

大量の板本を効率的に翻刻する ～ 「武鑑全集」の経時的差分翻刻による 江戸200年参勤交代データセットの構築



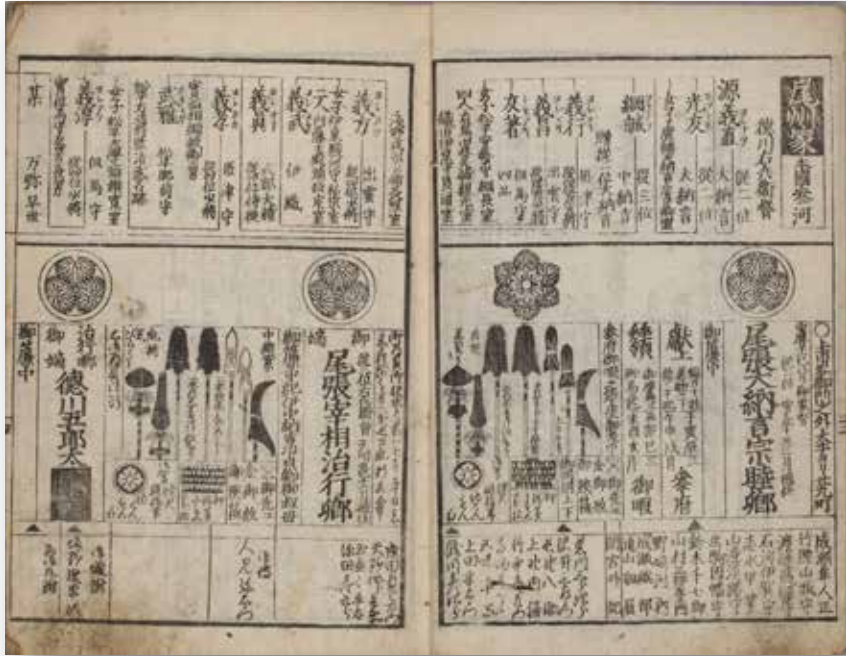
北本 朝展 (Asanobu KITAMOTO)

情報・システム研究機構 データサイエンス
共同利用基盤施設 (ROIS-DS) 人文学オープン
ンデータ共同利用センター(CODH)

国立情報学研究所

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

「武鑑」とは？



寛政武鑑 (1789)、日本古典籍データ
セット (国文研蔵、CODH公開)

<http://codh.rois.ac.jp/pmjt/book/200018823/>

1. **データブック**：江戸の大名や幕臣に関するリッチな構造化データを記録（紋などの「絵」も含む）
2. **200年**：江戸時代初期から末期まで出版され続けたベストセラー
3. **多目的**：役職移動情報、江戸観光みやげ、大名行列観覧ガイド等
4. **高頻度更新**：情報の更新は最大で1月に数回

参考：藤實 久美子、江戸の武家名鑑 武鑑と出版競争、吉川弘文館、2008

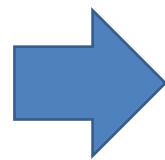
武鑑全集

日本古典籍データセットの武鑑381点

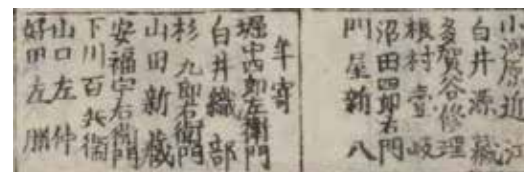
<http://codh.rois.ac.jp/bukan/>

日本古典籍
データセット
(国文研蔵)

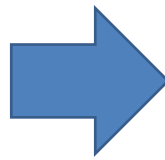
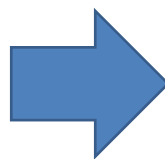
データ駆動型手法を
用いた武鑑の経時的
解析



参勤交代の可視化



大名家・幕臣の人物
データベースの構築



江戸デザインの分析

人手で大名家に関
する構造化データ
を作成

- 居城
- 参勤交代
- 上屋敷
- 菩提寺
- 紋 / 道具
- 献上品 / 拝領品



経時的差分翻刻のアイデア

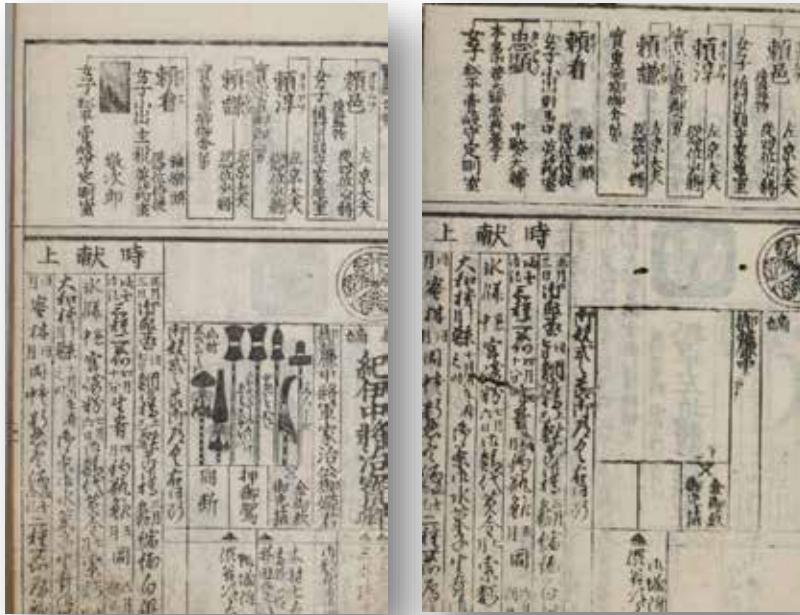
問い：200年にわたる史料をどう翻刻するか？



解決：前後の版本で変化した部分のみを翻刻し経時的データを構築する



木版印刷と板本書誌学



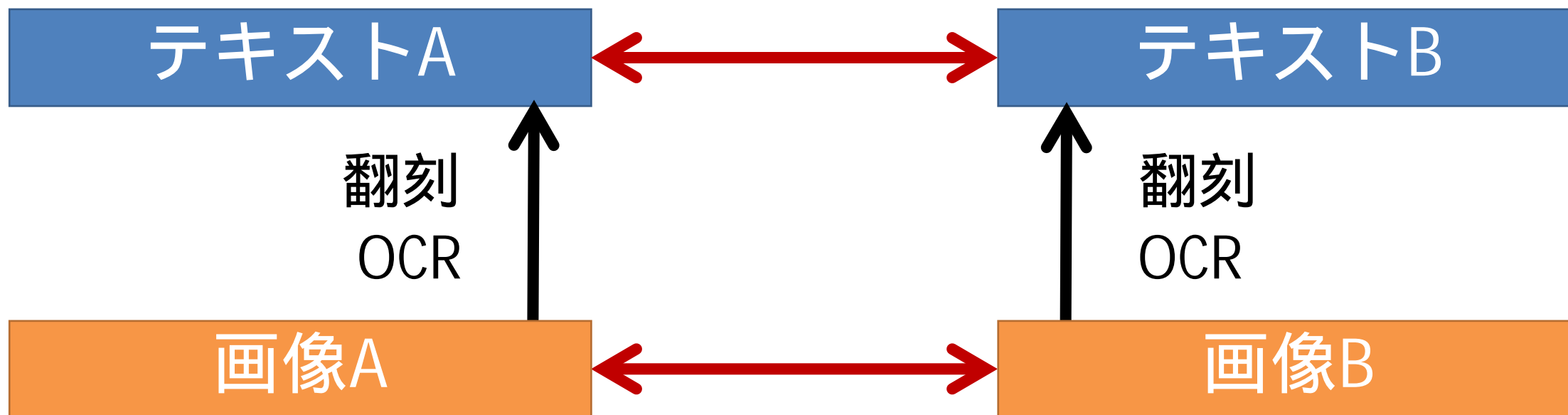
左：寛政武鑑（1789）
右：寛政武鑑（1791）

江戸時代の出版は**木版印刷が主流**
（活字もあるが限定的）

1. **刊**：板木を最初から作り直す（メジャーバージョン）
2. **印**：同一の板木から多くの印刷を行う（インスタンス）
3. **修**：板木を埋木で修正する（マイナーバージョン） = **細部の変化**

テキスト照合と画像照合

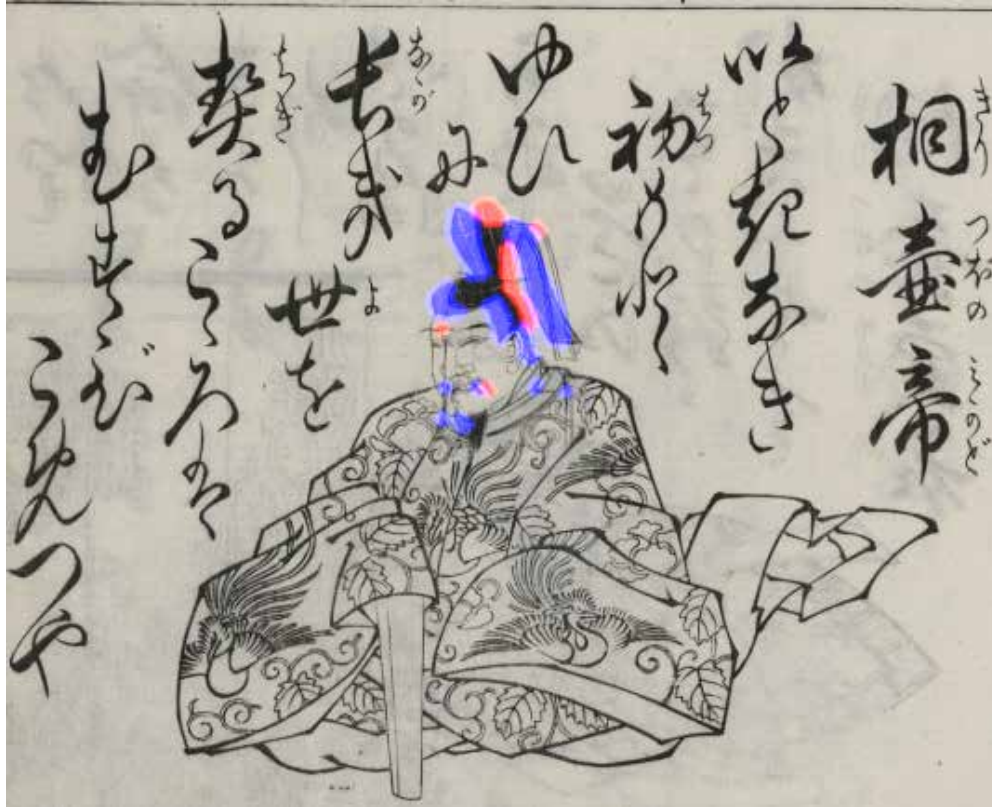
テキスト照合 = 多くのツールが存在



画像照合 = **標準的なツールがない**
(目視による「**並列比較 side-by-side**」が主流)

画像照合と差読（Differential Reading）

<http://codh.rois.ac.jp/differential-reading/>



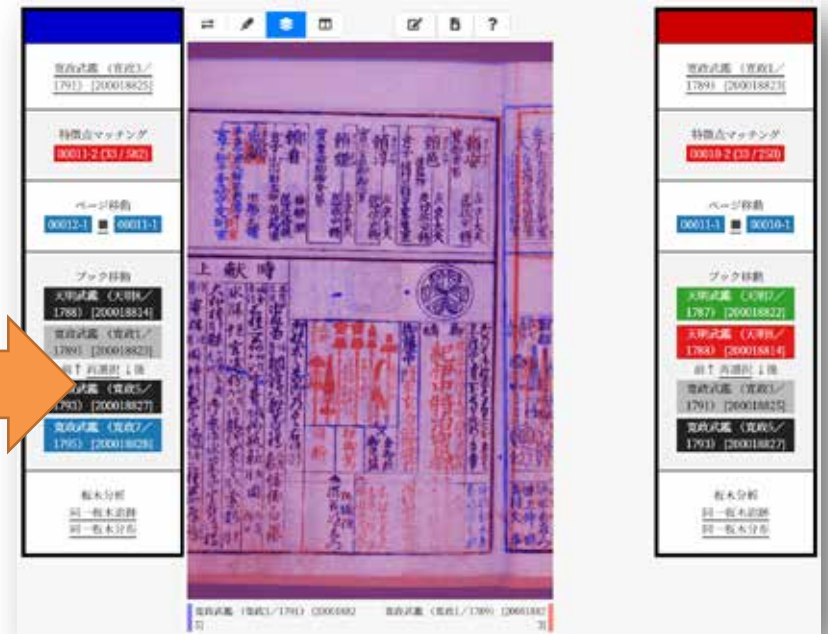
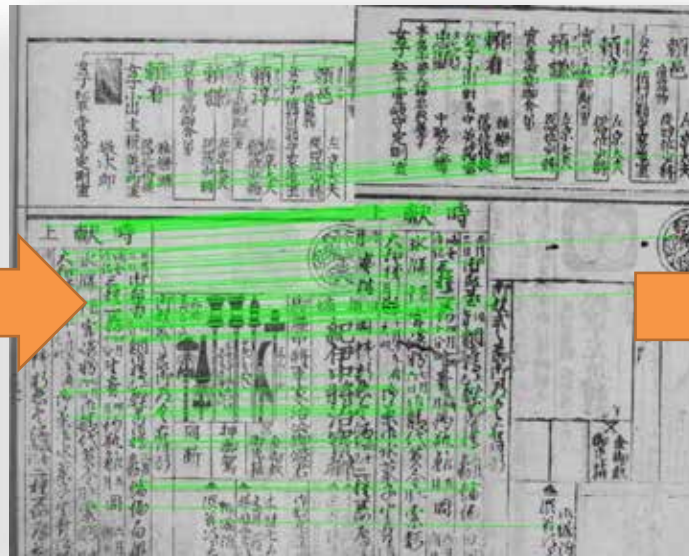
源氏百人一首（パタパタ顔比較）、東京大学総合図書館

1. 2枚の画像を照合した結果を表示・共有する
ビューア **vdiff.js** を構築
 - ウェブ版 = URL を指定
 - ファイル版 = ローカルフォルダを指定
2. 「2冊の差分を読む」**差読サービス**として公開

武鑑全集と「差読プラットフォーム」

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

多数の版から差分を取り出すプラットフォーム



武鑑の網羅的なデジタル化

コンピュータビジョン技術を用いた画像マッチング

版間の差分を強調するビューと様々な関連機能

ブック照合（ページペア表示）

入替	寛政武鑑（寛政3/1791）[200018825]					寛政武鑑（寛政1/1789）[200018823]				
ページリスト	00000-1	00000-2	00001-1	00001-2	00002-1	00000-1	00000-2	00001-1	00001-2	00002-1
	00002-2	00003-1	00003-2	00004-1	00004-2	00002-2	00003-1	00003-2	00004-1	00004-2
	00005-1	00005-2	00006-1	00006-2	00007-1	00005-1	00005-2	00006-1	00006-2	00007-1
	00007-2	00008-1	00008-2	00009-1	00009-2	00007-2	00008-1	00008-2	00009-1	00009-2
	00010-1	00010-2	00011-1	00011-2	00012-1	00010-1	00010-2	00011-1	00011-2	00012-1
	00012-2	00013-1	00013-2	00014-1	00014-2	00012-2	00013-1	00013-2	00014-1	00014-2
	00015-1	00015-2	00016-1	00016-2	00017-1	00015-1	00015-2	00016-1	00016-2	00017-1
	00017-2	00018-1	00018-2	00019-1	00019-2	00017-2	00018-1	00018-2	00019-1	00019-2
	00020-1	00020-2	00021-1	00021-2	00022-1	00020-1	00020-2	00021-1	00021-2	00022-1
	00022-2	00023-1	00023-2	00024-1	00024-2	00022-2	00023-1	00023-2	00024-1	00024-2
	00025-1	00025-2	00026-1	00026-2	00027-1	00025-1	00025-2	00026-1	00026-2	00027-1
	00027-2	00028-1	00028-2	00029-1	00029-2	00027-2	00028-1	00028-2	00029-1	00029-2
	00030-1	00030-2	00031-1	00031-2	00032-1	00030-1	00030-2	00031-1	00031-2	00032-1
	00032-2	00033-1	00033-2	00034-1	00034-2	00032-2	00033-1	00033-2	00034-1	00034-2

1. 2点の本を選び、ページ照合の結果を一覧
2. 青 > 緑 > 赤 > 黒の順番で、ページ照合の推定品質が高い
3. 灰はページ照合に失敗した場合で、落丁・増丁などに相当
4. 紫は人手で照合済み

ページ照合表示

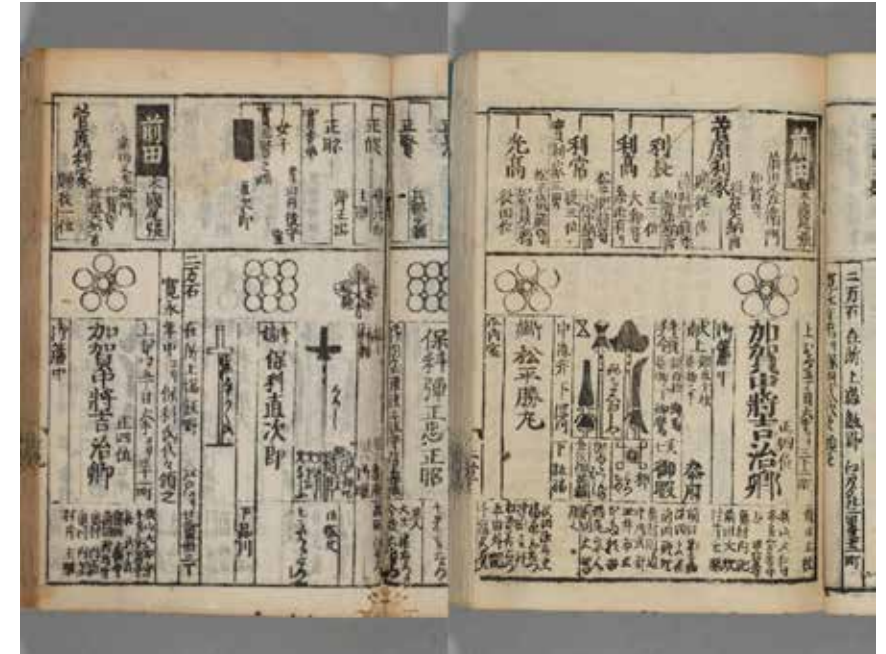
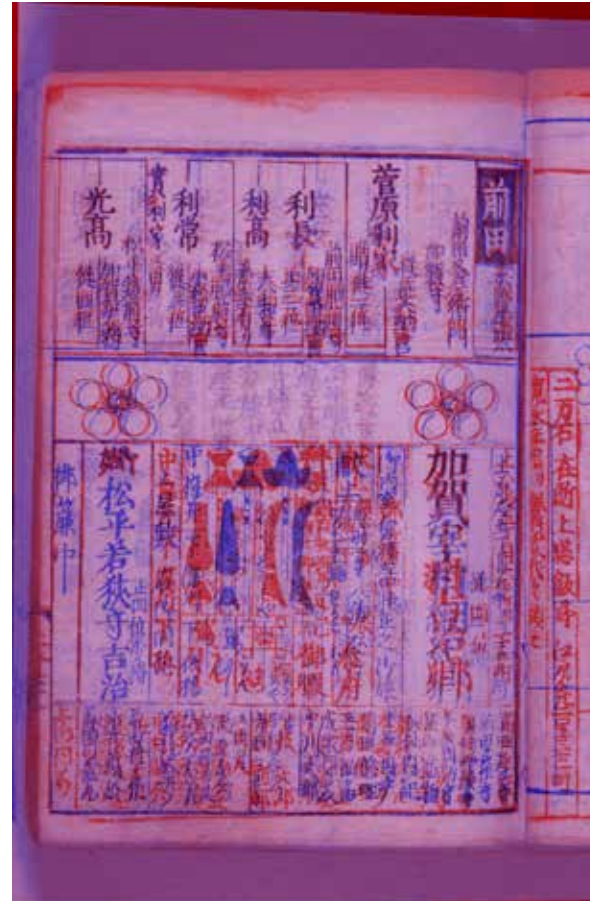
1. 2ページの画像を読み込み、比較結果を **vdiff.js** で表示
2. ページ移動、ブック移動、板木分析などの機能を提供
3. 最も重要な機能：**同一板木を保持しながら、前後のブックに移動する機能**

The screenshot displays a digital manuscript comparison interface. At the top, there are navigation icons for back, forward, search, and help. The main area is divided into three vertical panels. The left panel shows a list of manuscript entries with their respective page numbers and a 'ページ移動' (Page Move) button. The right panel shows a similar list with a 'ブック移動' (Book Move) button. The central panel displays a large image of a manuscript page with red and blue annotations. An orange arrow points from the text in the third list item to the 'ブック移動' button in the right panel. The bottom of the interface shows the current page numbers and book identifiers for both pages being compared.

Page 1 (Left)	Page 2 (Right)
寛政武鑑 (寛政3/1791) [200018825]	寛政武鑑 (寛政1/1789) [200018823]
特徴点マッチング 00011-2 (33 / 582)	特徴点マッチング 00010-2 (33 / 250)
ページ移動 00012-1 ■ 00011-1	ページ移動 00011-1 ■ 00010-1
ブック移動 天明武鑑 (天明8/1788) [200018814] 寛政武鑑 (寛政1/1789) [200018823] 前↑再選択↓後 寛政武鑑 (寛政5/1793) [200018827] 寛政武鑑 (寛政7/1795) [200018828]	ブック移動 天明武鑑 (天明7/1787) [200018822] 天明武鑑 (天明8/1788) [200018814] 前↑再選択↓後 寛政武鑑 (寛政3/1791) [200018825] 寛政武鑑 (寛政5/1793) [200018827]
板木分析 同一板木追跡 同一板木分布	板木分析 同一板木追跡 同一板木分布

寛政武鑑 (寛政3/1791) [20001882] 51 | 寛政武鑑 (寛政1/1789) [20001882] 31

ブック移動の例



同一板木でないのでマッチングできない

同一板木なのでマッチングできる

同一板木ではないがマッチングできる

江戸200年参勤交代データセット

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

1. 「武鑑全集」の中から、**大名家に関するデータを含む武鑑354点（1658-1868）**を選択
2. 参勤交代時期を表す「**参府**」「**御暇**」、参考情報として「**当主**」「**居城地**」を経時的に差分翻刻
3. **経時的な情報の連続性**に基づき、情報が不自然に変化する箇所を特定
4. 情報の連続性を向上させるため、**出版順の入れ替え**や**書誌情報の妥当性**を検討

	A	D	E	F	G	I	J	K	L	M
1		武鑑基礎情報				翻刻				
2	番号	武鑑名	DOI	出版年 (和暦)	出版年 (西暦)	記載ページ	当主名	参府年月	御暇年月	居城地
65	63	正徳武鑑	200018763	正徳4		2-2	加賀宰相綱紀卿	なし		
66	64	正徳武鑑	200018764	正徳5		3-2	加賀宰相綱紀卿	なし		
67	65	享保武鑑	200018765	享保2	1717	00022-2	加賀宰相綱紀卿	なし		
68	66	享保武鑑	200018766	[享保4]	1719	00022-2	加賀中将吉治卿 (吉徳)	辰9		
69	67	享保武鑑	200018768	享保6	1721	00023-2	加賀宰相綱紀卿	なし	なし	加州金沢
70	68	享保武鑑	200018769	享保11	1726	00024-2	加賀中将吉治卿 (吉徳)	午9	巳3	加州金沢
71	69	享保武鑑	200018770	享保14	1729	00024-2	加賀中将吉治卿 (吉徳)	午9	巳3	加州金沢
72	70	享保武鑑	200018771	享保17	1732	00022-2	加賀中将吉治卿 (吉徳)	なし	亥7	加州金沢
73	71	元文武鑑	200018772	元文1	1736	00020-2	加賀中将吉治卿 (吉徳)	辰7	巳7	加州金沢
74	72	元文武鑑	200018773	元文5	1740	00021-2	加賀中将吉治卿 (吉徳)	申7	未7	加州金沢
75	73	寛保武鑑	200018774	寛保1	1741	00021-2	加賀宰相吉徳卿	申7	未7	加州金沢

おそらく間違い

66は67より後、73より前に入るはず

板木の連続性を根拠に書誌情報を精査

1. 板木の変更はコストが高いため、コストを最小化するには**板木を連続して使う**必要がある
2. 出版年順で情報が連続しない場合は、**出版年の誤り**（または異なる出版年の板本の混在）を示唆する
3. 多くの大名家の翻刻で生じる矛盾を解消することで、**武鑑の出版順を確定**させる
4. **出版順に基づき出版年を推定し直す**ことで、より精度の高い書誌情報に更新する

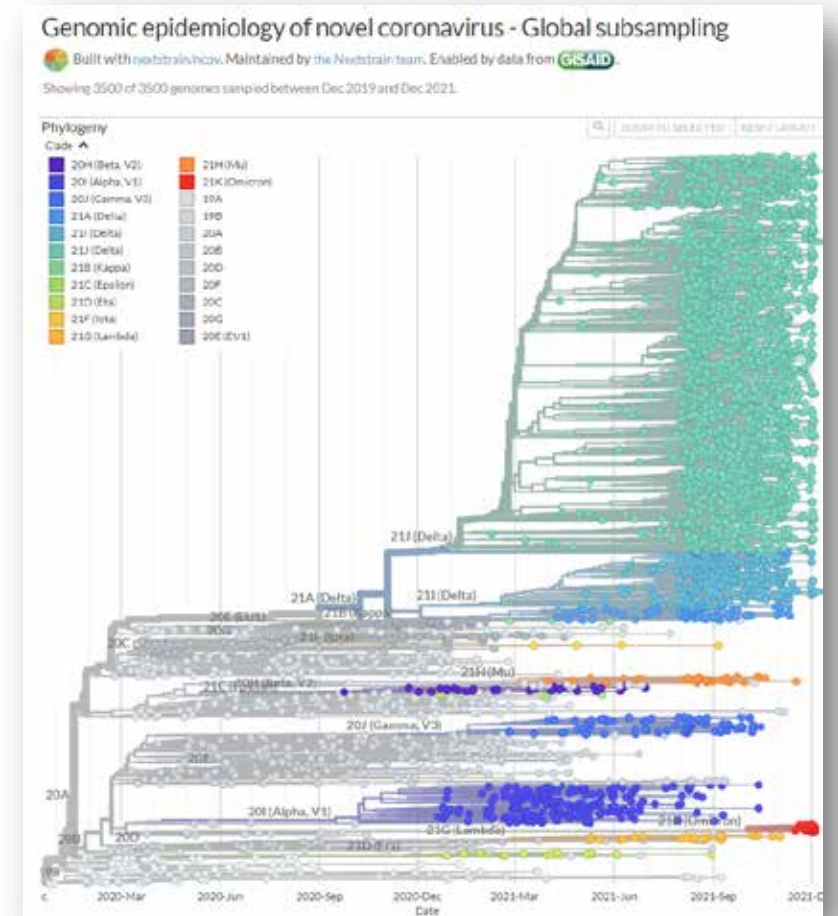
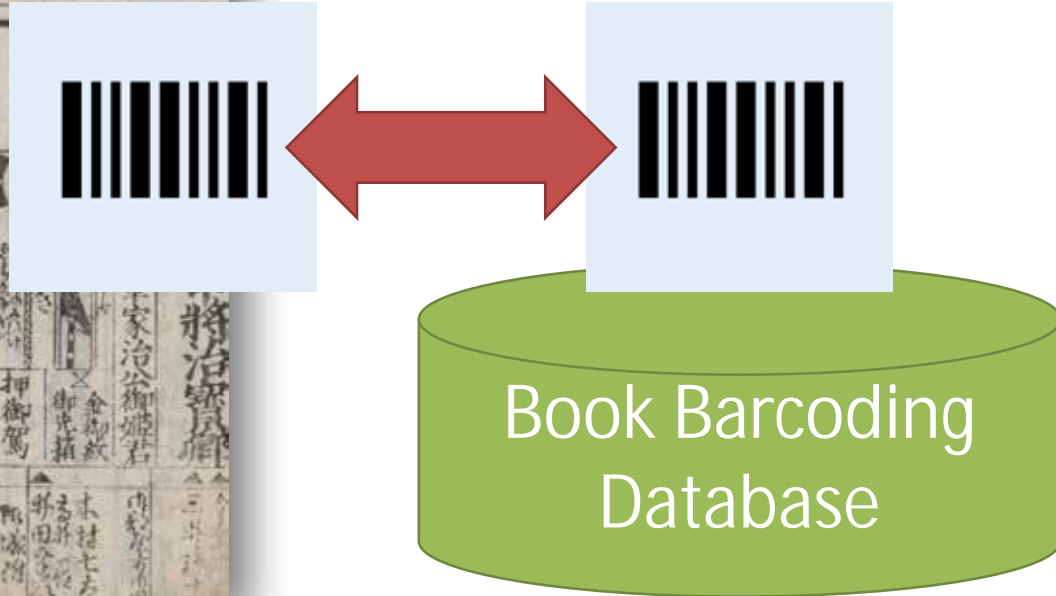
経時的差分翻刻の効率化と拡大

1. すでに尾張、加賀、薩摩の3家を翻刻済み。さらに紀州を現在翻刻中
2. 1家1冊1分（1家全冊1日）ペースでの差分翻刻の実現に向けて、ワークフローの効率化を試行中
3. 予定：効率化したワークフローに基づき、主要な大名（40-50程度）の翻刻を完了させ、出版順を確定
4. 予定：全家を対象とした参勤交代データセットや、その他の構造化データセットを作成

将来像：板本比較研究プラットフォーム



未知の板本のバーコードを計算し、データベースと比較



<https://nextstrain.org/ncov/gisaid/global>

謝辞

1. 「武鑑全集」の翻刻やシステム改良には、**合同会社AMANE（小川歩美氏、寺尾承子氏）**に協力いただきました。
2. 「武鑑全集」の学術的検討には、**国文学研究資料館の藤實久美子教授**に協力いただきました。
3. 「武鑑全集」の画像データには、**日本古典籍データセット（国文学研究資料館蔵）**を活用しています。
4. 「武鑑全集」で活用するvdiff.jsおよびvdiff-seq.jsの開発には、**本間淳氏（フェリックス・スタイル）**がコア開発者として貢献しました。