頭職御勢波  
補地頭

下 伊勢国波出御厨  
 (頼朝) (花押)  
 補任 地頭職事  
 (異筆) 左兵衛尉惟宗忠久  
 右件所者故出羽守平信兼黨類領也、而信兼依發謀反、令追討畢、仍任先例、爲令勤仕公役、所補地頭職也、早爲彼職、可致沙汰之狀如件、以下、

例: 『大日本古文書』

- 史料集
- DB公開

# 史料編纂所のデータベース

https://wwwap.hi.u-tokyo.ac.jp/ships/

The screenshot shows the search interface of the HI database. At the top, there is a search bar with the text "データベース検索" and a search button labeled "横断検索". Below the search bar, there are four main categories of search results:

- 史料の所在** (Location of Documents):
  - 所蔵史料目録データベース (HI-CAT)
  - Hi-CAT Plus
  - 日本古文書ユニオンカタログ
  - 正倉院文書マルチ支援データベース
  - Catalogue Database of the Batavia's Uitgaand Briefboek, 1621-1792
- できごとを主題に** (By Event):
  - 大日本史料総合データベース
  - 編年史料カード (古代関係) データベース
  - 近世編年データベース
  - 維新史料綱要データベース
  - 近世史編纂支援データベース (摺物データベース)
  - 錦絵データベース
  - 編年史料 (古代) 編纂支援資源化データベース MIDOH
- 画像をさぐる** (Browse Images):
  - 歴史絵引データベース
  - 肖像情報データベース
  - 史料編纂所所蔵肖像画模本データベース
  - 史料編纂所所蔵荘園絵図模本データベース
  - 金石文拓本史料データベース
  - 摺物データベース
  - 錦絵データベース
  - 古写真データベース
- 全文の検索** (Full-text Search):
  - 古記録フルテキストデータベース
  - 古文書フルテキストデータベース
  - 奈良時代古文書フルテキストデータベース
  - 平安遺文フルテキストデータベース
  - 鎌倉遺文フルテキストデータベース
- 人を主題に** (By Person):
  - (大日本史料総合データベース)
  - 中世記録人名索引データベース
  - 花押データベース
  - 志形見データベース
  - 大日本史料7編人名カードデータベース (肖像情報データベース)
  - (史料編纂所所蔵肖像画模本データベース)
- 辞典・字典** (Dictionaries):
  - 応答型翻訳支援システム(ダウンロードページへ)
  - 欧文日本古代史料解題辞典データベース(Dictionary of Sources of Classical Japan) (Dictionnaire des sources du Japon)
  - 電子くずし字字典データベース
  - 史的文字データベース連携検索システム
  - 木簡・くずし字解読システム (MOJIZO)

■ DB数: 32

## ■ 種類

- 目録
- 本文
- 画像・図像・写真
- 知識
  - 人名
  - 地名
  - 和暦
  - 文字

## ■ 範囲

- 日本前近代
- 奈良～維新

## ■ 史料集編纂において必要な情報をデータベース化

登録件数(2024年1月時点):  
約730万レコード, 約2110万画像, 約9.6億文字  
アクセス数(2022年度):  
データベース 約781万件, 画像 約880万件

# 所蔵史料と所外の史料

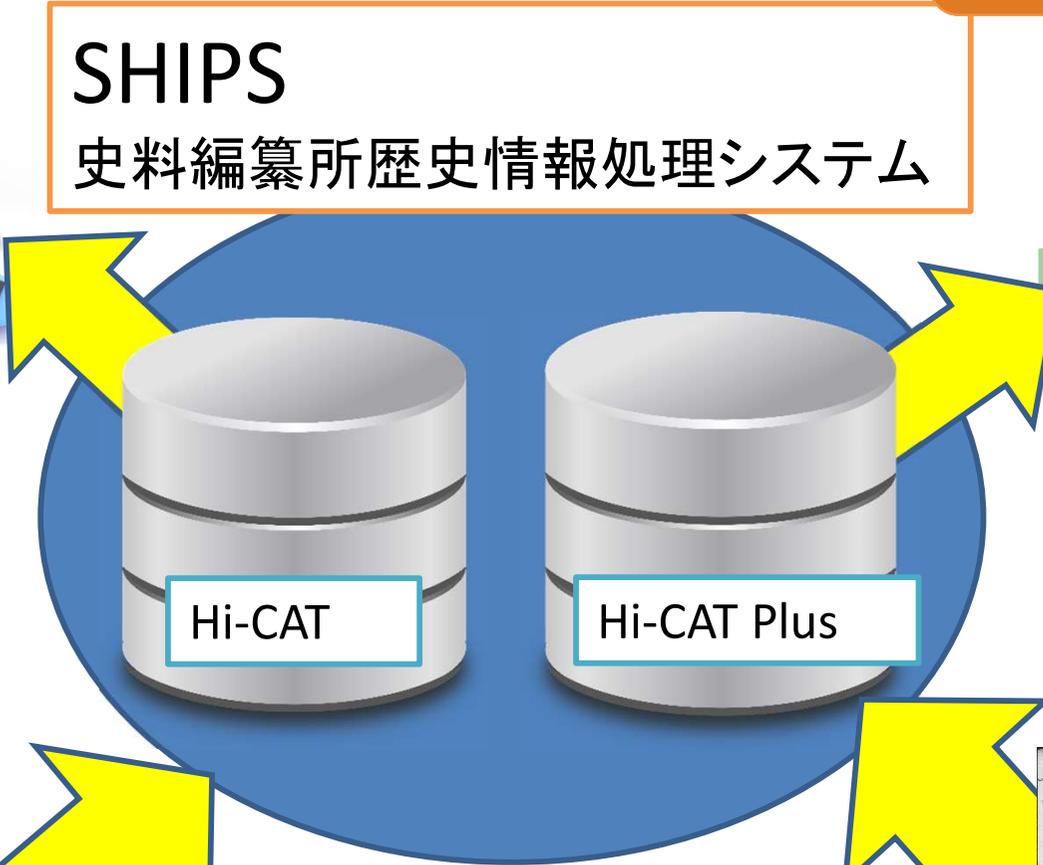
2020年3月よりweb  
公開(一部)

## SHIPS

史料編纂所歴史情報処理システム



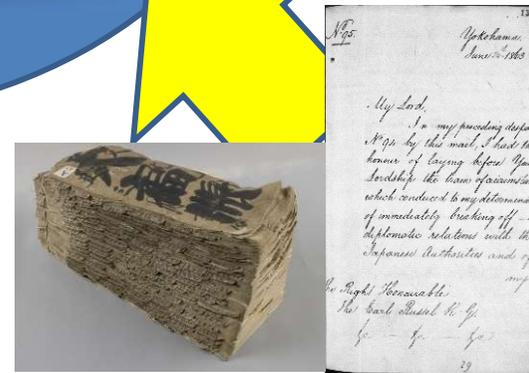
Web公開



閲覧室での公開  
+ 連携機関での閲覧  
例: 秋田県公文書館.  
横浜開港資料館



史料編纂所所蔵史料の画像



国内・海外から収集した史料の画像

# 編纂と史料複製

## ■史料採訪 約50件(約200史料群) /1年

- 史料編纂所以外にある史料が対象
- 史料の調査及び収集(複製作成・撮影)
- 1885年より本格化, 135年間

## ■収集

### – 当初:

- 影写本(敷き写し), 謄写本(見取り写し), 模写など
- 並行してガラス乾板による撮影

### – マイクロフィルム:

- 1950年より開始, 1970年ごろから本格化. ~ 2010年
- スキャニング 2008年- 2015年. 約 12,000 リール
- 国外分: 1930年頃から収集開始. デジタル化は2014年度より本格的に開始

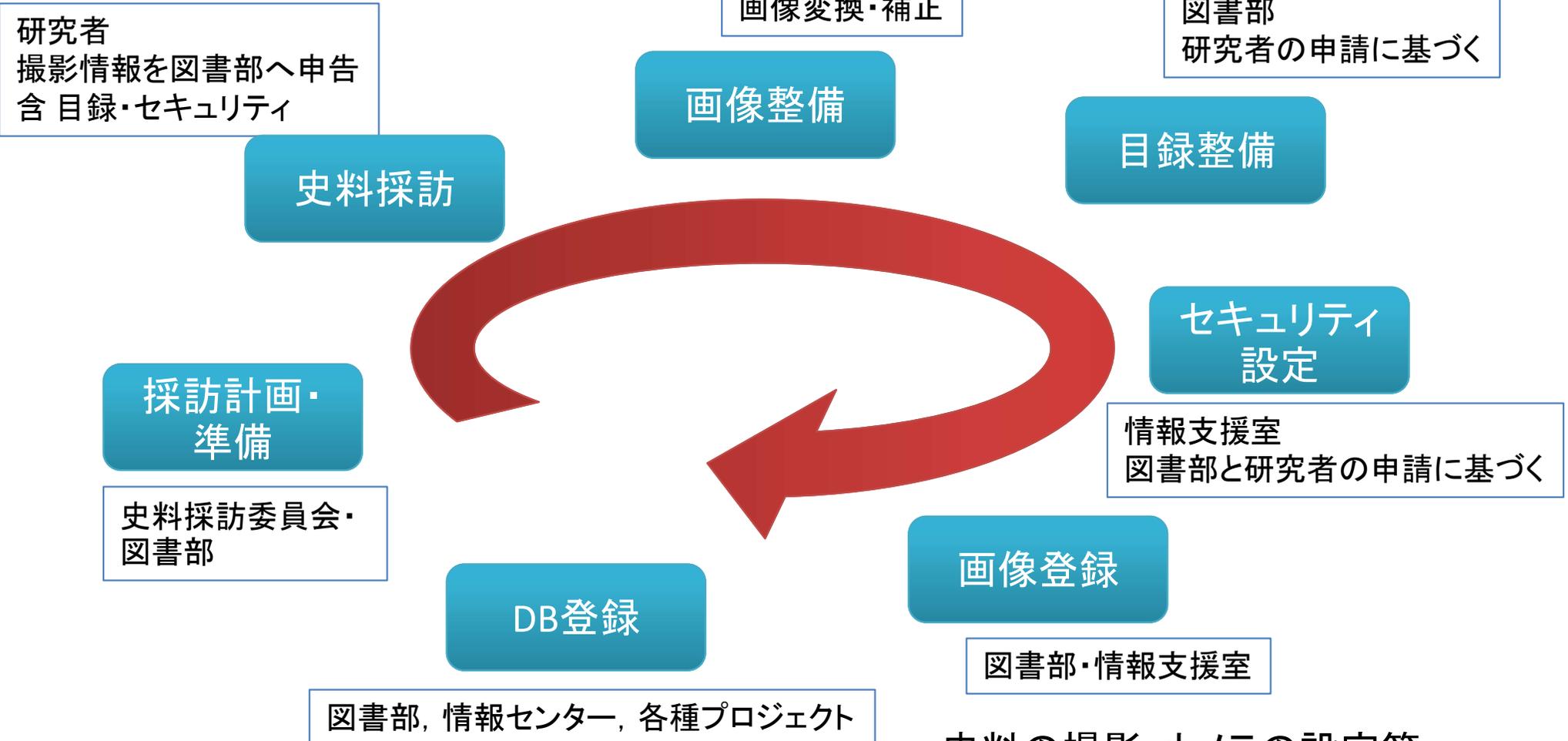
### – デジタル撮影:

- 2010年頃から

# 史料採訪の流れ

規約化は 2008年開始(3年かかった)  
 試行は2010年, 実施は2011年から  
 運用後10年経過, 見直し中

史料編纂所内で規約化  
 デジタル採訪小委員会でフローを策定



史料の撮影・カメラの設定等  
 のマニュアルも用意。

# 史料画像デジタル化進捗管理システム

史料画像デジタル化進捗管理システム ログアウト ヘルプ

検索結果一覧画面 > 画像管理情報参照画面

【画像管理情報参照画面】

< 前へ 次へ >

画像管理情報 (本所所蔵史料)

印刷表示 進捗確認 データコピー 画像管理情報編集

No.	項目	記入・選択項目
1	B D 番号	BD2099-YYYYYY
2	B D 番号申請者	図書部
3	画像管理情報記録者	図書部
4	画像管理情報記録年月日	2019-04-03
5	撮影対象	本所所蔵史料
6	撮影対象 (備考)	(本所所蔵史料) ←必須
7	史料 (群) 名	【記入例】 加判之列勤方覚 ←書名もしくは史料名(冊次、巻次がある場合は付け加える)を入力
8	史料 (群) 名ヨミ	カハンノレツツトメカタオボエ
9	原蔵者	東京大学史料編纂所
10	原蔵者ヨミ	トウキョウダイガクシリョウヘンサンジヨ
11	原蔵者所在地	東京都文京区本郷7丁目3-1
12	撮影代表者	図書部、[所外者撮影] ←必須
13	撮影メンバー	堀内カラー[撮影] ←業者名必須
14	撮影年月日	2019-04-01~2019-04-02
【画像種別】		
15	撮影時データ備考	RAW
16	ファイル数	113
17	原蔵番号	備後福山阿部家史料-02-014 ←所蔵史料目録の架番号もしくは未整理番号を入力
18	フォルダ分化	有
19	分化後フォルダ名	BD2099-YY0000~BD2099-YY0098 ←※別途撮影リストを作成してください
20	目録注記事項	
21	撮影目録の有無	
22	利用条件確認先 (担当者名・所属)	

史料画像デジタル化進捗管理システム ログアウト ヘルプ

検索結果一覧画面 新規登録

検索条件

B D 番号

撮影対象

史料 (群) 名

原蔵者

撮影代表者

現在の進捗状態

ステータス

クリア 検索

2 3 4 5 次へ > 最終へ >>

1件 ~ 20件 / 3458件中

B D 番号	撮影対象	史料 (群) 名	原蔵者	撮影代表者	撮影年月日	進捗状態	画像管理情報
BD2099-YYYYYY	本所所蔵史料	【記入例】 加判之列勤方覚 ...	東京大学史料編...	図書部、[所外...	2019-04-01...	(7) 作業中	参照 編集
BD2099-XXXXXX	本所所蔵史料	【記入例】 扶桑略記 廿五 ...	東京大学史料編...	図書部	2019-04-01...	(2) 未着手	参照 編集
BD2021-007200	所外史料	醍醐寺文書 380函	醍醐寺	高橋慎一朗	2021-02	(10) 未着手	参照 編集
BD2021-007100	所外史料	小川八幡神社大般若経	小川八幡神社	遠藤基部	2021-07-23	(3) 未着手	参照 編集
BD2021-007000	本所所蔵史料	宗家史料	東京大学史料編...	荒木裕行		(6) 未着手	参照 編集
BD2021-006900	本所所蔵史料	宗家史料	東京大学史料編...	荒木裕行		(6) 未着手	参照 編集
BD2021-006800	本所所蔵史料	史談会本	東京大学史料編...	保谷 徹	2020-10-01~	(6) 未着手	参照 編集
BD2021-006700	本所所蔵史料	万里小路惟房書状	東京大学史料編...	村井	2021-07-01~	(5) 未着手	参照 編集
BD2021-006600	本所所蔵史料	称光天皇口宣案	東京大学史料編...	村井	2021-07-01~	(5) 未着手	参照 編集
BD2021-006500	本所所蔵史料	伊達政宗消息	東京大学史料編...	村井	2021-07-01~	(5) 未着手	参照 編集

もともとは独自の保存モデル  
-> OAIS参照モデルに準拠

# 画像閲覧システム

The screenshot displays the SHIPS Image Viewer interface. At the top, it shows the University of Tokyo logo and the title 'SHIPS Image Viewer'. A navigation bar includes buttons for '全画面', 'ホーム', '拡大', '縮小', '左回転', '右回転', 'ナビ表示 OFF', '印刷プレビュー', and '一覧へ'. Below this is a control panel with a '隠す' button, a page indicator '11 / 21', and navigation arrows. A table of contents on the left lists image IDs from 00000003 to 00000018, with 00000008 highlighted. The main area shows a document page with vertical Japanese text. A ruler at the bottom indicates the page is from '0014-1 三宝院伝法灌頂私記 内道場' and is part of the '東京大学史料編纂所' collection. A PermaLink is provided at the bottom: <https://clioimg.hi.u-tokyo.ac.jp/viewer/image/ldata/000/0014/1/00000008.jpg>

## 1. 史料編纂所が所蔵する原本等の史料画像データ

- 史料種別(区分)が「貴重書(原本・古写本類)」・「特殊蒐書」(一部を除く)・「写本」であればCC BY相当
- “オープンデータ”

## 2. 史料編纂所出版物の版面画像データ

- 史料種別(区分)が「本所出版物」の画像データはCC BY-NC-SA相当

## 3. 上記以外の史料等画像データ

- 史料編纂所図書室へ申請

■ 上記は2019年4月1日より開始

# 構築推進事業

## ■共有化

- デジタルデータの長期保存・長期利用: OAI参照モデル適用
- データ利用条件の整備

## ■国際化

- 海外発信の強化英語版「維新史料綱要DB」

## ■連結化

- API整備: IIFの利用, 画像データの二次利用促進
  - 天皇家・公家文庫収蔵等の史料画像web公開
  - くずし字連携: 史的文字データベース連携システム (主幹: 奈良文化財研究所)

## ■JDCatへの寄与

- 史料編纂所所蔵史料「備後福山藩阿部家史料」: 約55,000件
- 神奈川県立金沢文庫所蔵「国宝 金沢文庫文書」: 1,024件
  - 神奈川県立金沢文庫とのシステム連携

## JPCOARによる記述

```
<jpcoar:jpcoar xmlns:datacite="https://schema.datacite.org/meta/kernel-4/" xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcdnl="http://ndl.go.jp/dcdnl/terms/" xmlns:dcterms="http://purl.org/dc/terms/" xmlns:jpcoar="https://github.com/JPCOAR/schema/blob/master/1.0/"
xmlns:oaire="http://namespace.openaire.eu/schema/oaire/" xmlns:rdflib="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rioxxtterms="http://www.riox.net/schema/v2.0/rioxxtterms/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="https://github.com/JPCOAR/schema/blob/master/1.0/jpcoar_scm.xsd">
  <dc:title xml:lang="ja">西御丸当番其外勤方留書</dc:title>
  <dc:title xml:lang="ja-Latn">nishiomarutobansonohokatsutomekatatomegaki</dc:title>
  <jpcoar:creator>
    <jpcoar:creatorName xml:lang="ja">阿部, 正右</jpcoar:creatorName>
  </jpcoar:creator>
  <jpcoar:contributor contributorType="Distributor">
    <jpcoar:nameIdentifier nameIdentifierScheme="ISNI" nameIdentifierURI="http://isni.org/isni/0000000121877782">0000000121877782</jpcoar:nameIdentifier>
    <jpcoar:contributorName xml:lang="ja">東京大学史料編纂所</jpcoar:contributorName>
    <jpcoar:contributorName xml:lang="ja">Historiographical Institute, the University of Tokyo</jpcoar:contributorName>
  </jpcoar:contributor>
  <dcterms:accessRights rdf:resource="http://purl.org/coar/access_right/c_abf2">open access</dcterms:accessRights>
  <jpcoar:subject subjectScheme="NDC" xml:lang="ja">日本史</jpcoar:subject>
  <jpcoar:subject subjectScheme="NDC" xml:lang="en">General History of Japan</jpcoar:subject>
  <datacite:description descriptionType="Abstract" xml:lang="ja">宝暦2年5月4日初当番。宝暦2年5月4日～6月12日</datacite:description>
  <datacite:description descriptionType="Other" xml:lang="ja">史資料: テキスト</datacite:description>
  <datacite:description descriptionType="Other" xml:lang="en">materials: text</datacite:description>
  <dc:language>jpn</dc:language>
  <dc:type rdf:resource="http://purl.org/coar/resource_type/c_0040">manuscript</dc:type>
  <jpcoar:identifier identifierType="URI">https://wwwap.hi.u-tokyo.ac.jp/ships/w01/detail/T50/01/001</jpcoar:identifier>
  <jpcoar:relation relationType="references">
    <jpcoar:relatedIdentifier identifierType="Local">T50-1-1</jpcoar:relatedIdentifier>
  </jpcoar:relation>
  <jpcoar:relation>
    <jpcoar:relatedIdentifier identifierType="DOI">10.57459/hi.01.T50_01_001</jpcoar:relatedIdentifier>
  </jpcoar:relation>
  <dcterms:temporal>1752-06-15/1752-07-22</dcterms:temporal>
</jpcoar:jpcoar>
```

アイテム

### 西御丸当番其外勤方留書

<https://jdcatalog.jp/records/9241>

<b>Item type</b>	Multiple(1)
<b>公開日</b>	2023-01-18
<b>タイトル</b>	西御丸当番其外勤方留書 nishiomarutobansohokatsutomekatatomegaki
<b>作成者</b>	阿部, 正右
<b>寄与者</b>	Distributor 東京大学史料編纂所, Historiographical Institute, the University of Tokyo
<b>URI</b>	<a href="https://wwwap.hi.u-tokyo.ac.jp/ships/w01/detail/T50/01/001">https://wwwap.hi.u-tokyo.ac.jp/ships/w01/detail/T50/01/001</a>
<b>アクセス権</b>	open access
<b>内容記述</b>	Abstract 宝暦2年5月4日初当番。 宝暦2年5月4日～6月12日
<b>内容記述</b>	Other 史資料: テキスト
<b>内容記述</b>	Other materials: text
<b>トピック</b>	日本史 General History of Japan
<b>データの言語</b>	jpn
<b>整理番号</b>	T50-1-1
<b>整理番号</b>	10.57459/hi.01.T50_01_001
<b>対象時期</b>	1752/06/15 - 1752/07/22

**155**

views

total

[See details](#)

**Versions**

Ver.1 2021-11-09 13:58:32.942990

[Show All versions](#)

**エクスポート**

**OAI-PMH**

JPCOAR

DublinCore

DDI

**Other Formats**

[JSON](#) [BIBTEX](#)

[← 戻る](#)

# Web公開したデータ等

## ■ 史料編纂所所蔵史料

- 備後福山藩阿部家史料: 55,000件
- 東京大学デジタルアーカイブズ構築事業

## ■ 所外の史料

- 天皇家・公家文庫史料: 主に宮内庁書陵部所蔵史料, 56万件, 2020年3月公開
- 都城島津邸所蔵史料: 2021年7月公開
- 東寺文書(琵琶湖博所蔵): 2021年12月公開
- 岩倉具視関係史料(海の見える杜美術館所蔵): 2022年10月公開
- 松尾大社所蔵史料: 2023年1月公開
- 紀州菊池本家・新家史料(和歌山県湯浅町所蔵): 2023年5月公開
- 日本中世気象災害史年表稿(藤木久志編, 高志書院): 2022年12月公開

## ■ 電子くずし字字典, 連携検索システム

- 史的文字データベース連携検索システム 2020年10月公開
  - 全体で200万件以上の文字画像
  - 史料編纂所分(電子くずし字字典DB): 約37万件 (2024年3月時点)

## ■ 英語版「維新史料綱要DB」

- “Summary database of the Ishin Shiryo” : 2020年12月公開

# 強化事業では何をやるのか？

## ■構築推進事業で行ってきたことを継続

### —JDCatとの連携

- 史料編纂所所蔵史料：画像があるもの
- 他機関所蔵史料：JDCatへの登録を承認頂いたもの

### —他機関のデータ受入や受託

- これまでのご協力頂いている史料所蔵機関の追加分
- さらに協力機関を増やしていく

## ■日本史（人文）にかかるデータについて

### —データ利活用の促進

### —JDCatの運用・改修、メタデータスキーマや統制語彙の妥当性に係る継続的な検討等

### —「データ共有のための手引き」

URL: <https://lab.hi.u-tokyo.ac.jp/>

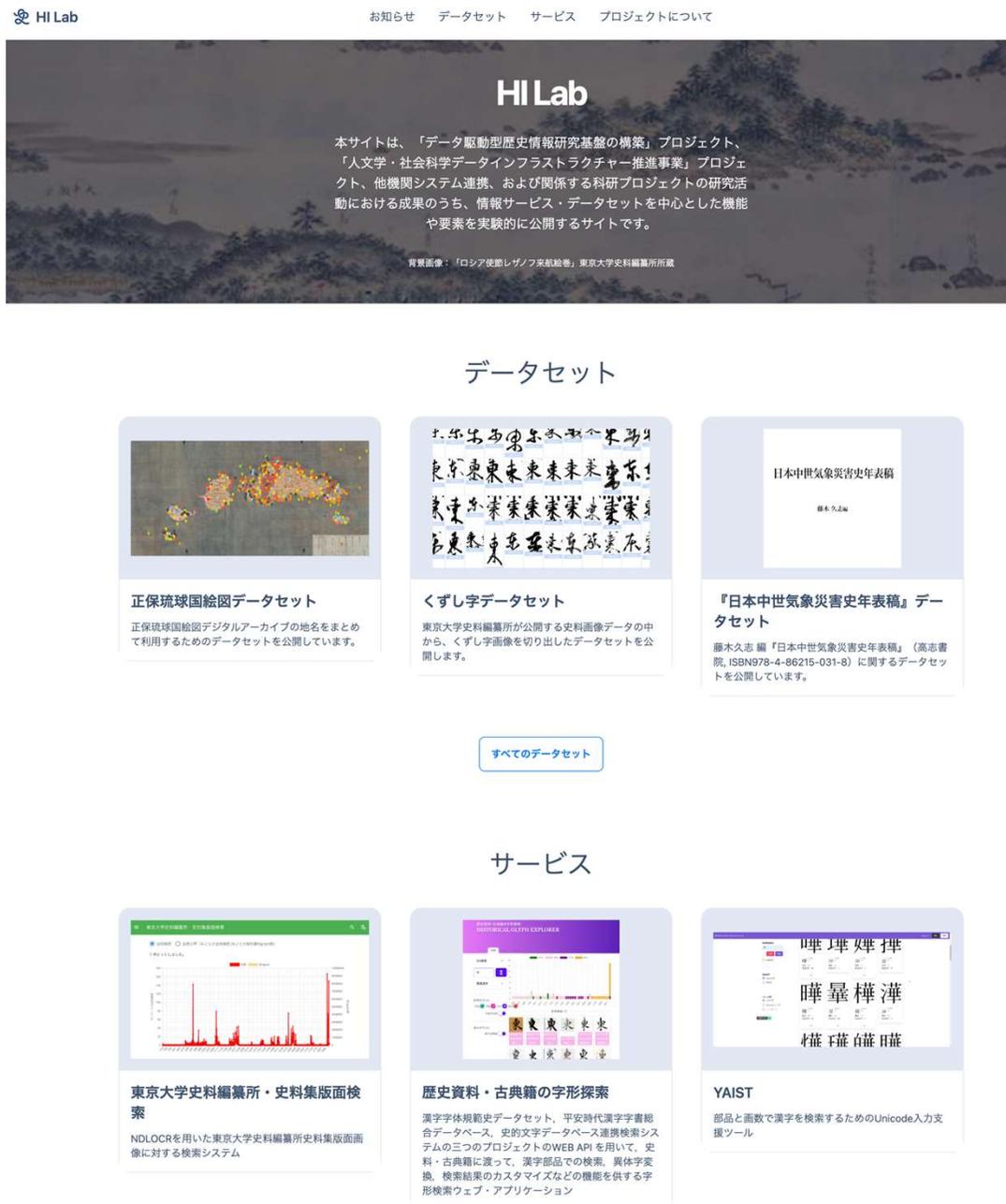
情報サービス・データセットを中心とした機能や要素を実験的に公開するサイト。データの有用性を探る。

データセット:  
AI・機械学習, Web アプリケーション等で利用可能なデータセットを公開

- ・国絵図データセット
- ・くずし字データセット
- ・...

サービス:  
実験的な情報サービス・ツール・ユーティリティを提供

- ・史料集版面検索
- ・YAIST, yaist-vscode



# おわりに

## ■人文学(日本史学)の研究においてもデータベース利活用は欠かせない

- 分野における“インフラ”になってきたかも

## ■データの収集・蓄積・管理・共有

- 史料編纂所は、130年間、形式は違えども同じ活動を行ってきた。
- 紙 -> マイクロフィルム -> デジタル画像
- データの利用条件整備の必要性

## ■他の分野でも利用できるデータへ

- JDCatを介したデータ提供の可能性

## ■Webを介した“次”へ : データ駆動型日本史研究基盤構築へ

- HI Labを介した発信
- AI・機械学習に耐えうるデータとしての整備
- 整備したデータがdata drivenな利用ができるか -> アプリケーションを作って確かめる
- このデータを学界・一般へ継続して提供していく環境整備

## ■日本史研究DX: 今まさに突入

- 研究資源のtransformation はこれまで行ってきた。
- DXへ突入