

第2回人文学・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
「研究データの利活用と流通を強化する」

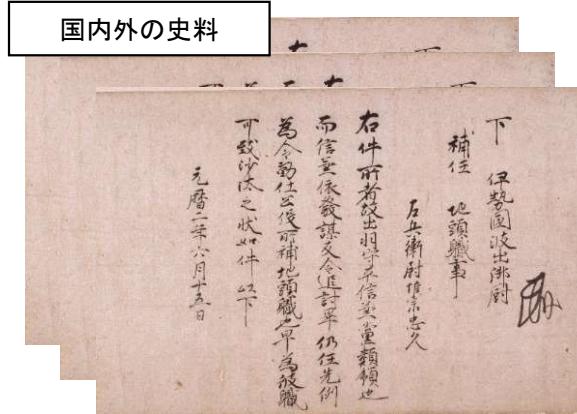
東京大学史料編纂所における取組

山田太造



史料編纂所における研究活動

史料採訪



収集・調査

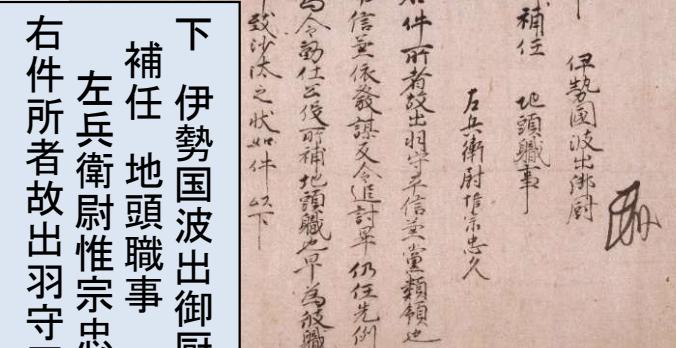
メタデータ

- 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



オブジェクト識別・同定

- 史料
- 人名
- 地名
- ...

出版

類ノ信兼
平
斯頭出ヲ頼
職御伊朝
ニ厨忠
補地波久

下
伊勢國波出御厨
(花押)
補任
地頭職事、
(異筆)
左兵衛尉惟宗忠久
(花押)

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、所補地頭職事也、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



■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

データベース検索

キーワード (例:島津家文書)

横断検索

史料の所在

- 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

図像をさぐる

- 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
- 史料編纂所所蔵図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

人を主題に

- (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
- 忘形見データベース
- 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)

できごとを主題に

- 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

全文の検索

- 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

辞典・字典

- 応答型翻訳支援システム(ダウンロードページへ)
- 歐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
- 史的文書データベース連携検索システム
- 木簡・くずし字解説システム (MOJIZO)

登録件数(2025年1月時点):

約750万レコード, 約2,120万画像, 約10億文字

アクセス数(2023年度):

データベース 約820万件, 画像 約1,070万件

所蔵史料と所外の史料

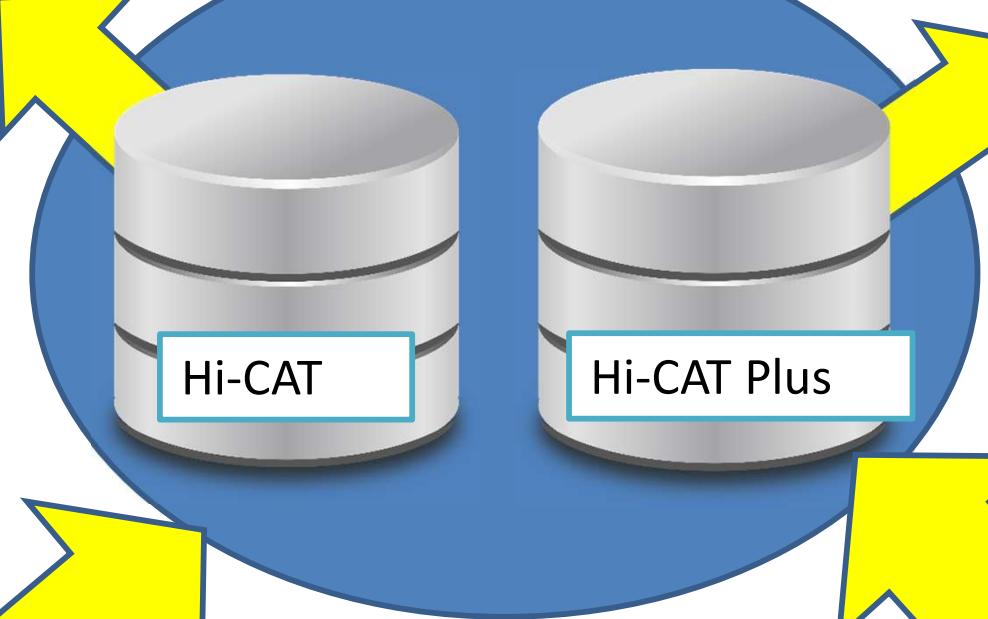
2020年3月よりweb
公開(一部)



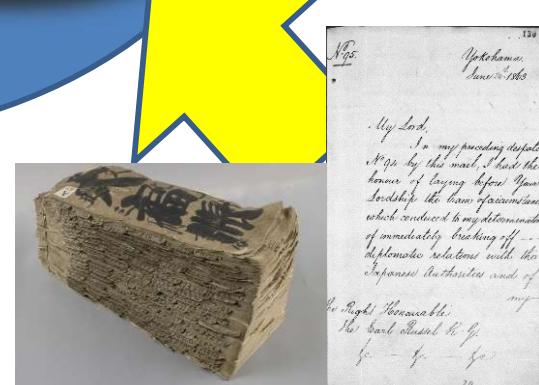
Web公開

SHIPS

史料編纂所歴史情報処理システム



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



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

画像閲覧システム

ビューア:OpenSeaDragonの利用,2015年12月~
画像サーバ:IIF, Legacy Image Pyramids対応

東京大学 SHIPS Image Viewer

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 一覧へ

所員用ログイン

11 / 21

隠す

_00000003

00000001

00000002

00000003

00000004

00000005

00000006

00000007

00000008

00000009

00000010

00000011

00000012

00000013

00000014

00000015

00000016

00000017

00000018

0014-1 三宝院伝法灌頂私記 内道場

PermaLink <https://clioimg.hi.u-tokyo.ac.jp/viewer/image/idata/000/0014/1/00000008.jpg>

東京大学史料編纂所

一初夜作法 金剛界

職衆平受有礼堂奉座乗通首參之後參阿闍梨齋
衍入堂自後戶入御西懶令修供養法取右箱置脇机
上又香呑相茶磬臺邊又仰奉住乞諸教等爐香燈明
等令拵之了出了着礼堂

次大阿闍梨光首便金二打 表白神會 祈歎 丑悔
勸請_{未句云}薦持大有或志迄立大願普供 三力次前供養讚_{讀安}誦之
四智心略 金剛薩埵錢如常 但第三版四一二不撞之
次普供養三刀 金二打 職衆誦佛眼真言

次供養法振鈴 職衆誦大日真言

聞鈴聲教授師迴堂自後戶入堂內正念誦畢起坐以秉
住法師礼盤空移他所三尊前數半疊大阿闍梨著座念
誦次秉住等令出堂內 次教授師立瓶一取各三遍迴禮
即置小壇瓶不遠本方中瓶白遙間誦大日真言置小壇之
中央宣籠青遙間誦普賢真言置東辰巳瓶黃誦詠
勸真言置南辰申瓶赤誦徐蓋障真言置西戌亥瓶黑
誦徐惠趣真言置北華嚴瓶本處

次教授師正面戸開出局加持香水呑新阿闍梨汝達舍身堂
手含香含丁子潔香水入門內以復面勿持之用不動明照復
告云聞一切諸惠趣門能聞清淨立眾次信三昧耶下口授三摩
耶薩怛錫明三支耶堅忍頸二度為針次引入體前越過香焉

一後夜特

史料編纂所所蔵史料のデータ利用条件

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

- 他機関所蔵史料
 - データ利用条件を提示, 一部**オープンデータ**
- 上記は2019年4月1日より開始

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
 - 神奈川県立金沢文庫「国宝 金沢文庫文書データベース」

—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」

今年度の取り組み(1)

■ JDCatとの連携：2,893件の史料データ

- 神奈川県立金沢文庫所蔵「国宝 金沢文庫文書」: 1,479件(455件の追加)
 - 神奈川県立金沢文庫とのシステム連携
- 史料編纂所所蔵史料「島津家本」: 639件 新規
 - 東京大学デジタルアーカイブズ構築事業との連携

■ 他機関のデータ受け入れや受託

- Hi-CAT Plus からの公開
 - 都城島津邸所蔵「都城島津家文書」: 3,923点追加 (合計: 14,905点)
 - 海の見える杜美術館所蔵「岩倉具視関係史料」: 843点追加 (合計: 6,884点)
 - 立花家史料館所蔵史料(大友家文書など): 3,587点 (新規)
 - ... x2

島津家本

旧鹿児島藩島津家の家史編纂・史料蒐集機関として存在した「公爵島津家編輯所」が、その編纂事業の過程において蒐集・作成し、第二次世界大戦後東京大学史料編纂所の所蔵となった書籍（刊本・写本・文書・写真等）の総称

JDCatでの表示

SHIPS -> JDCat のハーベストは ResourceSyncによる

アイテム

中書家久公御上京日記

Item type	Multiple(1)
公開日	2024-11-28
タイトル	中書家久公御上京日記 chushoehisakogojokyonikki
寄与者	Distributor 東京大学史料編纂所 Historiographical Institute, the University of Tokyo
URI	https://wwwap.hi.u-tokyo.ac.jp/ships/w01/detail/T42/_211sa_501r1_-12/33-124
アクセス権	open access
内容記述	Abstract 書目ID: 191670 史料区分: T42 書名よみ: ちゅうしょいえひさこうごじょうきょうにっき 架番号: T42/さI-12/33-124 冊ID: 411156 ディレクトリ: T42/_211sa_501r1_-12/33-124 注記 (書誌): [家譜関係史料 (全995冊)] 注記 (冊): マイクロフィルム: Hdup.M-38 「島津家文書マイクロ版集成 [No.島津家本]」 リール番号93, 開始コマ4 法量 (cm): 縦26.8×横19.2 マニフェストURL: https://clioapi.hi.u-tokyo.ac.jp/iiif/01/idata/T42/_211sa_501r1_-12/33-124/1/manifest 特殊蒐書: 島津家本
内容記述	Other 史資料: テキスト
内容記述	Other materials: text
トピック	日本史 General History of Japan
データの言語	jpn
整理番号	さI-12-33-124
整理番号	10.57459/hi.01.T42_211sa_501r1_-12_33-124
対象時期	1818 - 1897

DOI

所蔵史料目録データベース (Hi-CAT)

データベース検索 > 所蔵史料目録データベース

史料編纂所 SHIPSDBへリンク

«書目データ»

【史料種別】 貴重書 (特殊蒐書)

【請求記号】 島津家本-さ I-12-33-124

【書名】 中書家久公御上京日記

【著者名】

【出版事項】

【大きさ】 袋綴冊子。1冊

【注記】 [家譜関係史料 (全995冊)]

【解題】

【イメージ】

«冊データ»

【注記】

マイクロフィルム: Hdup.M-38 「島津家文書マイクロ版集成 [No.島津家本]」 リール番号93, 開始コマ4
法量 (cm): 縦26.8×横19.2

校正の紙片数、取り扱い注意

[10.57459/hi.01.T42_211sa_501r1_-12_33-124](https://doi.org/10.57459/hi.01.T42_211sa_501r1_-12_33-124)

【利用制限】

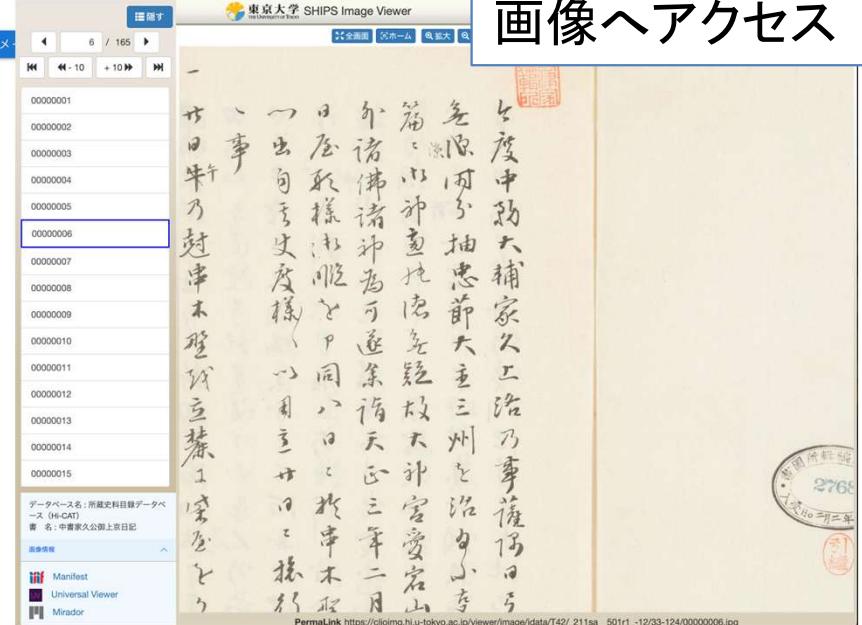
【DOI】

【細目データ】

【原蔵者分類番号】

【解題】

【イメージ】



画像へアクセス

今年度の取り組み(2)

■データ利活用の促進

- 人文学データシンポジウムの開催

- テーマ：「」
- 登壇者
 - 中村 覚氏 (東京大学史料編纂所)
 - 高田祐一氏 (奈良文化財研究所)
 - 橋本雄太氏 (国立歴史民俗博物館)
 - 上阪彩香氏 (大阪成蹊大学データサイエンス学部)
 - 佐藤健一氏 (滋賀大学データサイエンス学部)
- 2024.9.18, 東京大学本郷キャンパス, 国際会議JADHプレシンポジウムとして

- 国際会議EAJRS2024への参加・報告

- 2024.9.11-14, Sofia Bulgaria
- 社会科学研究所とのコラボレーション(中核機関として)
- ブースの設置
- 報告 : Initiatives for Sharing and Utilization of Humanities and Social Science Data at the University of Tokyo 東京大学における人文学・社会科学データの共有・利活用に関する取組



おわりに

■ 史料の“力”を再評価していく

- 高付加価値化へ

■ データの収集・蓄積・管理・共有

- 他の分野でも利用 → JDCatを介したデータ提供による可能性

■ JDCatでの利用を促進するには…

- データ件数:もっと増やす
- ユースケース:つくっていく

■ 史料をもっと利用していくためには

- 史料の目録のみではなく、本文の検索も
- 本文を作るのは大変…OCRの活用も検討していくべき

■ 史料所蔵機関の要望・相談

- 史料のデータ公開を行いたいという要望
- でも経費や技術がない…という課題

■ Webを介した“次”へ

- 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 → 確認を踏まえてアプリケーションを作ってみる
- 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