

地上気象観測原簿アーカイブ：現状と課題



北本朝展（ROIS-DS人文学オープンデータ共同利用センター / 国立情報学研究所）

<http://codh.rois.ac.jp/>

デジタル台風：地上気象観測原簿アーカイブ

<http://agora.ex.nii.ac.jp/digital-typhoon/genbo/>

デジタル台風：地上気象観測原簿アーカイブ

ホーム > Earth > デジタル台風 > 気象原簿アーカイブ

気象庁の気象官署が作成した地上気象観測原簿の画像ファイルを、閲覧しやすいようにIIIF形式でアーカイブ化し、IIIF Curation Viewerで閲覧可能としました。現在の画像枚数は4,890,691枚です。

1. 地上気象観測原簿

種類ごとのデータ概要

	種類	件数	画像数	開始	終了
1	気象月表	474	285,602	1879年01月	1961年12月
2	気象月表原簿	1,950	1,018,277	1878年01月	1966年12月
3	地上気象観測月原簿	10	71,858	1980年	1989年
4	普通気候観測月原簿	2,962	2,308,708	1953年01月	1982年12月
5	区内気象観測原簿	1,574	727,272	1898年01月	1979年06月
6	区内気象観測毎時降水量原簿	965	478,974	1950年01月	1978年12月

地点と西暦で検索

地点： 西暦：年

- 地上気象観測原簿の画像4,890,691枚を閲覧しやすくしたサイト
- IIIF形式を活用することで、IIIF関連OSS資産を活用
- 基本的な検索機能を提供

支援

- [ROIS-DS-JOINT \(027RP2021\)](#), 東日本に凶作をもたらした天候の時空間構造の文書記録を活用した解析, 研究代表者: 増田 耕一
- [ROIS-DS-JOINT \(032RP2020\)](#), 古日記天気記録の定量化に関する研究, 研究代表者: 庄 建治朗
- [JSPS科研費 \(JP22H04938\)](#), ミレニアム大気再解析プロダクトの創出, 研究代表者: 芳村 圭
- [JSPS科研費 \(JP21H03776\)](#), 江戸時代の東北地方における死亡危機の要因分析システムの構築, 研究代表者: 川口 洋
- [JSPS科研費 \(JP20K01152\)](#), 江戸時代の日記天候情報による高時空間分解能の日射量復元, 研究代表者: 市野 美夏

データ整理の方針

1. **大分類**：原簿を6種類に分類
2. **中分類**：気象業務支援センターが提供するデータが用いているフォルダを活用
3. **画像**：
 - 中分類に存在するサブフォルダはすべて展開。
 - ファイル名が重複しないようにサブフォルダ文字列を調整。
 - 中分類フォルダの画像ファイル名をソートし、全体を一つのIIIマニフェストとして作成。
4. **メタデータ**：中分類のみ。個々の画像にはない

データリストの例

	名称	地点	期間	画像数	IIIF
1	気象月表 網走1921～1930	網走	1921年01月 ～ 1930年12月	502	
2	気象月表 札幌～高知1882～1884	札幌,根室,寿都,函館,青森,秋田,宮古,野蒜,新潟,金沢,岐阜,浜松,東京,境,京都,下関,広島,大阪,和歌山,長崎,鹿児島,宮崎,高知,沼津	1882年10月 ～ 1884年12月	652	
3	気象月表 札幌～高知1885～1886	札幌,根室,寿都,函館,青森,秋田,宮古,野蒜,新潟,金沢,伏木,岐阜,浜松,東京,境,京都,下関,広島,大阪,和歌山,長崎,鹿児島,宮崎,高知,釜山,宗谷,沼津	1885年01月 ～ 1886年05月	628	
4	気象月表 札幌～高知1886～1887	札幌,根室,寿都,函館,青森,秋田,宮古,野蒜,新潟,金沢,伏木,岐阜,銚子,浜松,東京,境,京都,下関,広島,大阪,和歌山,嵯原,大分,長崎,鹿児島,宮崎,高知,釜山,宗谷,沼津,襟裳	1886年06月 ～ 1887年02月	657	
5	気象月表 札幌～高知1887～1887	札幌,根室,寿都,函館,青森,秋田,宮古,野蒜,新潟,金沢,伏木,岐阜,銚子,浜松,東京,境,京都,下関,広島,大阪,和歌山,大分,長崎,鹿児島,宮崎,高知,宗谷,沼津,襟裳	1887年03月 ～ 1887年10月	647	
6	気象月表 札幌～高知1887～1887	札幌,根室,寿都,函館,青森,秋田,宮古,石巻,新潟,金沢,伏木,岐阜,銚子,浜松,東京,境,京都,下関,広島,大阪,和歌山,大分,長崎,鹿児島,宮崎,高知,宗谷,沼津,襟裳	1887年11月 ～ 1887年12月	168	
7	気象月表 札幌1879～1890	札幌	1879年01月 ～ 1890年12月	405	
8	気象月表 札幌1904～1922	札幌	1904年01月 ～ 1922年12月	1882	
9	気象月表 札幌1923～1929	札幌	1923年01月 ～ 1929年12月	1285	
10	気象月表 札幌1930～1933	札幌	1930年01月 ～ 1933年12月	1019	

IIIF (トリプルアイエフ) とは？

International Image
Interoperability
Framework = 国際的な
画像配信方式



Web : HTML
画像 : IIIF

IIIFサー
ビス1

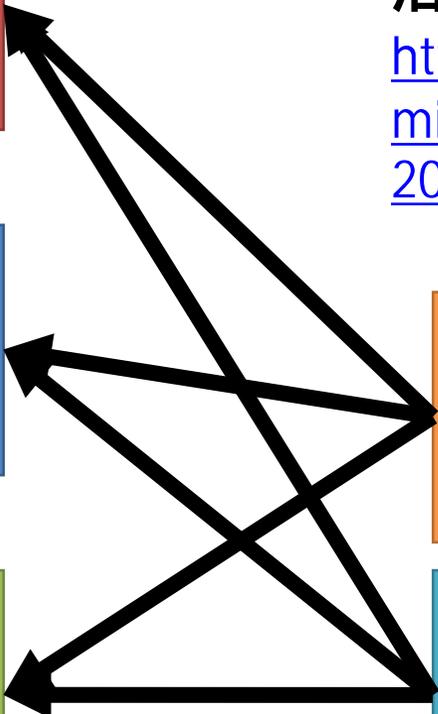
IIIFサー
ビス2

IIIFサー
ビス3

第14回CODHセミナー -
IIIF Curation Platform利
活用レシピ100連発
<http://codh.rois.ac.jp/seminar/icp-recipe-20210218/>

IIIFビュー
ア1

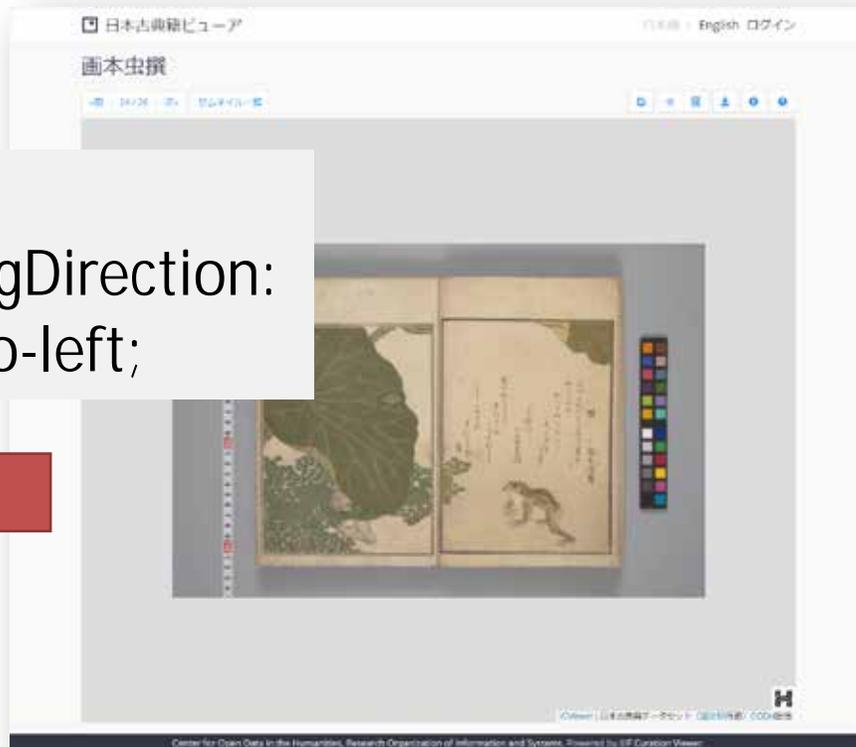
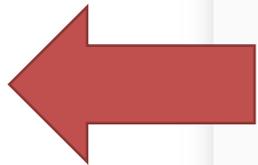
IIIFビュー
ア2



IIF Curation Viewerによる書籍の閲覧

和書

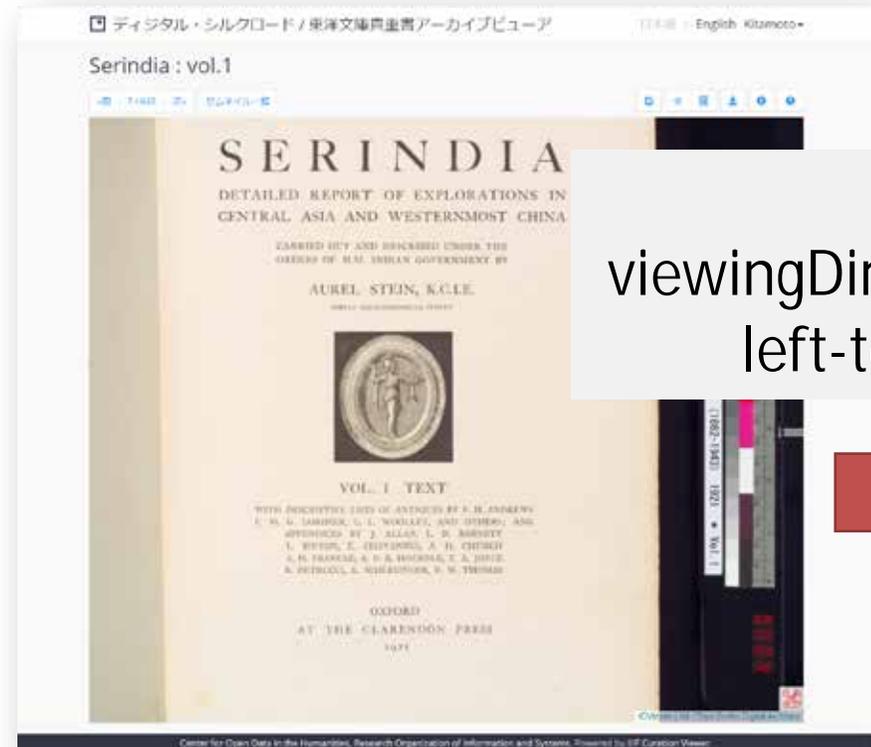
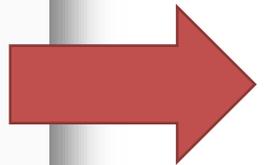
viewingDirection:
right-to-left;



日本古典籍データセット
<http://codh.rois.ac.jp/pmjt/>

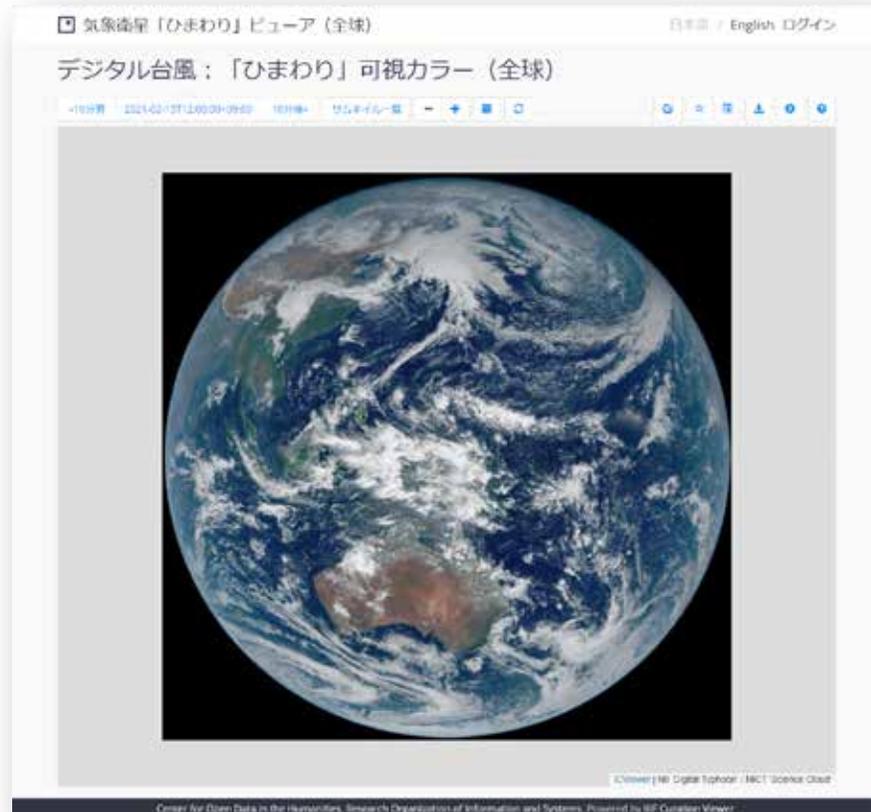
洋書

viewingDirection:
left-to-right;



『東洋文庫所蔵』貴重書アーカイブ
<http://dsr.nii.ac.jp/toyobunko/>

時系列画像の閲覧

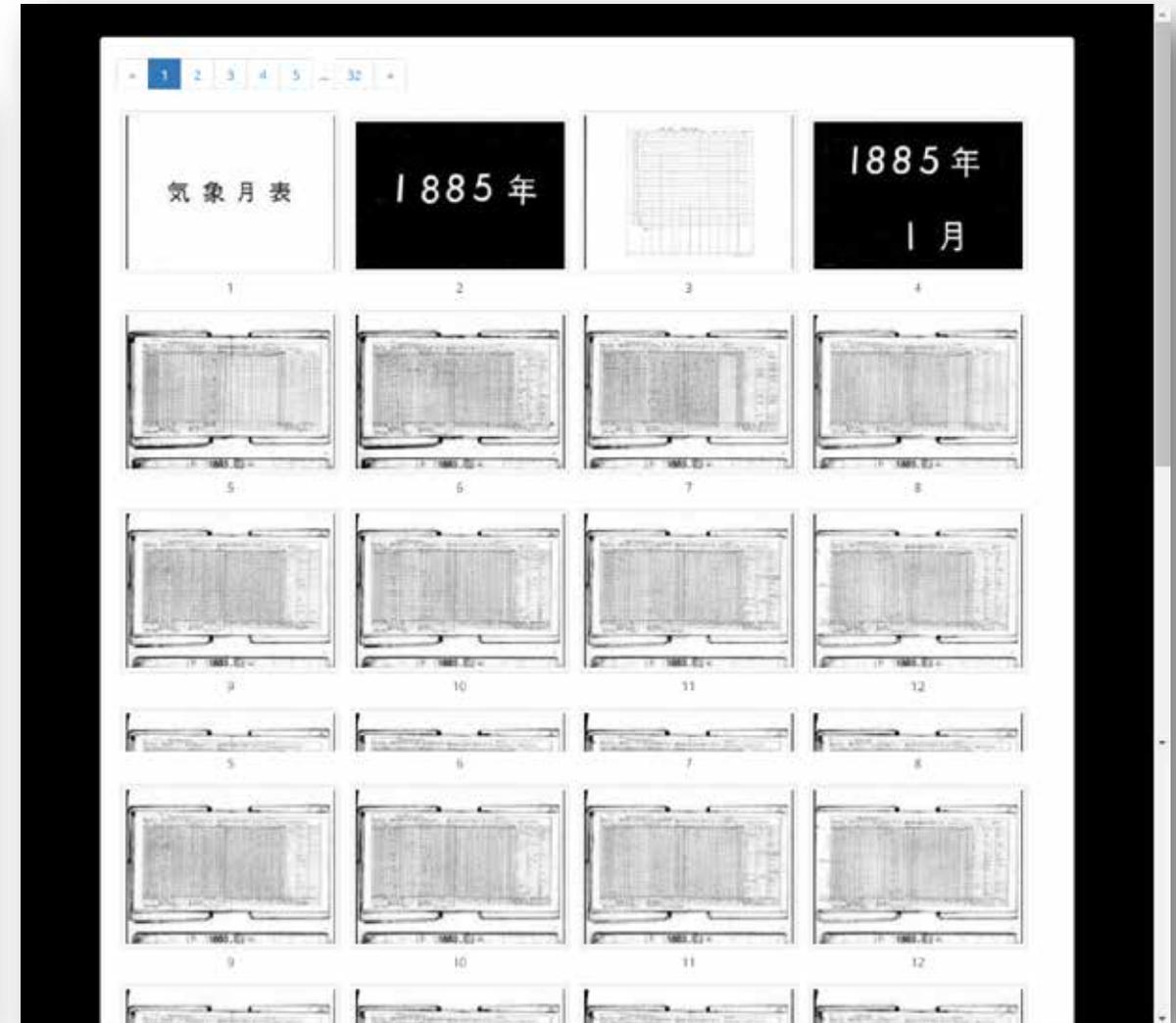


デジタル台風：気象衛星「ひまわり」ビューア

<http://agora.ex.nii.ac.jp/digital-typhoon/himawari-3g/clipping/>

1. 気象衛星「ひまわり」が観測した高解像度画像（11,000画素×11,000画素）を表示。
2. 「前」「次」は「ページ」ではなく「時間」になる。カレンダーからも日付選択可能。
3. 現在のIIIFの枠組みでは適切に扱えないため、Timeline APIとCursor APIを独自に拡張。

気象原簿の閲覧



課題1：画像回転

IIF Curation Viewer 日本語 / English ログイン

気象月表 網走1921～1930

2/502 サムネイル

年	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月	年計
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930

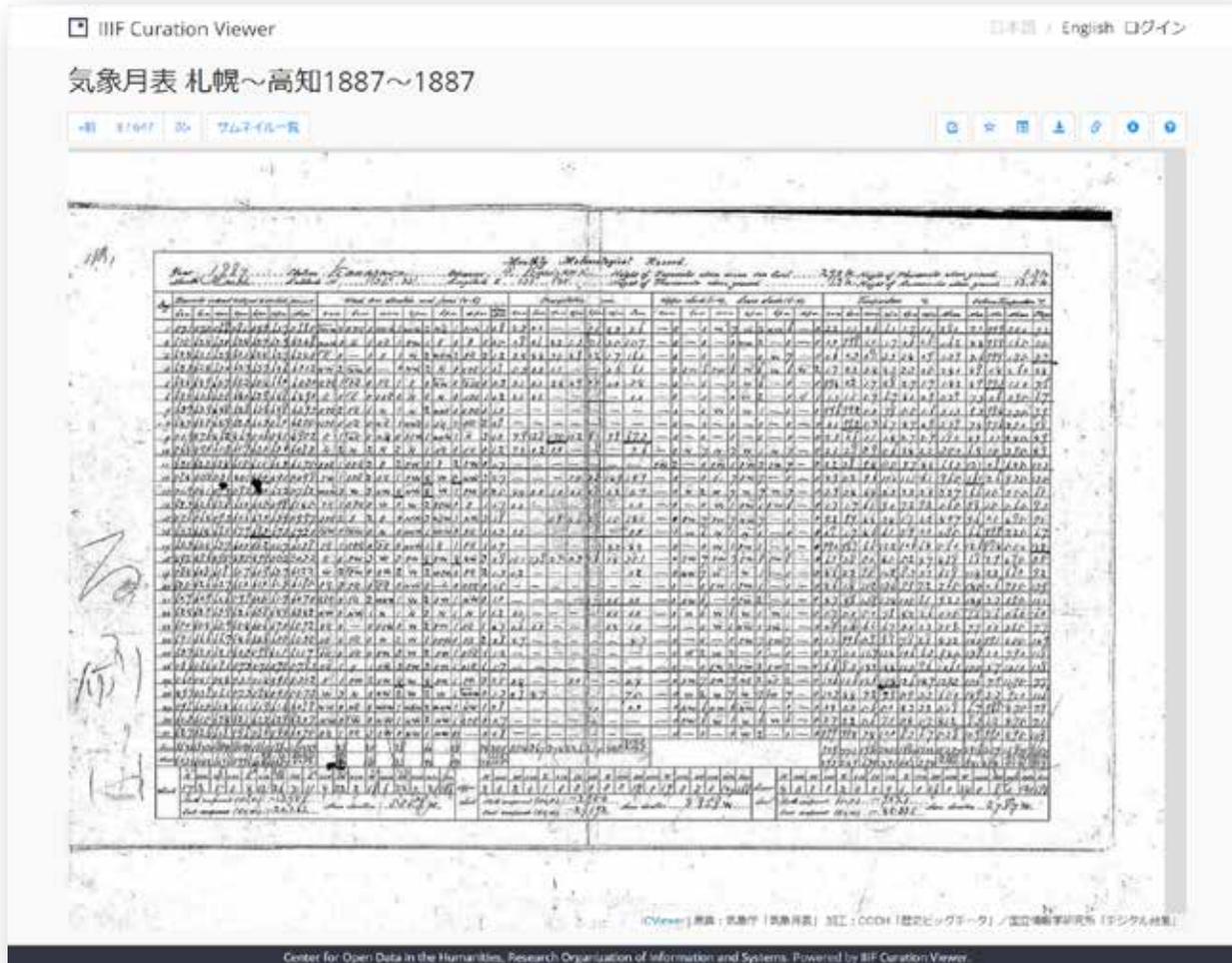
Center for Open Data in the Humanities, Research Organization of Information and Systems, Powered by IIF Curation Viewer.

1. **画像の向き**を揃えてデジタル化していないため、向きがバラバラで読みづらい
2. IIF Curation Viewerは、**画像回転ができない**という大きな制約あり
3. 回転するのは**サーバ側**か**クライアント側**か？

課題1：解決方法

1. **サーバ側回転**：手動でマーク = 手動のチェックは大変すぎる
2. **サーバ側回転**：機械学習などで自動検出 = 教師データの準備が必要
3. **サーバ側回転**：OCR結果を参考に自動検出 = 縦書きと横書きが混在する画像はどうする？
4. **クライアント側回転**：回転可能なIIFビューア (Mirador, Universal Viewer等) を活用 = 事前準備不要

課題2：画像メタデータ



1. **メタデータは中分類のみ「気象月表 札幌～高知 1887～1887」**
2. 個別の画像について、**いつの、どの地点の、どんな変数**があるのか、全くわからない
3. **画像メタデータ**があれば、検索も可能になる

課題2：解決方法

1. IIF Curation Viewerのキュレーション機能を用いて、個別の画像にメタデータを付けていく
2. NDL OCRなどを用いてテキストを抽出するか、手書き文字用のOCRを新規に開発する
3. 作成済みのデータをインデックスに用いることで、徐々に検索可能な範囲を広げていく
4. 画像特徴などを用いて機械学習による自動分類を行い、仮のメタデータを付与していく

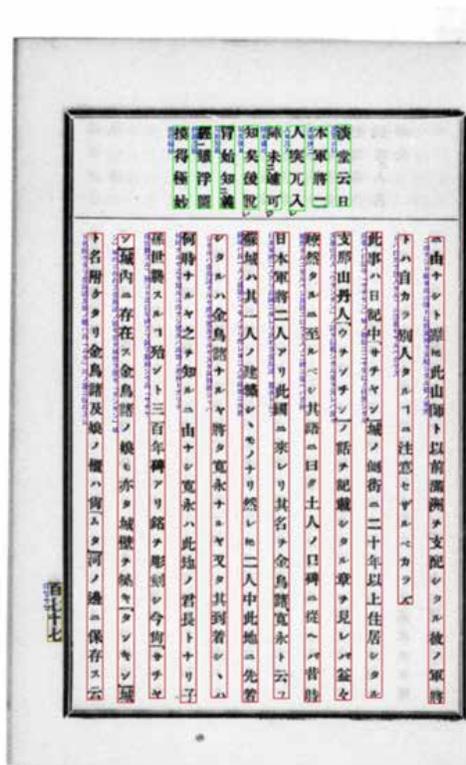
NDL OCR

https://lab.ndl.go.jp/data_set/ocr/r3_software/

旧字② (1909年以前の文系図書)



楠廷尉秘鑑：今古実録 第2編 中巻
<https://dl.ndl.go.jp/info:ndljp/pid/880917/13>



義経再興記
<https://dl.ndl.go.jp/info:ndljp/pid/782055/93>

1. 近代の活字を認識できる、オープンソースのOCR
2. 性能は高ければ90%以上に達するが、手書き文字は学習していないため、観測値の読み取りはできない
3. 処理負荷が高いため、全数OCRは困難

IIIF Curation Viewer

<http://codh.rois.ac.jp/software/iiif-curation-viewer/>

CODH開発



1. ' ' の「切り取り」ボタン → 四角領域を指定
2. ' ' の「お気に入り」ボタン → 欲しい画像を収集

キュレーションリスト

(1) <http://codh.rois.ac.jp/iiif/iiif-curation-viewer/?pages=200014778/8-9,24:200011824/4-6>



(3) × 全てクリア

(4) ↓ JSON

(5) 閉じる

メタデータ

IIIF Curation Viewer 日本語 / English ログイン

伊勢物語絵巻 (伊勢物語絵巻)

12 / 35

Curation Metadata (Canvas 2)

Metadata

label	value
テーマ	顔貌
性別	女
身分	貴族
向き	四分の三
原典	伊勢物語絵巻
DOI	10.20730/20
所蔵	

顔貌コレクション (顔コレ)

日本語 / English ログイン ヘルプ リスト

伊勢物語絵巻

戻る + リストに入れる

DOI: 10.20730/200024363

テーマ: 顔貌 性別: 女 身分: 貴族 向き: 四分の三 原典: 伊勢物語絵巻 DOI: 10.20730/200024363 所蔵: 国文学研究資料館

制作年: 制作年: 制作年: タグ:

コレクションとメタデータ

顔貌コレクション (顔コレ)

検索

テーマ / 性別 / 向き / 身分 / 原典 / 制作年 / タグ / 所属 / 京大貴重資料デジタルアーカイブレコードID / DOI / 慶応義塾大学請求記号

テーマへ
顔貌 (9683)

性別へ
男 (7406) / 女 (2228) / 不明 (49)

向きへ
四分の三 (5658) / 横 (3031) / 正面 (112) / 不明 (11) / 後 (868)

身分へ
貴族 (4556) / 武士 (3015) / 平民 (885) / 化身 (826) / 僧侶 (313) / 動物 (20) / 神様 (10) / 不明 (58)

原典へ
伊勢物語 (32) / 伊勢物語絵巻 (85) / 巖島縁起 (17) / 宇津保物語 (479) / 浦島太郎 (44) / 永ほしおりさつし (260) / 転寝草紙 (11) / 花鳥風月 (53) / 囃の草子 (15) / 成膳百 (104) / 妓王 (75) / 熊野権現縁起 (67) / 車巻絵巻 (53) / 御所まこと (35) / 五節調音之屏風絵 (180) / 西行物語 (86) / 狭衣 (21) / ささやき竹 (112) / [三十六]歌仙 (36) / 四十二の物あらそひ (153) / しつか (259) / 酒呑童子 (450) / しゃん童子 (172) / しゆてんとうし (359) / 水宮けい (144) / 住吉物語 (121) / 大織冠 (288) / 玉たすき (9) / 玉水物語 (17) / 依藤太 (148) / 長恨歌絵巻 (90) / 付喪神絵巻 (149) / 土くも (99) / 常盤の囃 (23) / ときはのうは (33) / ともなが (246) / 判官都ばなし (42) / 火おけのさつし (20) / ふくろふ (38) / 文正草子 (831) / ふんせう (147) /

江戸観光案内

検索

出典 / キーワード / 名所 (統一地名) / 名所 (原本表記) / 歴史地名データID / 江戸マップID / 江戸観光ID

出典へ
江戸名所図会 (759) / 絵本江戸土産 (247) / 江戸名所百人一首 (100) / 江戸名所記 (80) / 伊賀似顔東錦絵 (62) / 絵本江戸橋 (23) / 絵本江戸の見図 (21) / 絵本江戸橋 (17) /

キーワードへ
寺社 (775) / 水 (456) / 鳥居 (262) / 植栽 (153) / 橋 (104) / 歌舞伎 (67) / 坂 (63) / 富士山 (59) / 松 (51) / 桜 (48) / 祭 (48) / 花見 (43) / 伝説 (42) / 市場 (32) / 漁業 (30) / 食事 (29) / 墳墓 (22) / 宿場 (19) / 芸能 (18) / 武家屋敷 (17) / 農業 (17) / 門 (16) / 広場 (14) / 梅 (13) / 遊郭 (12) / 工芸 (12) / 武芸 (11) / 城 (10) / 紅葉 (8) / 五重塔 (6) / 富士塚 (6) / 藤樹 (5) / 衣料 (5) / 医療 (5) / 湯仏 (5) / 納涼 (4) / 稲木屋 (4) / 御旅所 (4) / 塩 (4) / 花火 (4) / 海苔 (3) / 塩業 (3) / 崖 (3) / 月見 (3) / 祭礼 (3) / 竹 (3) / 酒 (2) / 火除地 (2) / 潮干狩 (2) / 蕎麦 (2) / 炭 (2) / 御園 (2) / 蓮 (2) / 筑波山 (2) / 榎 (2) / 材木 (2) / 椎 (2) / 松浦家 (2) / 葛藤 (2) / 料理 (2) / 柳 (2) / 紙團会 (2) / 紅葉 (1) / 米 (1) / 大久保家 (1) / 出版 (1) / 鶴屋善右衛門 (1) / 山王祭 (1) / 海鼠 (1) / 鮎 (1) / 鷹場 (1) / 笹 (1) / 林業 (1) / 鮎 (1) / 三重塔 (1) / 桃 (1) / 杉 (1) / 卯花 (1) / 焼き米 (1) / 蓮見 (1) / 岡大師遺屋 (1) / 大黒組 (1) / 聴虫 (1) / 錆物 (1) / 橋枝 (1) / 般若講 (1) / 芝 (1) / 蕨業 (1) / 榎 (1) / 萩 (1) /

江戸買物案内

検索

出典 / 仲間 / 版面サイズ / 商人名 / 職種 (原本表記) / 居所 (原本表記) / 居所 (歴史地名大系) / 江戸マップID / 備考 / 江戸買物ID

出典へ
江戸買物控案内・問屋の部 (2368) / 江戸買物控案内・飲食の部3 (87) / 江戸買物控案内・飲食の部2 (18) / 江戸買物控案内・飲食の部1 (16) /

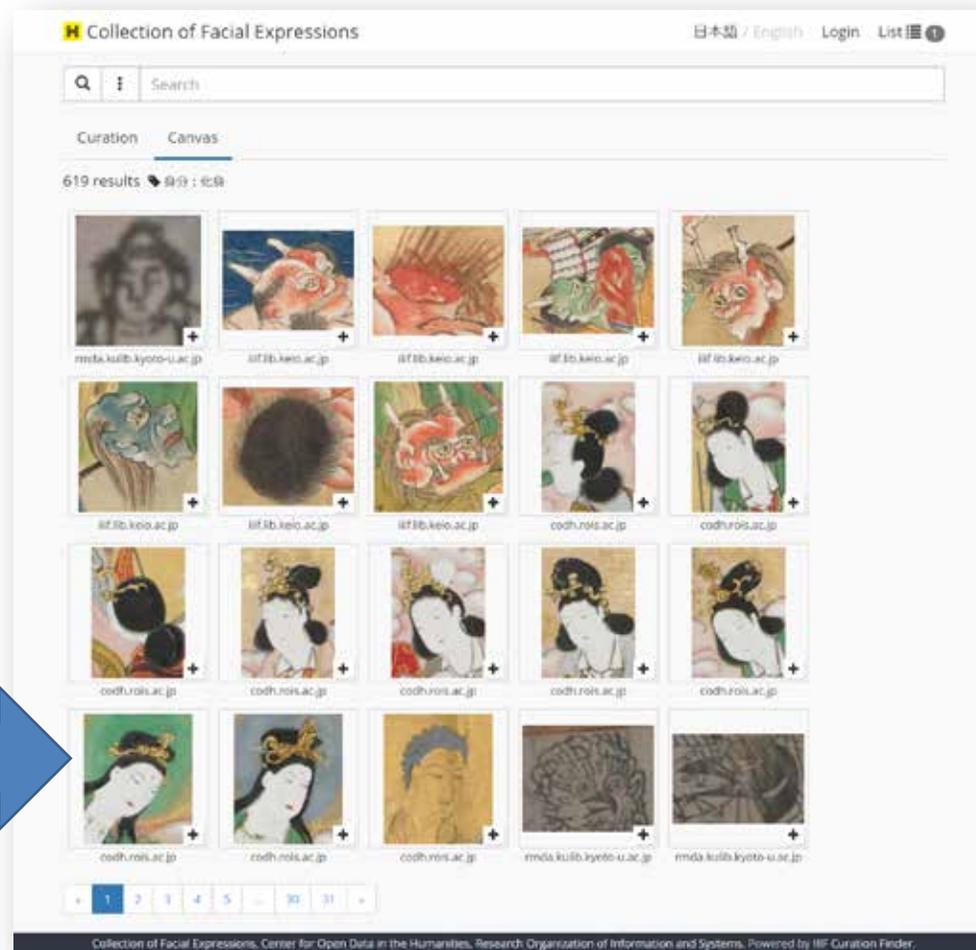
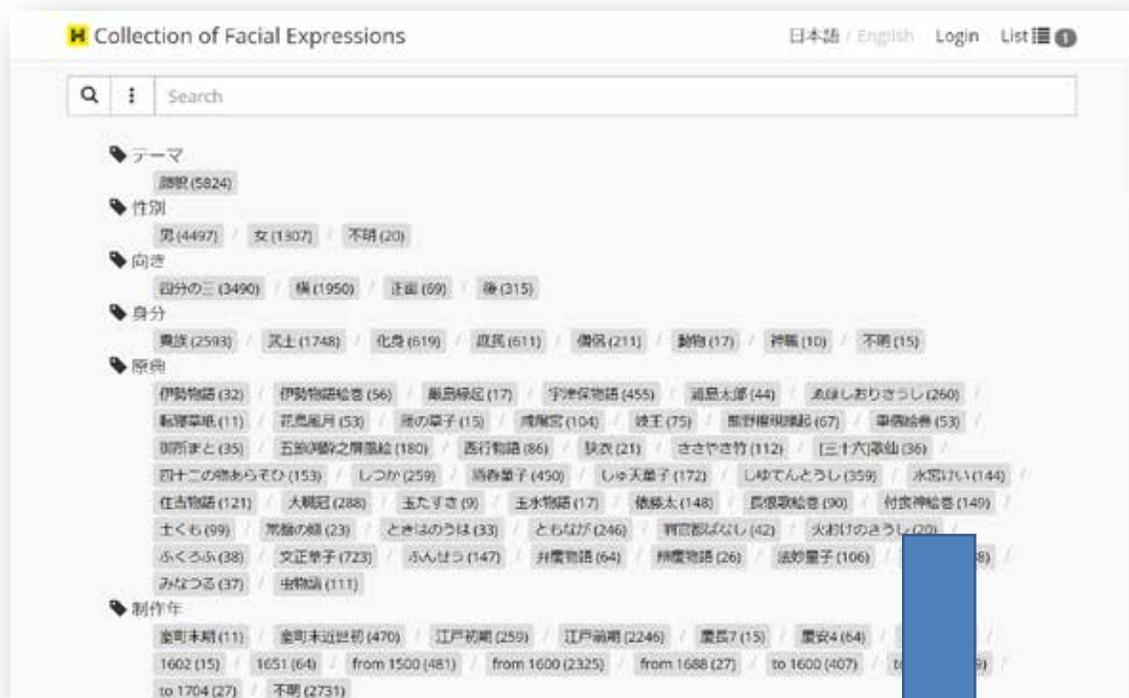
仲間へ
十組 (1122) / 丸合組 (60) / 釘店組 (44) / 傳馬町組 (24) / 浜吉組 (21) / 三十軒組 (16) / 通町組 (14) / 内店組 (14) / 二番組 (14) / 茅町組 (12) / 丸合古組 (12) / 元相本組 (3) / 本組 (3) / 三十組 (1) / 住吉組 (1) /

版面サイズへ
1/3 (1415) / 1/2 (469) / 1/4 (212) / 1/6 (123) / 1 (85) / 1/5 (71) / 2/3 (35) / 2 (32) / 4 (6) / 4/5 (3) / 1+1/2 (3) / 1+1/3 (2) / 1+1/5 (1) / 1+1/5 (1) / 3/4 (1) /

商人名へ
伊勢屋吉兵衛 (10) / 伊勢屋六右衛門 (7) / 助屋長左衛門 (7) / 恵比須屋八郎左衛門 (7) / 白木屋彦太郎 (6) / 森田屋半兵衛 (6) / 伊勢屋八兵衛 (5) / 柏屋孫左衛門 (5) / 日野屋善兵衛 (5) / 熊野屋作兵衛 (5) / 白子屋弥兵衛 (5) / 大坂屋孫八 (5) / 松本屋彦四郎 (5) / 多田屋新兵衛 (5) / 伊坂屋市右衛門 (5) / 伊勢屋半兵衛 (5) / 遠州屋半兵衛 (5) / 越後屋八郎兵衛 (4) / 鷹木屋七兵衛 (4) / 大黒屋三郎兵衛 (4) / 辻屋新兵衛 (4) / 鳩屋半兵衛 (4) / 木屋九兵衛 (4) / 小西長左衛門 (4) / 大坂屋庄左衛門 (4) / 伊勢屋平八 (4) / 伊勢屋宇兵衛 (4) / 岸部屋安兵衛 (4) / 中村屋伊兵衛 (4) / 近江屋小兵衛 (4) / 伊勢屋徳三郎 (4) / 白子屋勘兵衛 (4) /

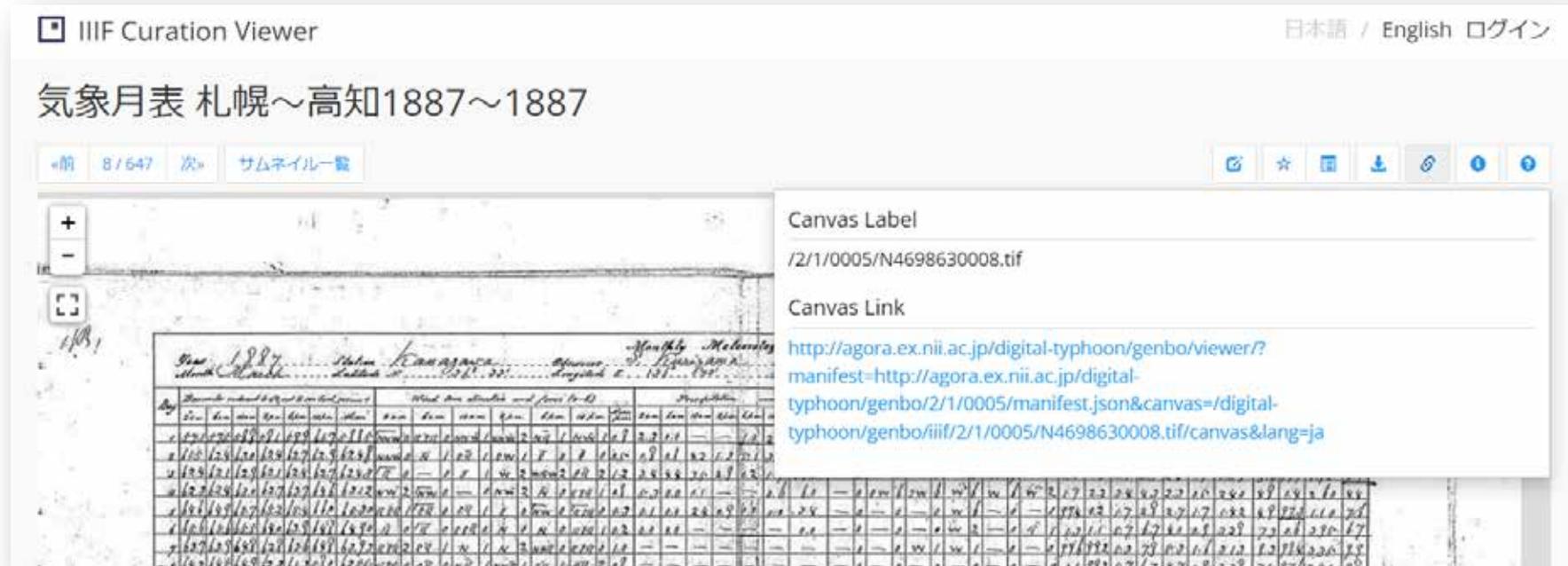
メタデータによるファセット検索

<http://codh.rois.ac.jp/face/iiif-curation-finder/>



個々の画像にメタデータを付与すると、メタデータの値ごとに画像を検索できる

課題3：画像とデータのリンク



1. IIF Curation ViewerのCanvas Link機能を用いることで、**データから画像をリンク**することは可能
2. **画像からデータをリンク**するには別の方法が必要

課題4：気象アーカイブとのリンク

<http://agora.ex.nii.ac.jp/digital-typhoon/weather-chart/>



1. 天気図（1883-）と気象原簿（1878-）は、ほぼ同時期に始まっている
2. 両者をリンクすることで、マクロな総観気象とミクロな観測値とを比較できる
3. 個別画像にメタデータを付与する、またはデータ公開サイトを作るといった方法がある

まとめ

1. 地上気象観測原簿アーカイブの構築方法と現状を紹介した
2. 回転やメタデータなどの課題解決には、III技術をどう活用するかという視点が不可欠
3. OCRや機械学習を活用する余地もあるが、効果が明確に見通せず、着手する価値があるか不透明
4. 活用の拡大には、個別画像のメタデータ付与、機械可読化したデータの公開サイトが課題ではないか？