



session 1 くずし字初歩





自己紹介：山本和明 (YAMAMOTO Kazuaki)



1962年生

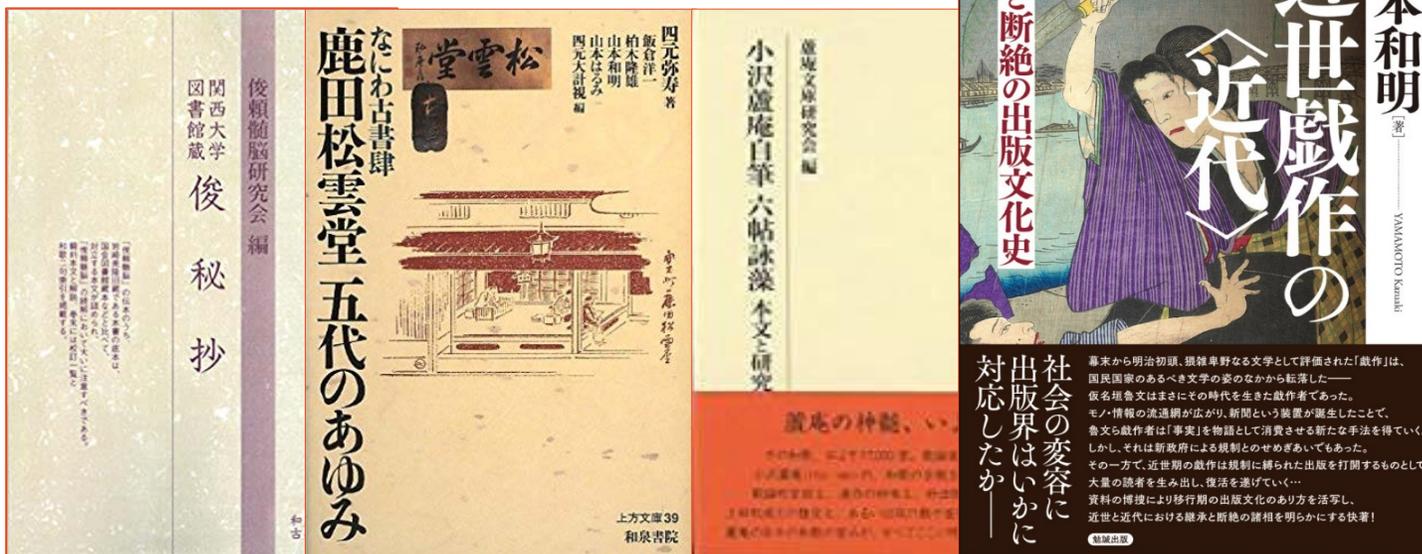
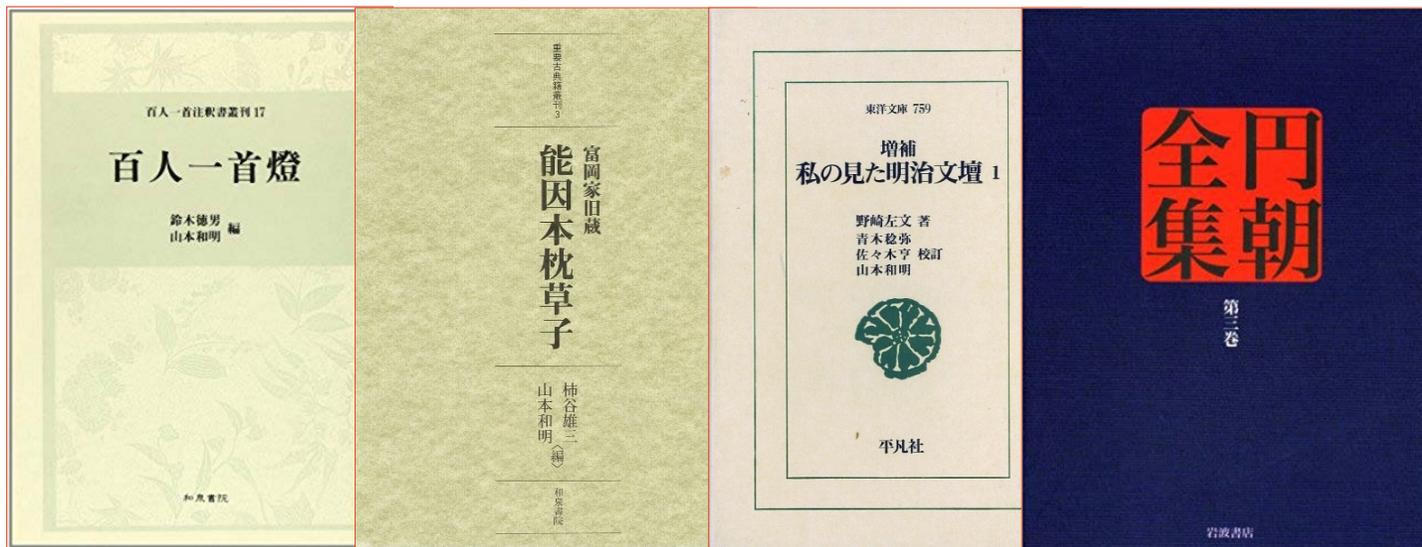
大阪の茨木高校、神戸大文学部を経て、
神戸大学大学院文化学研究科文化構造
学専攻単位修得退学、博士（文学）

大阪の相愛大学教授等を経て2013年10
月国文研に着任

歴史的典籍NW事業の推進に従事

専門：19世紀日本文学

加藤千蔭・山東京伝・仮名垣魯文等の
研究（とはいえもろもろ）

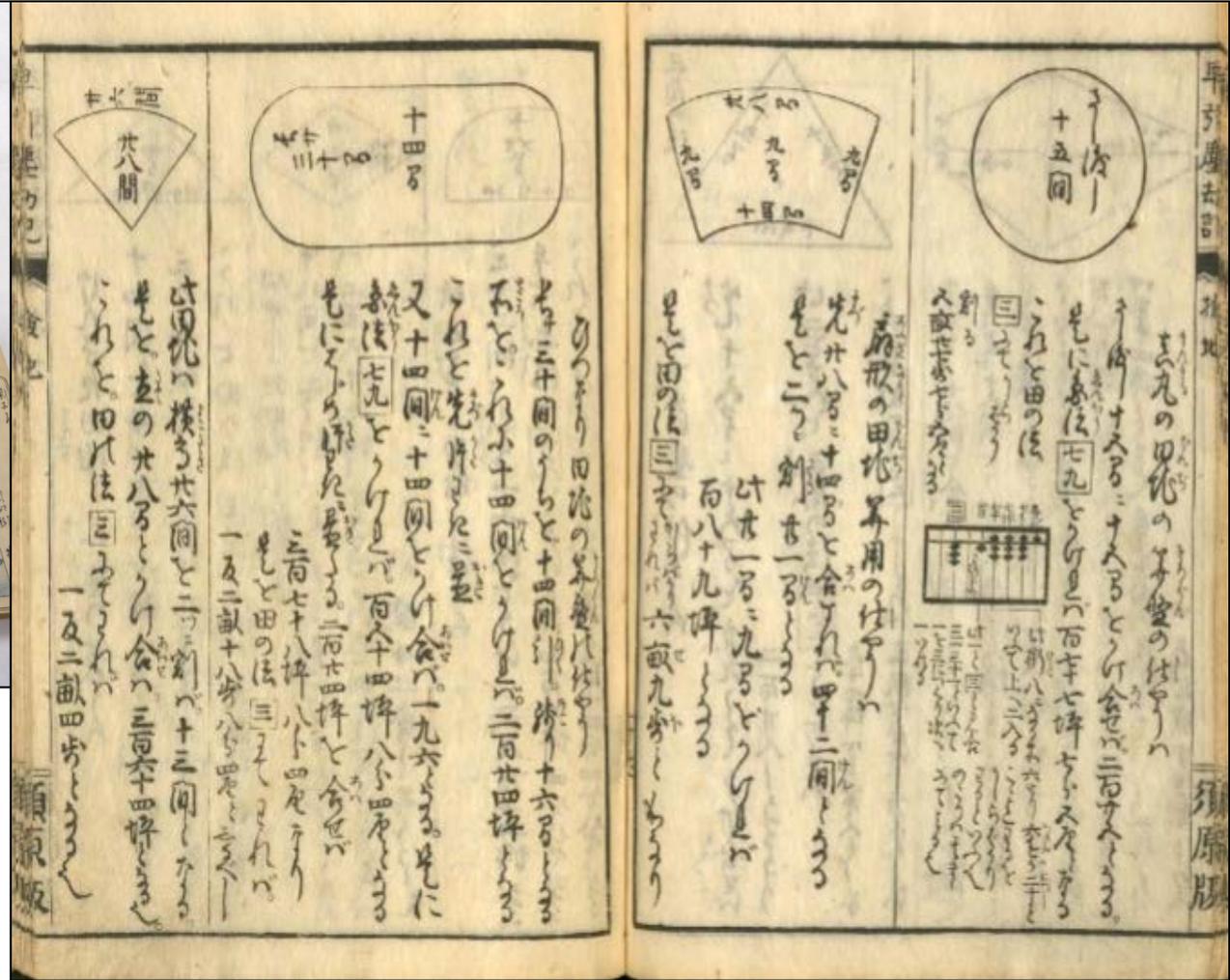




古典籍のあれこれ



むかしの数学の本
(早引塵劫記大成)



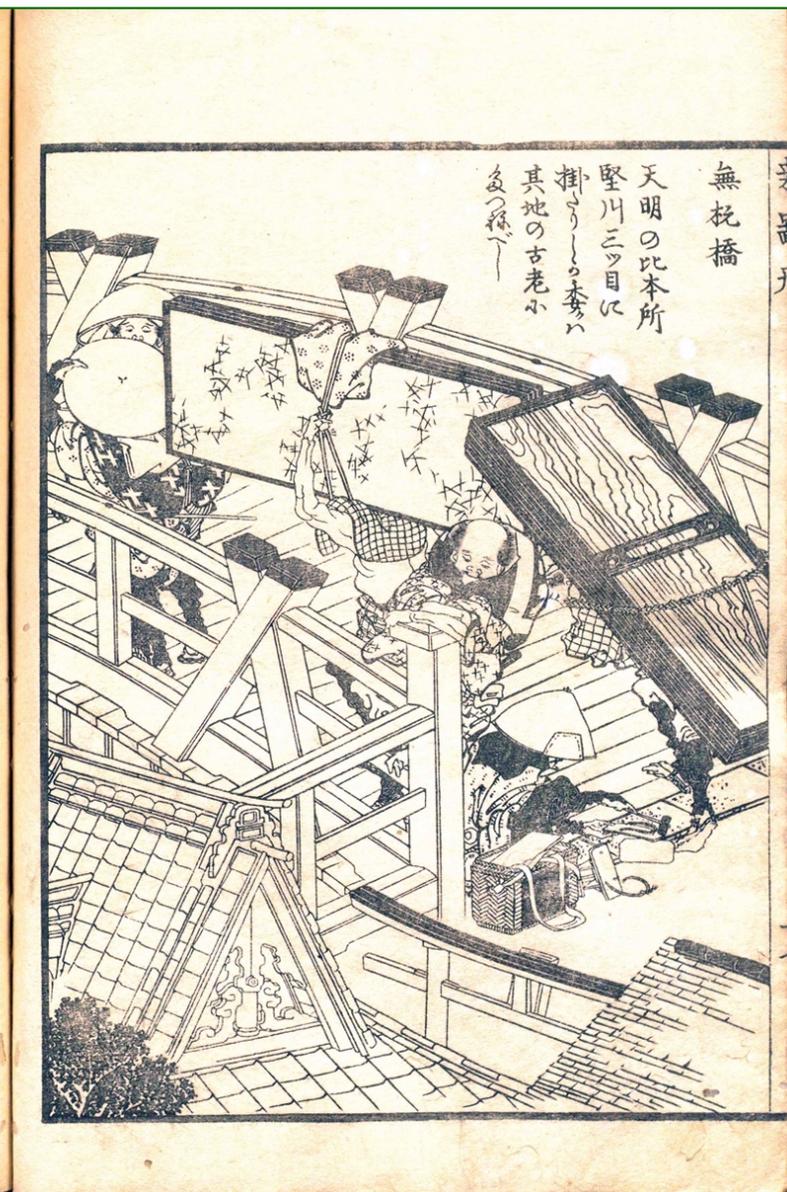
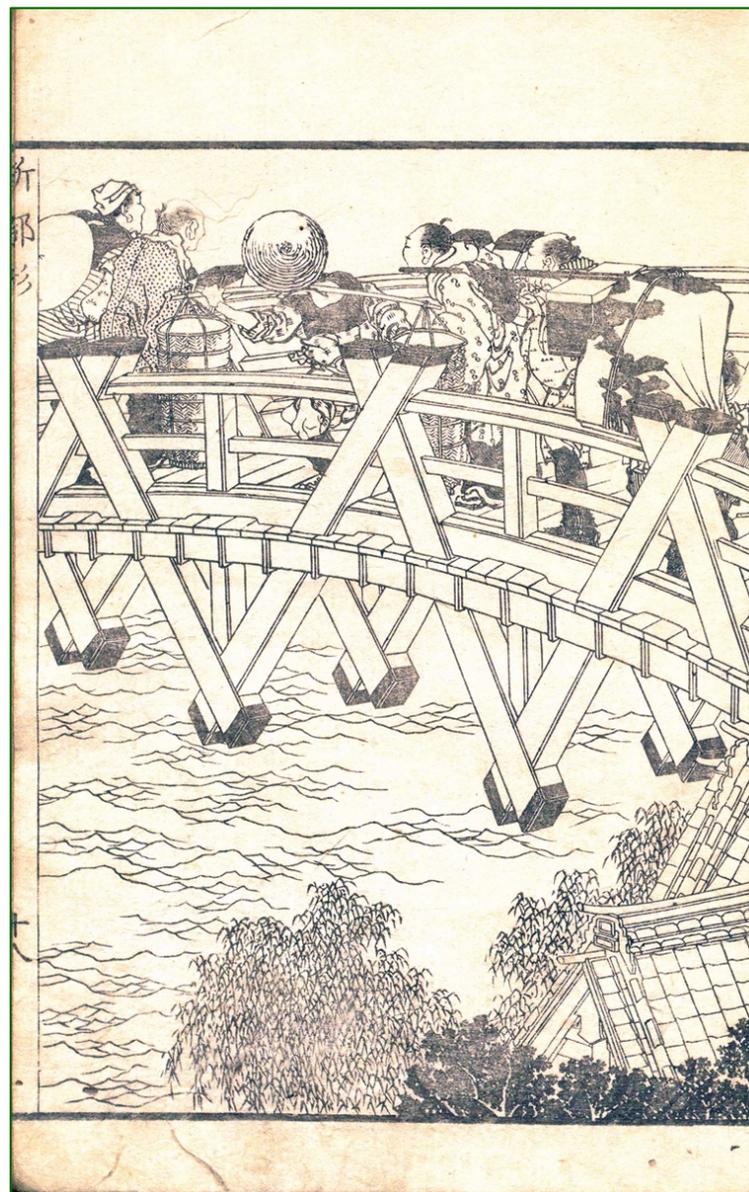
あらゆる情報が古典籍のなかには詰まっている



古典籍のあれこれ

『新鄙形 (しんひながた)』 (諸職絵本 / 葛飾新鄙形)

画を志す人の便をはかるため、神社仏閣の写法などを示した書





古典籍のあれこれ

『源氏ひながた』 織物の図案集

源氏物語などを題材にした図案など載せ、「桐つほのかうゐ」「若むらさき」などにつき説明を付す。





古典籍のあれこれ

『大和本草』 貝原益軒編纂の 本草書



大和本草

諸品圖
五

482
24
国文学研究資料館

蚊子木^{ウズノキ} 其枝ニ如木實内ニ
物ナク空虚ニシテ如瓢ナル物
アリ俗猿瓢ト云大ナルハ周五
寸五分許又小ナルアリ是眞實ニ
アス或其内ニ
腰蜂アリ
別ニ眞實アリ

臘梅^{ウツギ} 其葉繁茂兩々相對
有光高不過四尺枝葉
繁多如荊至冬葉落臘月
發小黃花不
美其臭如
蘭

常山樹高不過數尺其枝
繁茂有惡臭其花細小青
白色不美其葉有光與臘
梅相似
臘梅葉ハ兩々
相對ス

ヤレホヤドリ木也

葉相似無光有惡臭



古典籍のあれこれ

画本虫撰 (狂歌絵本)

1788年に蔦屋重三郎より
喜多川歌麿画で出た狂歌集



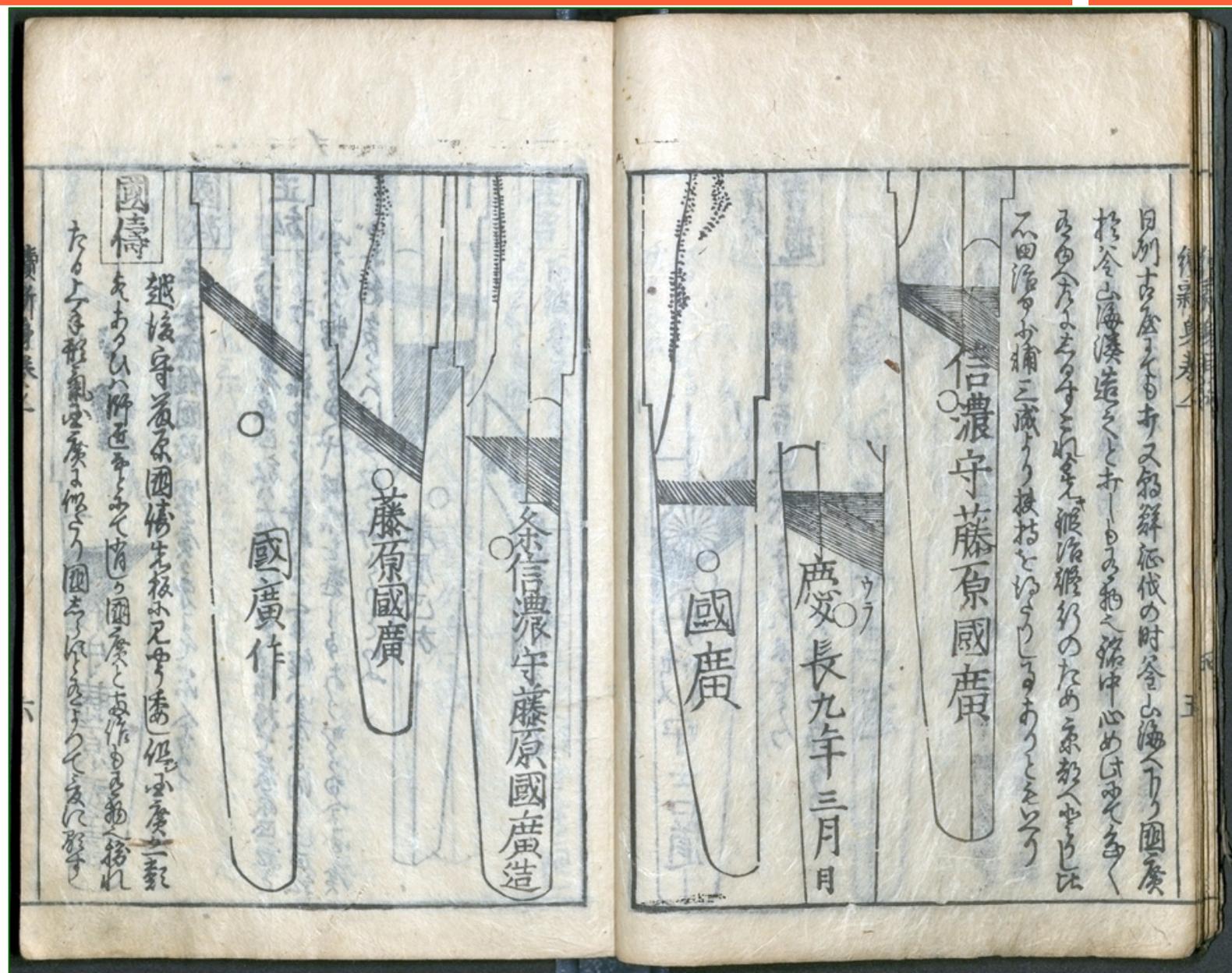


古典籍のあれこれ

『新刃銘尽後集』

(あらみめいづくしこうしゅう)

国ごとに名刀といわれる刀剣を絵入りで紹介したもの。刀剣を好む者や刀剣を商う人の一助としてまとめられた。



國傳

越後守藤原國傳生板小兄也(倭)但少廣之親
夫あついは解通平とて清の國廣と云作も多物へ授れ
たの上形氣は廣より作り國をいふをうつてまに形す

國廣作

○藤原國廣

○余信濃守藤原國廣造

○信濃守藤原國廣

○國廣

廣^{ウラ}長九年三月月

日刃十の盛とも亦又新羅征伐の時谷山海より國廣
於谷山海渡造ると云一も其物之銘中心めけりて多く
るゆゑたよるすそれをも彼流渡りのため東都へせりは
不田流や少補三藏より技持とゆふしるありとをう



古典籍のあれこれ

『紙漉重宝記』

石州半紙の製法を絵入りで紹介したもの。詳細な挿絵が多いので江戸時代の製法を知る上で貴重な資料である。





古典籍のあれこれ

『平治物語絵巻』 卷子本



中世では、戦乱のたびに語り伝えられた英雄伝などが記録され、軍記物語が誕生しますが、鎌倉初期の『平治物語』などは、和漢混交文で生き生きと武将たちの活躍が描かれており、その絵巻物です。



古典籍のあれこれ

『豆腐百珍続編』

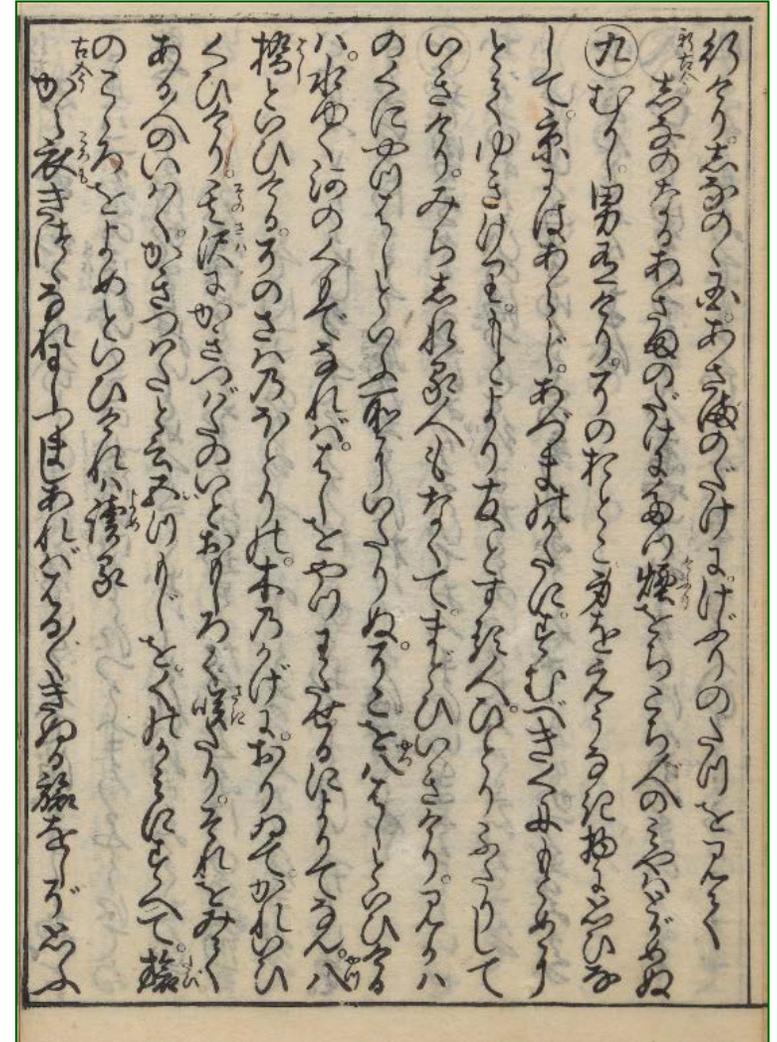
豆腐の料理法100種類を解説した料理本。豆腐料理を尋常品・通品・佳品・奇品・妙品・絶品に分類して説明

右図は豆腐田楽（石川県）と田楽法師の構図が一致しているという見立て





なぜ読めないか？



読める文字／読めない文字

伊勢物語第九段



なぜ読めないか？

- 読もうとしない 本が汚い！！／他に興味有り
- そもそも本を読まない
- 読みたくても読めない文字だから
⇒「くずし字」の問題とされがち



技術の進展とともに情報の淘汰がなされ、
画一化していった
(だからこそ識字率が高いとも言える)



なぜ読めないか？

読めないのには
訳がある

字形は
固定していない





読めないのには訳がある

⇒ 異体字を知らないから

○くずし字：くずし書きにした文字（人それぞれ）

○変体仮名（異体字）：かつては1音節に種々の字体あり

平仮名	片仮名	平仮名	片仮名
あいうえお	アイウエオ	らりるれろ	ラリルレロ
かきくけこ	カキクケコ	わゐうゑを	ワヰウヱヲ
さしすせそ	サシスセソ	ん	ン
たちつてと	タチツテト	がぎぐげご	ガギグゲゴ
なにぬねの	ナニヌネノ	ざじずぜぞ	ザジズゼゾ
はひふへほ	ハヒフヘホ	だぢづでど	ダチヅデド
まみむめも	マミムメモ	ばびぶべぼ	バビブベボ
やいゆえよ	ヤユエヨ	ぱぴぷぺぽ	パピプペポ

小學校令施行規則

（しょうがっこうれいしこうきそく）

- 法令番号：明治33年文部省令第14号
- 公布：1900年（明治33年）8月21日

音と文字の対応関係 1対多 ⇒ 1対1へ



異体字の存在

字母

丹

爾

仁

に

「平仮名」「カタカナ」も漢字の崩れた文字に他ならない。

くずし字だから読めないのではない！
食わず嫌い??

丹

丹

丹

爾

仁

仁

丹

丹

丹

爾

仁

仁

丹

丹

爾

仁

仁

ひらがな：平安初期に成立した日本独特の音節文字。それぞれの字源は万葉仮名に用いた漢字でその草書体を一層簡素化したもの。女手。

カタカナ：かたは完全で無い、一部分の意味。万葉仮名を簡略化して用いたところに発する。略体仮名。



あ あ (安) 愛 (愛) 阿 (阿) 悪 (悪)

い い (以) 伊 (伊) 意 (意) 移 (移)

う う (宇) 宇 (宇) 憂 (憂) 有 (有) 雲 (雲)

え え (江) 盈 (盈) 縁 (縁) 衣 (衣) 衣 (衣) 要 (要)

お お (於) 於 (於) 隠 (隠)

か 佳 (佳) 加 (加) 可 (可) 可 (可) 嘉 (嘉) 我 (我) 歎 (歎) 賀 (賀) 閑 (閑) 香 (香)
駕 (駕) 家 (家)

き 喜 (喜) 幾 (幾) 幾 (幾) 支 (支) 木 (木) 祈 (祈) 貴 (貴) 起 (起) 期 (期)

く 久 (久) 久 (久) 九 (九) 供 (供) 俱 (俱) 具 (具) 求 (求)

け 介 (介) 介 (介) 希 (希) 氣 (氣) 計 (計) 遣 (遣) 家 (家)



こ 古 (古) 故 (故) 許 (許) 期 (期) 子 (子)

さ 乍 (乍) 佐 (佐) 佐 (佐) 左 (左) 差 (差) 散 (散) 斜 (斜) 沙 (沙)

し 之 (之) 之 (之) 事 (事) 四 (四) 志 (志) 新 (新)

す 受 (受) 壽 (壽) 數 (數) 數 (數) 春 (春) 春 (春) 須 (須) 須 (須)

せ 世 (世) 世 (世) 世 (世) 勢 (勢) 聲 (聲)

そ 所 (所) 所 (所) 曾 (曾) 曾 (曾) 楚 (楚) 蘇 (蘇) 處 (處)

た 堂 (堂) 多 (多) 多 (多) 當 (當)

ち 千 (千) 地 (地) 智 (智) 知 (知) 知 (知) 致 (致) 遲 (遲)

つ 川 (川) 川 (川) 津 (津) 都 (都) 徒 (徒)

て 亭 (亭) 低 (低) 傳 (傳) 天 (天) 天 (天) 天 (天) 帝 (帝) 豆 (豆) 轉 (轉) 而 (而)



と 土 (土) 度 (度) 东 (東) 登 (登) 登 (登) 砥 (砥) 徒 (徒) 忒 (等)

な 南 (南) 名 (名) 奈 (奈) 奈 (奈) 奈 (奈) 菜 (菜) 那 (那) 那 (那) 難 (難)

に 丹 (丹) 二 (二) 仁 (仁) 兒 (兒) 爾 (爾) 爾 (爾) 耳 (耳) 而 (而)

ぬ 努 (努) 奴 (奴) 怒 (怒)

ね 年 (年) 年 (年) 年 (年) 根 (根) 熱 (熱) 禰 (禰) 子 (子)

の 乃 (乃) 濃 (濃) 能 (能) 能 (能) 農 (農)

は 八 (八) 半 (半) 婆 (婆) 波 (波) 盤 (盤) 盤 (盤) 破 (破) 者 (者) 者 (者) 葉 (葉)
頗 (頗)

ひ 悲 (悲) 日 (日) 比 (比) 避 (避) 非 (非) 飛 (飛) 飛 (飛)

ふ 不 (不) 婦 (婦) 布 (布)



へ	倍 (倍) 弊 (弊) 弊 (弊) 遍 (遍) 邊 (邊) 辺 (邊) へ (部)
ほ	保 (保) 洵 (保) 報 (報) 奉 (奉) 寶 (寶) 本 (本) 本 (本) 豊 (豊)
ま	万 (万) 末 (末) 末 (末) 満 (満) 満 (満) 萬 (萬) 麻 (麻) 馬 (馬)
み	三 (三) 微 (微) 美 (美) 美 (美) み (美) 見 (見) 身 (身)
む	武 (武) 無 (無) 牟 (牟) 舞 (舞) 无 (无) 无 (无)
め	免 (免) 面 (面) 馬 (馬)
も	母 (母) も (毛) 毛 (毛) も (毛) 茂 (茂) 裳 (裳) 无 (无) 无 (无)
や	也 (也) や (也) 屋 (屋) 耶 (耶) 耶 (耶) 夜 (夜)
ゆ	游 (游) 由 (由) 由 (由) 遊 (遊)
よ	代 (代) 余 (余) 与 (與) 与 (與) 与 (與) 餘 (餘) 夜 (夜)



ら 羅 (羅) 良 (良) 良 (良) ろ (良) 𠂔 (等)

り 利 (利) 𠂔 (利) 李 (李) 梨 (梨) 理 (理) 里 (里) 離 (離)

る 流 (流) 畱 (留) 𠂔 (留) 𠂔 (留) 𠂔 (累) 𠂔 (類)

れ 禮 (禮) 𠂔 (禮) 連 (連) 麗 (麗)

ろ 呂 (呂) 𠂔 (呂) 婁 (婁) 樓 (樓) 路 (路) 露 (露)

わ 倭 (倭) 和 (和) 𠂔 (和) 王 (王) 𠂔 (王)

ゐ 井 (井) 𠂔 (井) 居 (居) 為 (為) 遺 (遺)

ゑ 惠 (惠) 衛 (衛) 𠂔 (衛) 𠂔 (衛)

を 𠂔 (平) 𠂔 (平) 尾 (尾) 結 (緒) 越 (越) 𠂔 (遠) 𠂔 (遠) 𠂔 (惡)

ん 𠂔 (无) 𠂔 (无)



覚えておくと宜いぐずし字

159 候 こう (こうがし)

候 候 候 候 候
 候 候 候 候 候
 候 候 候 候 候
 候 候 候 候 候

被 ひ (ひがし)

被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被
 被 被 被 被 被

39 也 なり

也 也 也 也 也
 也 也 也 也 也
 也 也 也 也 也
 也 也 也 也 也

1901 歟 やか

歟 歟 歟 歟 歟
 歟 歟 歟 歟 歟
 歟 歟 歟 歟 歟
 歟 歟 歟 歟 歟

32 乍 な

乍 乍 乍 乍 乍
 乍 乍 乍 乍 乍
 乍 乍 乍 乍 乍
 乍 乍 乍 乍 乍

書簡体文字

合略仮名

- 候の畧字 (さふらふ)
- 候の畧字
- 奉の草體きの畧字 (まゐらせ)
- 被の畧字
- よりの合略字
- さほの合略字 (さま)



古文書と古典籍

古文書

特定の対象に意思を伝えるために作成されたもの（差出人／受取人）

そのほか目録・帳簿なども古文書の対象に考えることも多い。

⇒だから何処の由来のものかが大事！
価値が変わる

「好古日録」（1797）

古文書の世に存する、或は朽敗、或は魚食、又は水火に逢て日に減ず

※魚食 紙魚（しみ）

※和紙（世界文化遺産）

古典籍

辞書にあまり掲載されることのない言葉！

広辞苑 和書・漢籍のうち、近代以前の写本・版本の総称

典籍…和書・漢籍・仏典など書物・書籍

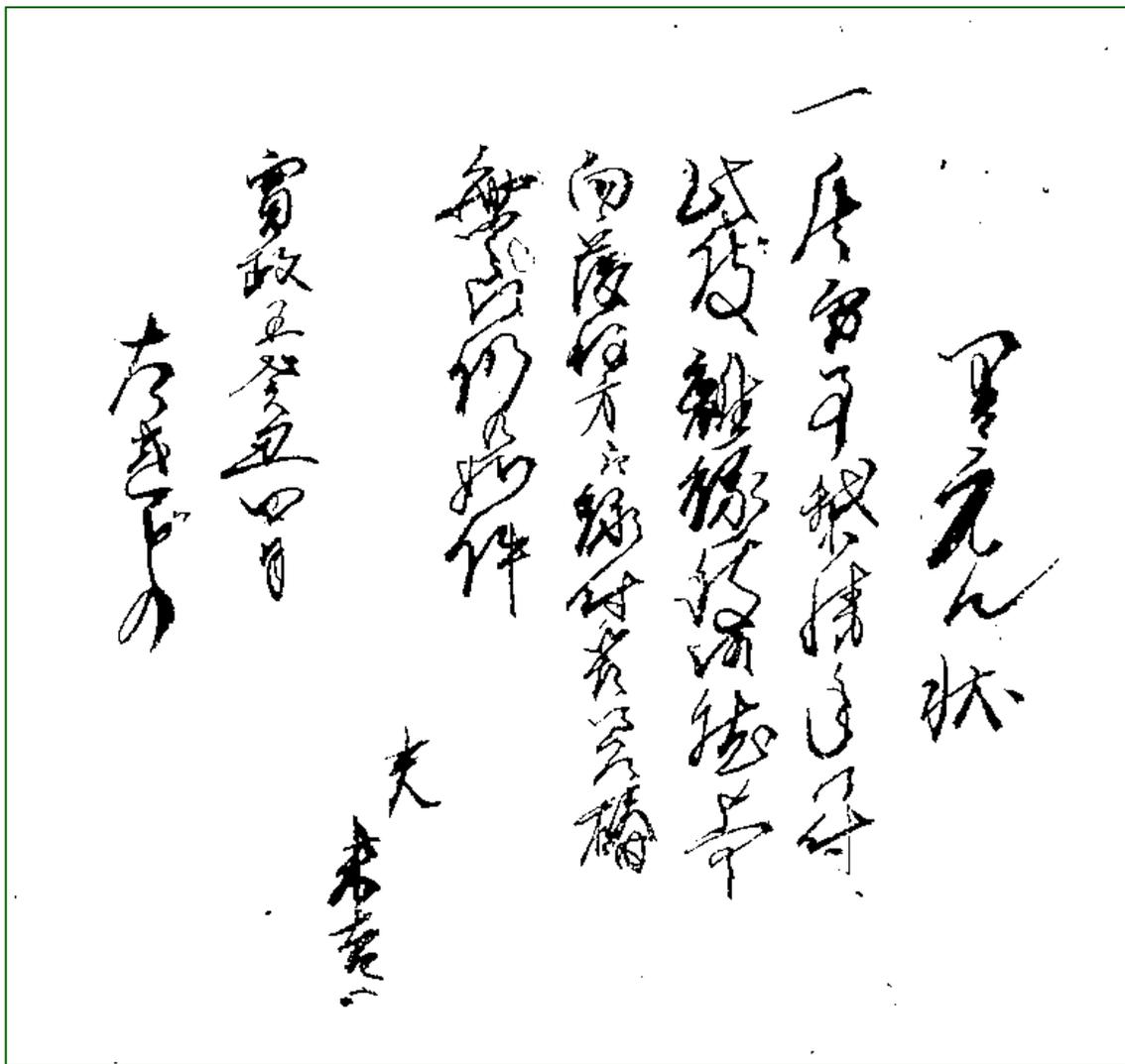
古典…古い典籍、昔の経典

島崎藤村「落梅集」の用例が古く、優れた著述や作品で過去の長い年月にわたって多くの人々の模範となり、また愛好されてきたもの、の意が出てきている（近代の産物）。



くずし字を読むために

パターンを覚える (特に古文書)



りえん状

一其方事我等勝手に付

此度離縁致候然上は

向後何方江縁付候共差構

無之候 仍而如件

夫 末吉 (爪印)

(みずのとうし)

寛政五癸丑四月

たきどの

いわゆる三下り半

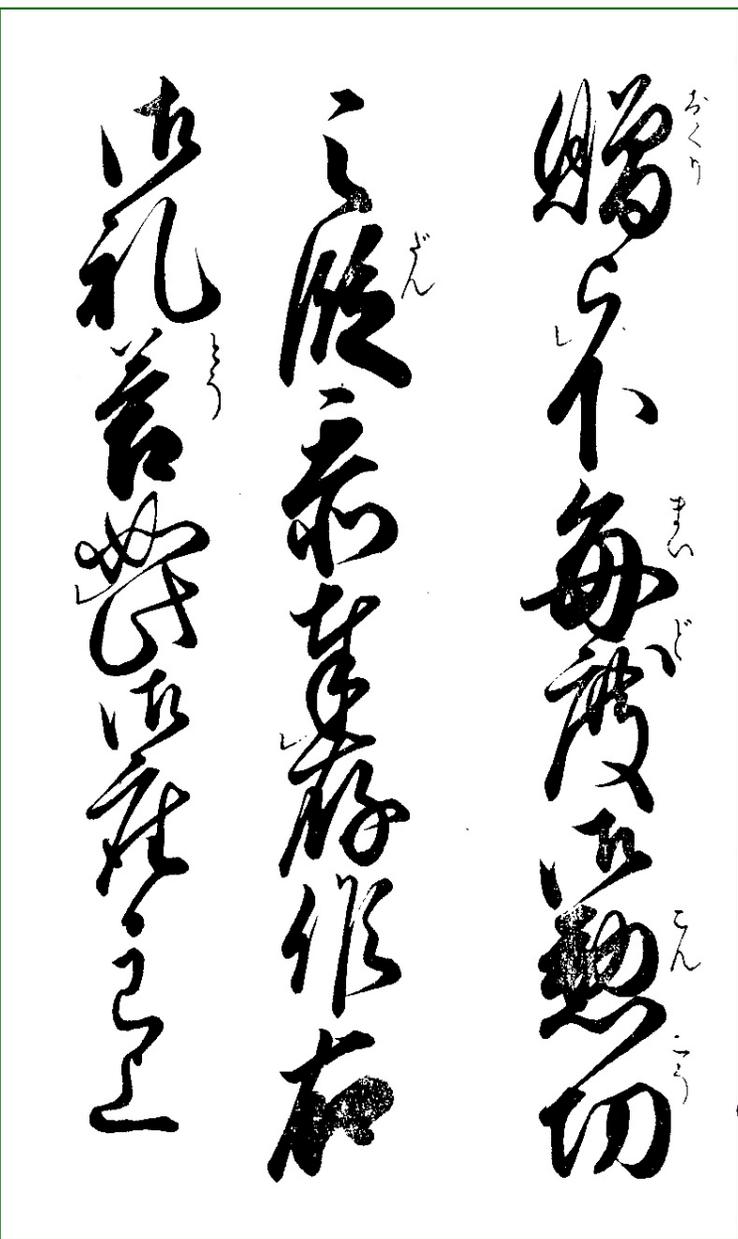
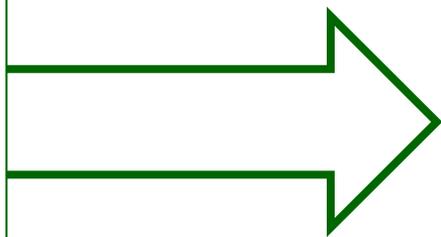


くずし字を読むために

書きならわしを知る

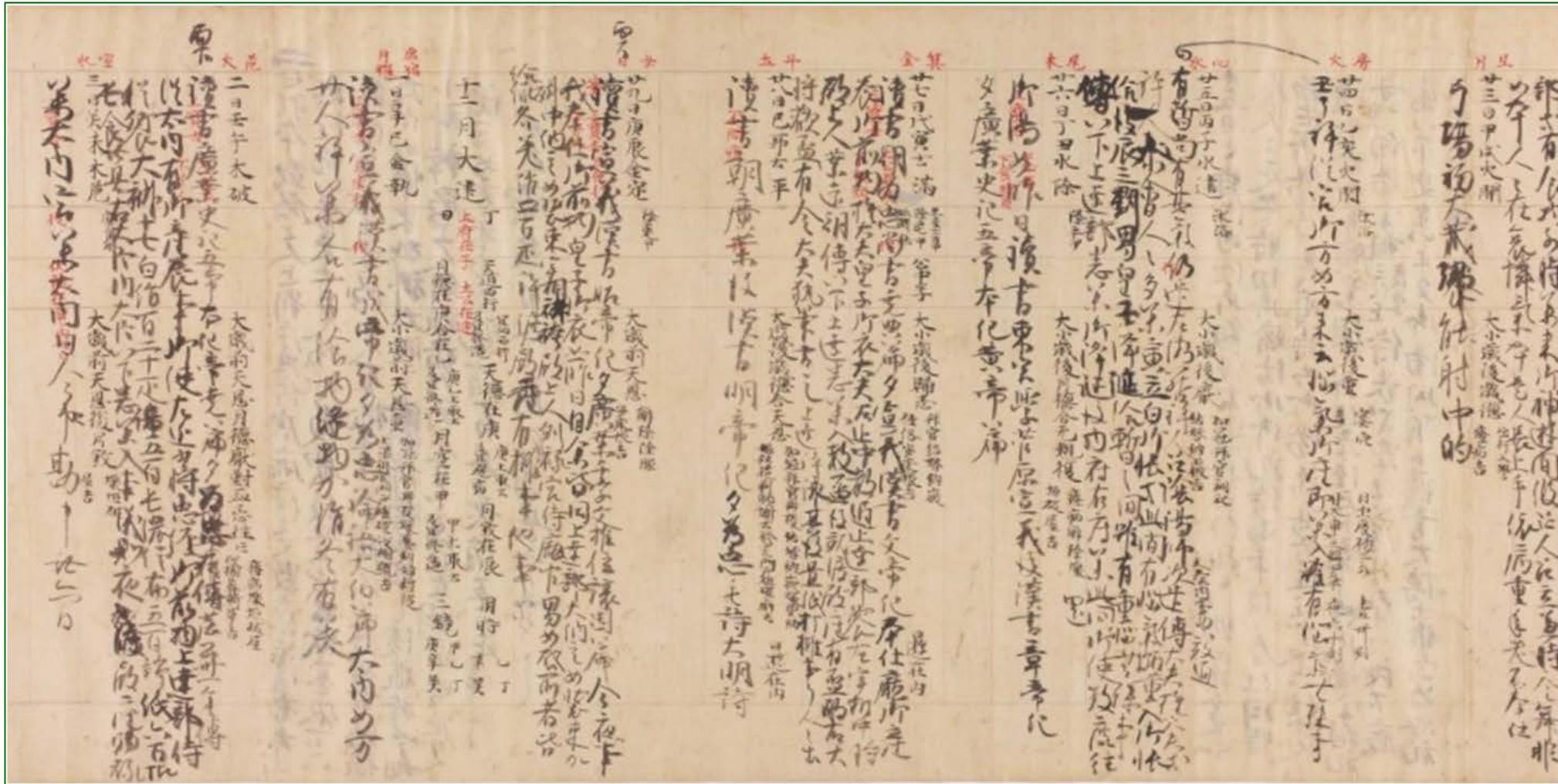
『随一用文章』
往来物。年頭から歳暮までの
季節の手紙の例文十四通と、
その返信を載せている。

贈被^レ下毎度御懇切^(懃)
之段忝奉^レ存候右
御礼答如^レ此御座候已上





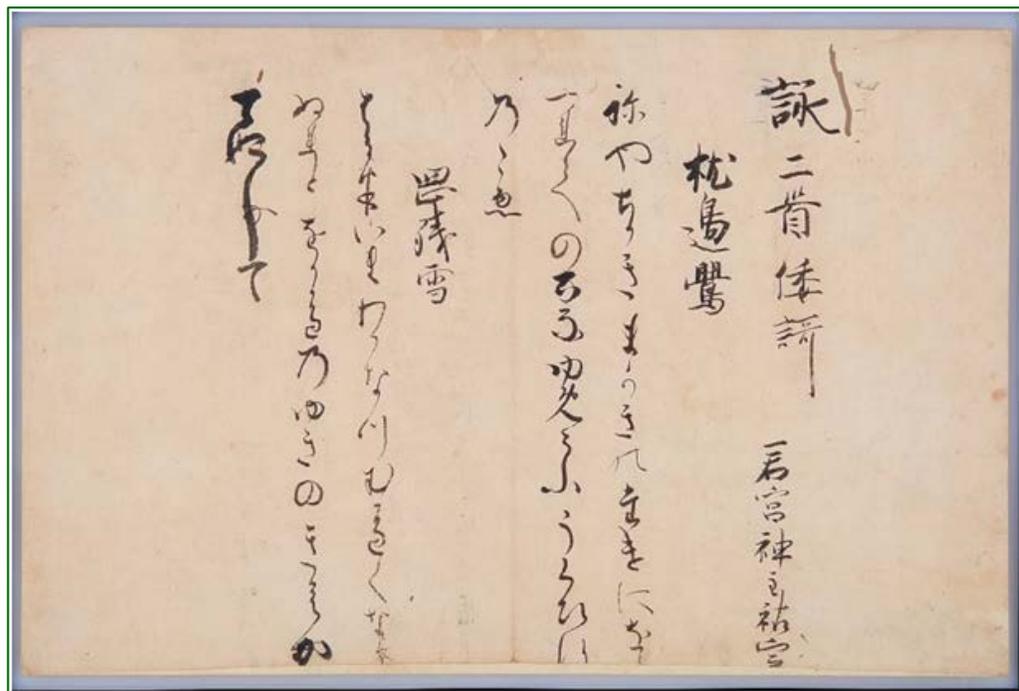
古典籍？古文書？



藤原道長自筆 寛弘6年（1009）「御堂関白記」寛弘六年下巻（部分）
漢字を主として使用（真名）



古典籍の対象あれこれ

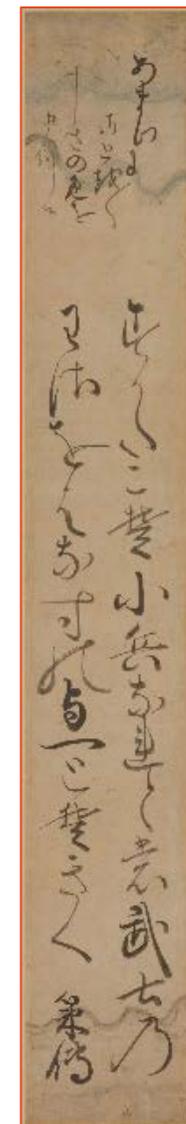


春日懷紙 (重文)

三十六歌仙帖



書物の形態だけではなく様々に古典籍の対象として扱われているのが現状



短冊



り川のなる神を祀りて
本の水よりあつた
こころ大空木川流
とまるるの形一
世の中みあ

嵯峨本
方丈記

好色
一代男

其の面影
とて思ふ世の人
とて思ふ世の人
とて思ふ世の人

四五年おとせ前まへ童話わらわ比出ひだおととつづりてかきとらるるも
いまだ考れ足たざるこころおととつづりてかきとらるるも
よみだれども打出うちだの小櫃こづの事とまてらるるも

骨董集

八犬伝

八や犬いぬ伝でん
歳としのと月つきのと山やまよりいりまり日よひりまり
うたのを何なによもと有身みはなき

正ただしく記國史こくしの體はなりてなり
ばあらまり取見みる人も罕らまりなり世々
ちよくも大おも言ひてなり一撮しつ

宣長
古事記伝

篤胤
古史伝

て大土おほつちを云ひと聞きえて古言ふるごありてはりて
を泥土ひぢは堅かたはりて國土くにを成なれはりてはりて
小こくも大おも言いひてなり一撮しつ
廣ひろく海うみは對かへる陸地くわがをも云ひてなり



典籍はステータスシンボルであり、「文化」だった



伊勢物語／絵入読曲

書物を所蔵することは家のステータスを示すものだった

中国 皇帝 四庫全書

焚書 毛沢東

日本 庄屋 蔵に保管された書物

床の間 書物を見せる 知識人

伊勢物語・古今和歌集・万葉集

源氏物語△

中国書なら古文真宝・唐詩選



典籍はステータスシンボルであり、「文化」だった（実例）

141
八ッ橋水禽図印籠 銘（梶川作）
印籠美術館



印籠の意匠

小袖の意匠



9 白繪子地八橋模様紋縫小袖 野村コレクション「小袖屏風」 国立歴史民俗博物館



典籍はステータスシンボルであり、「文化」だった（実例）

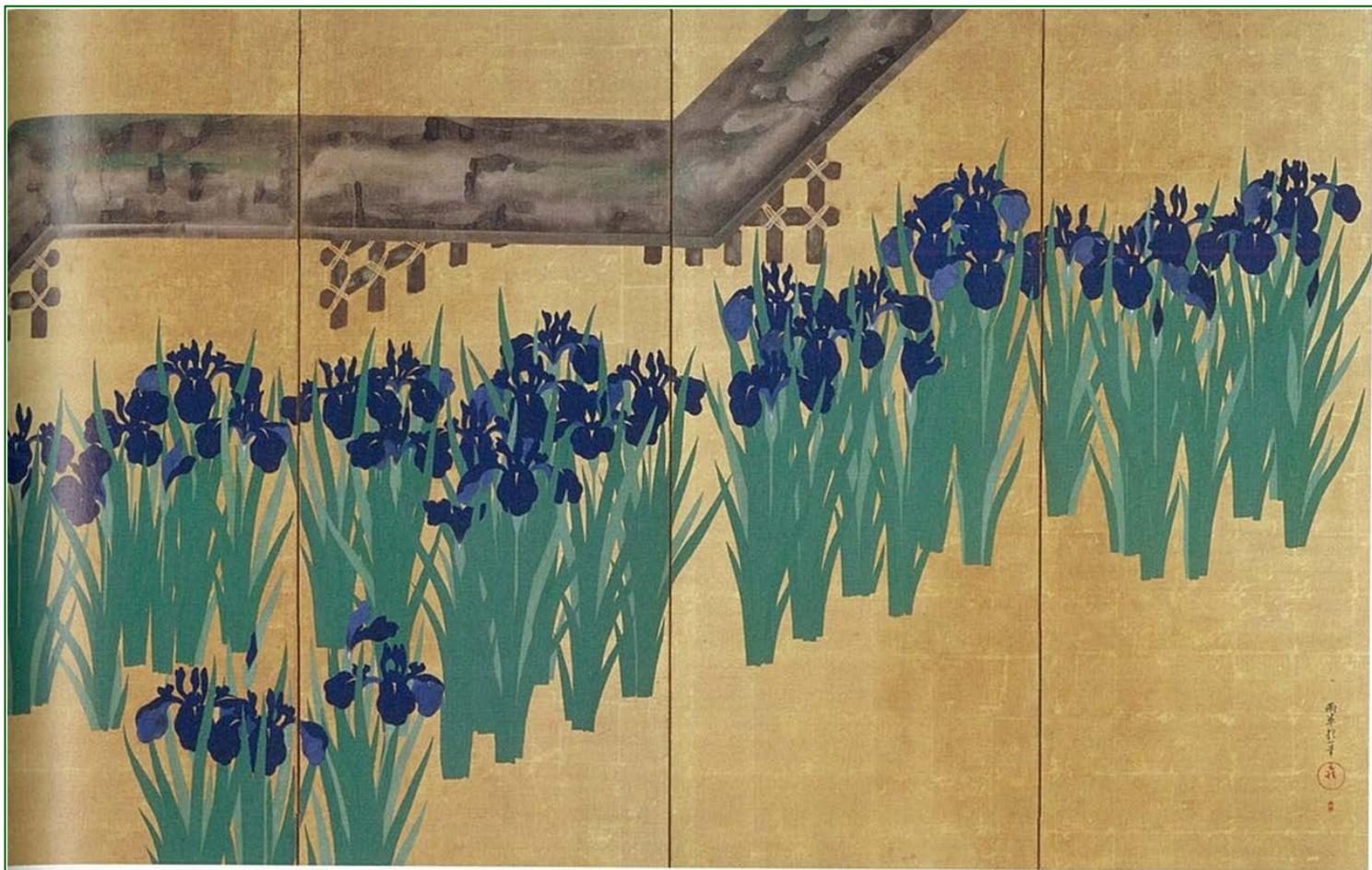
伊勢物語第九段（口語訳 一部抜粋）

昔、男がいた。その男、自分を世の中には無用の人間であると思ひ込んで、「京には住まないつもりだ。東国の方に住める国を探しに行く」ということので出かけて行った。昔からの友人一人二人と連れ立って行った。道を知っている人もいなくてあちこち迷いながら行ったのだった。三河の八橋という所に着いた。そこを八橋というのは、水が流れていく川が蜘蛛の手のように八方に分かれていて、橋を八つ渡してあることから八橋と

いったのである。その沢のほとりの木蔭に馬から降りて座って乾飯を食べた。その沢に、かきつばたがともきれいに咲いていた。それを見てある人が、「かきつばたという五文字を各句の頭に置いて（＝折句）、旅の心情を詠みなさい」と言ったので、男は詠んだ。

唐衣 きつゝ 馴にし つましあれば はるばる来ぬる 旅をしぞ思ふ

（馴れ親しんだような妻が都に居るものだからはるばるとこんなに遠くまで来てしまった旅を悲しく思うのです）



江戸琳派酒井抱一「八橋図屏風」（一部）



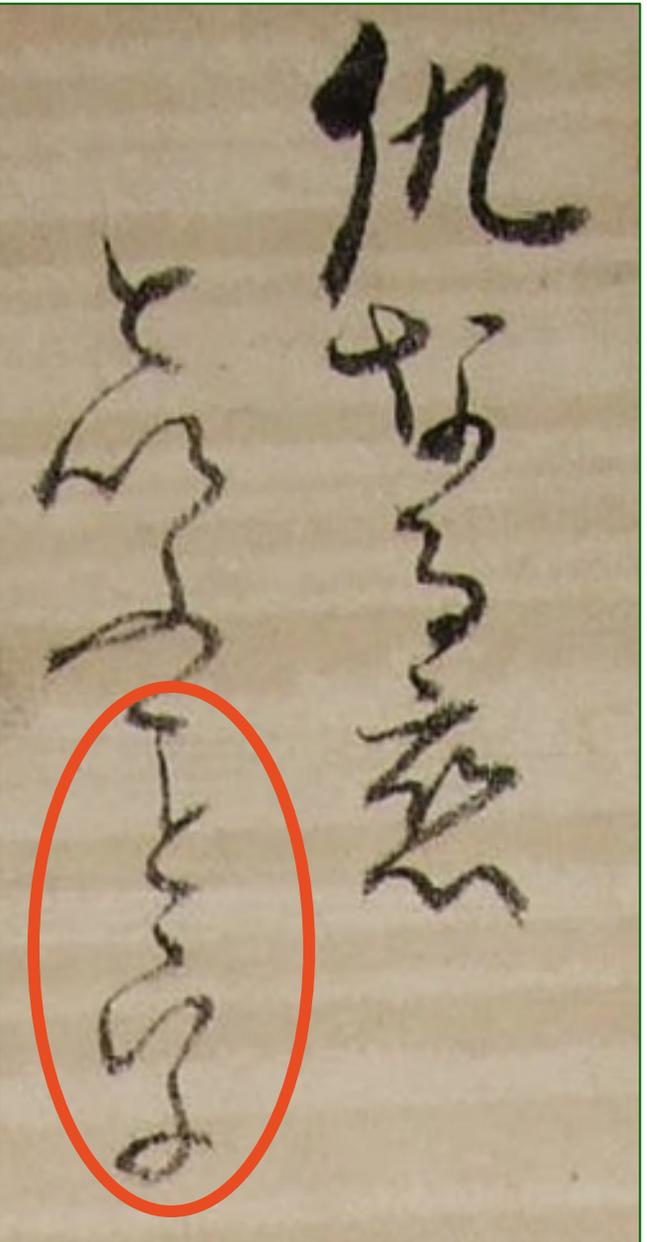
短冊を読む

八犬伝を書いた曲亭馬琴 (1767~1848)
の短冊をよんでみよう



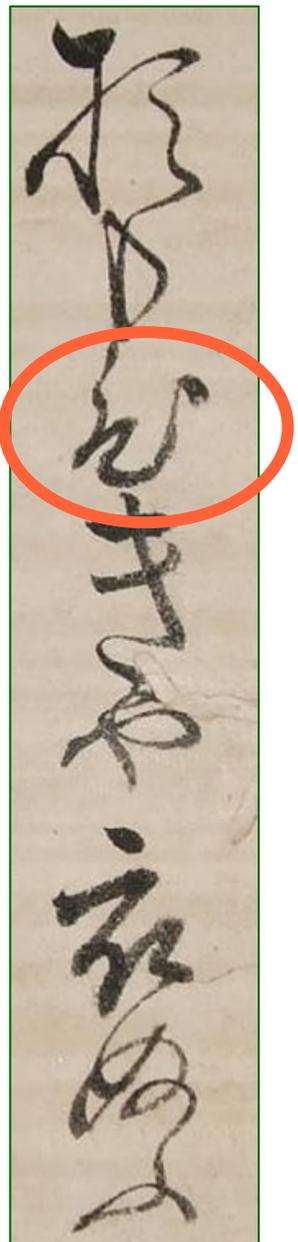


仇（あだ）なる恋
といふことを

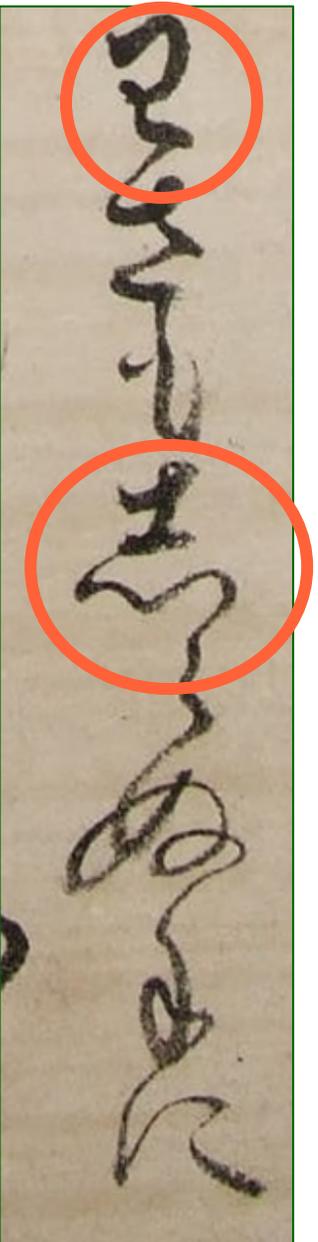




おもひきや衣ぬふ



おもひきや衣ぬふ





針もて君か名を

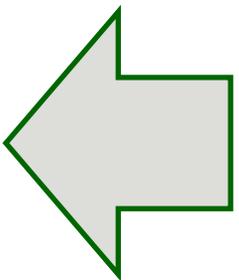


ほらんとは 馬琴

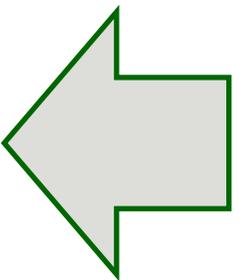




まとめ



仇なる恋といふことを
おもひきや／衣ぬふわさも／しらぬ
手に／針もて君が／名をほらんとは
馬琴



かりそめの恋ということを
だれが思い描いたことだろうか／
着物も縫うことも／知らない私の
腕に／針でアナタの／名前を彫ろ
うなんて 馬琴

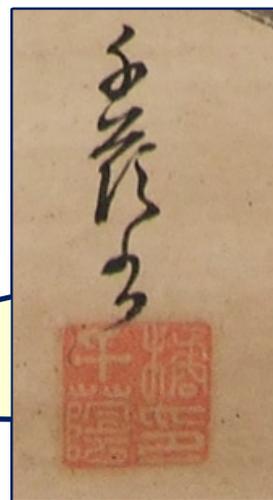
(遊女に仮託して艶っぽい情景を描いたか)



掛け軸を読む

画賛

画の余白に書き添える文章や詩歌



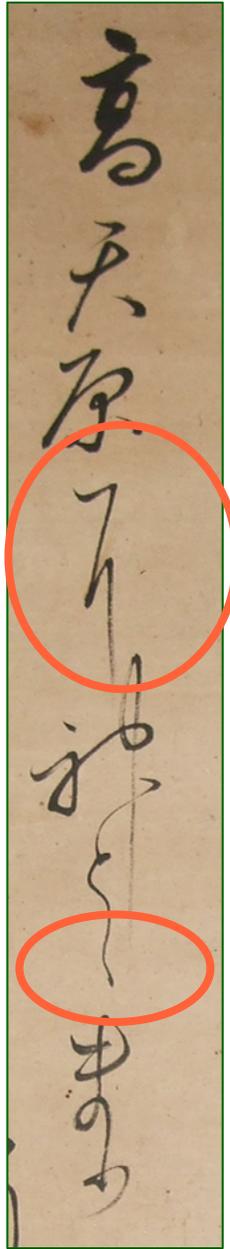
加藤千蔭

江戸の町与力、歌人、国学者、書家。賀茂真淵に学び『万葉集略解』などを記した。

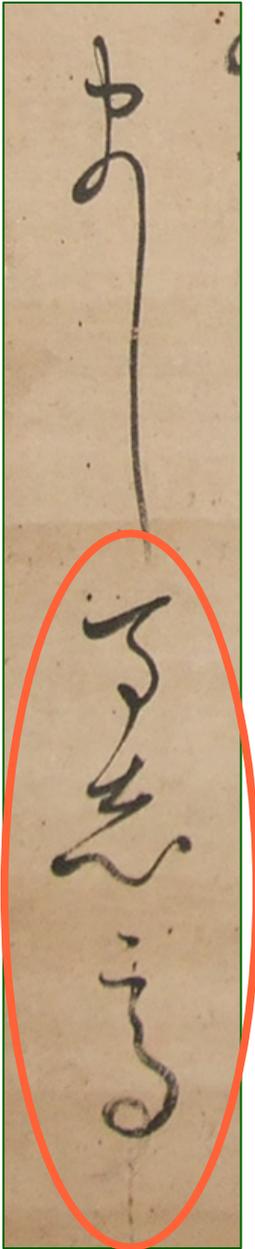
1735～1808



高天原に神とまり



高天原に神とまり





画文共鳴

※高天原
 (たかまのはら・たかまがはら)
 日本神話の聖地で神々が住んでいるところ。アマテラスの支配する天上界

神官が幣(ぬさ)を手に持ち
 所在なげに空を仰ぎ見る図
 高天原に神とゞまり 千蔭書
 ましまして





くずし字アプリの活用 K u L A

アプリで学ぶくずし字

くずし字学習支援アプリ

KuLA

クーラ

ダウンロードはこちらから！



くずし字学習支援アプリKuLA
Yuta Hashimoto 教育 ★★★★★ 51
インストール

はじめに
「まなぶ」機能の使い方
くずし字学習の基本
文字検索
実体仮名をまなぶ
実体仮名をまなぶ (ア～ヤ行)
実体仮名をまなぶ (エ～ハ行)

あ (安)
あ (阿)
い (以)
う (宇)
え (衣)
お (於)
か (加)
か (可)

2016年3月9日: 本アプリが朝日新聞関西版および共同通信で取り上げられました！

浮世絵に書かれた文字を読みたいと思ったことはありませんか？
江戸時代の刀剣書を読みたいと思ったことは？

江戸時代以前に出版された書籍を読むために、現代では使われることのない文字や書体で書かれた「くずし字」を学ぶためのアプリ「KuLA」を開発しました。

もっと見る

<https://kula.honkoku.org/>



AI手書きくずし字検索

AI 手書きくずし字検索

下の枠内に1文字だけ手書きしてみてください↓

消 去

部首指定なし

※ブラウザでJavaScriptを有効にしてください
※IE11には対応していません
○このサイトの使い方
○このサイトについて

<http://www.ai-kuzushiji.net/>



データセットに利用された古典籍

餅菓子即席手製集
 好色一代男
 鼎左秘録
 菓子話船橋
 おらが春
 歌学提要
 万宝料理秘密箱
 地震考
 新板料理秘伝抄
 世間胸算用
 椿説弓張月
 春色梅児与美
 料理方心得之事
 膳部料理抄
 料理珍味集
 かてもの
 うそなるべし
 傾城買四十八手
 南総里見八犬伝
 武家義理物語
 新編異国料理
 日本永代蔵
 吉利支丹物語
 諸国方言物類称呼
 御前菓子秘伝抄
 雨月物語
 日用惣菜俎・不時珍客即席包丁合刻
 玉くしげ
 当世料理
 二人比丘尼
 料理物語
 竹斎物語
 比翼連理花迺志満台
 養蚕秘録
 浮世風呂
 源氏物語
 伊曾保物語
 をみなへし
 うすゆきものかたり
 安倍晴明物語天文巻



くずし字字形データの活用例

<https://mp.ex.nii.ac.jp/kuronet/>

KuroNetサービス KuroNet くずし字OCR CODH ログインして下さい

KuroNetくずし字認識サービス

日本古典籍くずし字データセットを用いたAIくずし字OCRサービスとして、「KuroNetくずし字認識ビューア (International Image Interoperability Framework)」に準拠した画像であれば、世界中で公開されるくずし字データセットから、指定した領域のくずし字を抽出し、認識した一文字を返すサービスです。

KuroNetくずし字認識サービスは、「KuroNetくずし字認識ビューア」と「ダッシュボード」でくずし字OCRを行う領域を指定し、「ダッシュボード」でくずし字OCRの結果を確認できます。

ログイン後にダッシュボードへ移動
(クリックしても何も反応しない場合は、ログインして下さい)

KuroNetくずし字認識ビューアを起動

利用方法

1. 利用登録の後、ログインします。なおログインについては、「KuroNetくずし字認識サービス」の「利用登録」ページをご覧ください。
2. 「KuroNetくずし字認識ビューア」に、認識したい本のIIIFマニフェストをドロップして、認識したい領域を指定し、認識した一文字を返すサービスです。

人文学オープンデータ共同利用センター Center for Open Data in the Humanities 日本語 / English メニュー

日本古典籍くずし字データセット / くずし字アプリ・サービス / KogumaNetくずし字認識サービス (一文字)

KogumaNetくずし字認識サービス (一文字)

KogumaNetくずし字認識ビューア 日本語 / English ログイン

IIIF Curation Viewer / Drop a manifest URL to open

使い方

ディープラーニング(機械学習/AI)を用いてくずし字を一文字認識するサービスです。IIIF (International Image Interoperability Framework)による画像配信に対応していれば、世界のどこのサイトでもお試しできます。

まず、KogumaNetくずし字認識ビューアにIIIF Manifestを読み込みます。次に画像の一部領域を切り取ります。具体的には、ビューアの右上に表示される「■」ボタンを使って、認識したい領域を四角で囲みます。次に四角をクリックすると、その中に書いてある一文字(単文字)を認識し、その結果をポップアップウィンドウ内に表示します。

<http://codh.rois.ac.jp/char-shape/app/single-mobilenet/>

1 文字認識の使い方 https://researchmap.jp/blogs/blog_entries/view/90992/f7e0a6ae6ddb61744ce34c564d92f87d?frame_id=426673



くずし字字形データの活用例 教育ツール

豊田高専くずし字翻刻WWWサービス (<http://vpac.toyota-ct.ac.jp/hayasaka/kuzushiji/>)

国立高等専門学校機構 豊田工業高等専門学校の早坂太一氏、大野互氏、加藤弓枝氏らによる、くずし字認識を行うWWWアプリケーションです。日本古典籍データセット等の文書画像を読み込ませ、1文字をマウスで選択すると、ディープラーニングによる画像認識技術により、翻刻結果が出力されます。

《例》『絵本見立仮名尽』(書誌ID: 200015136/DOI:10.20730/200015136)で試した場合

深層学習によるくずし字認識 (ver.0.5.1)

文字の対応

Recognition of Kuzushiji (Hentaigana and cursive script) by Deep Learning (ver.0.5.1)

画像選択 (カメラ起動) Image Selection

反時計回り CCW

時計回り CW

くずし字を認識 Recogniton

どる Reset

文書画像を読み込み
(カメラ撮影も可能)
翻刻したい文字をマウス
や指で選択してください

Please read a document image
file (camera photography image
is also possible) and extact
the *Hentaigana* letter which you
want to know with a mouse or
your fingers.

本研究はJSPS研究助成(21660213)の助成により平成28年度から令和2年度まで実施されたものです。
This work was supported by JSPS KAKENHI Grant Number JP16020433 and The Naoto Research Grant 2016.

豊田高専くずし字翻刻WWWサービスの画面を開きます。

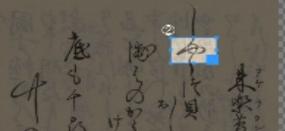
画像選択 (カメラ起動) Image Selection

反時計回り CCW

時計回り CW

くずし字を認識 Recogniton

どる Reset



ディープラーニングによるくずし字翻刻 (豊田高専)
学習データに字形データ利用 山本研究協力者
<http://vpac.toyota-ct.ac.jp/hayasaka/kuzushiji/>

2020年01月27日 中日新聞 夕刊 9頁

くずし字 AI で 活字に

豊田高専 翻刻システム開発

日本の古文書、古典文学などの「くずし字」を人工知能(AI)が高精度で認識し、現代の文字に置き換える「翻刻」のシステムを豊田高専(愛知県豊田市)などの研究チームが開発した。持ち運びに便利なコンパクトさで、古民家での調査や学校の国語教育で気軽に利用できそうだ。(谷村卓哉)



研究チームの中核を担った早坂太一准教授(左)と竹内正広さん(右)愛知県豊田市の豊田高専で

よく「ミスがはったよう」と形容される「くずし字」は、明治以降に学ばれなくなるまで千年間も使われた。江戸末期までに印刷または書写された膨大な書籍や文書は、過去の文化、社会状況から災害、自然現象まで幅広い研究に役立つ情報宝库だが、解読できるのは限られた専門家だけで多くが死蔵、廃棄されている。そこで、くずし字を誰でも読める形に効果的に変換する研究が、さまざまな手法で試みられている。

高専の早坂太一准教授(左)、専攻科二年竹内正広さん(右)らが国の科学研究費補助金を得て二〇一六年から開発を進めたシステムは、カメラ撮影などで取り込んだ史料画像の字を一つ一つ検出し、それが今のどの字と同じか、認識結果を確率順に五位まで提示。変換したい候補を選ぶと画像上に反映する。

元はプログラミング教材として開発され、六千円程度で市販されているカード大のコンピュータ「ラズベリーパイ」にAIを搭載。文字列の検出に四〜五秒、認識に一字あたり〇・〇八秒程度の処理速度だ。江戸時代に出版された「源氏物語」の約一万字を試したところ、はじき出された確率一位の字と正解の字が97%一致。書写された別の藩政記録では、字のく

江戸の源氏物語では精度97%

せや墨のじみ、かすれもあって成績を下けたものの、70%以上の精度を示した。

AIの構築には、膨大なデータから物事の傾向や特徴をかませるディープラーニング(深層学習)を利用。「くずし字」のデータを集積、提供する情報・システム研究機構「人文学オープンデータ共同利用センター」(CODH、東京)の公開データに、東京の民間企業「TRC A D E A C」提供の図書館、博物館所蔵史料のデジタル画像から高専の学生が約四年かけて抽出した一万字を加え、四千種以上、百一十万字を学ばせた。

学習量は成績に直結し、七万字を学び終えた段階で「元」の解読は正解率40%強だったが、昨年十一月にCODHが追加公開した四十万字のデータを加えると95%に改善されたという。

解読結果は最終的には国文学者の確認が必要だが、早坂准教授は「掛け軸や文書に何が書かれているのか、ちょっと調べたい」と話す。利用の依頼があれば対応する。利用の依頼が深まるので、たくさん使ってほしい」と期待する。

中学校までくずし字の出前授業を行う愛知県立大の三宅宏幸准教授(近世文学)は「くずし字に親しむ人が増え、価値が分らず捨てられることも多い古文書などの保存につながる。天文現象や地震などの情報は専門家でも引き出すのが難しく、理系分野の研究の助けになる可能性もある」と指摘する。



小学堂(台湾)

金文

首頁 簡介 收錄現況 凡例 使用說明 參考書目 聯絡信箱

簡易查詢 進階查詢

字號

字形 海

頁碼

器號

器名

器目

時代

1801 𣎵 (海)

共搜尋到3字/1頁, 每頁 20 字, 字形大小: 36 點

 小臣謎簋 西周早期 集成4238	 小臣謎簋 西周早期 集成4238	 小臣謎簋 西周早期 集成4239		
--	---	---	--	--

(點選字形可取得字形編輯資訊)

相關連結

- [小學堂異體字表](#)、[字形演變](#)、[小篆](#)、[楚系簡帛文字](#)、[傳抄古文字](#)
- [小學堂上古音](#)、[中古音](#)、[官話](#)、[晉語](#)、[吳語](#)、[徽語](#)、[贛語](#)、[湘語](#)、[閩語](#)、[粵語](#)、[平話](#)、[客語](#)、[其他土話](#)
- [小學堂韻書集成](#)、[廣韻](#)、[集韻](#)、[附釋文互註禮部韻略](#)、[增修互注禮部韻略](#)、[中原音韻](#)、[中州音韻](#)、[洪武正韻](#)

相關索引

<http://xiaoxue.iis.sinica.edu.tw/>



くずし字字形データの活用例

くずし字、いろいろ。

あなたのことをくずしてみよう。

くずしじとええあい

- 1回に変換できる文字数は25文字までです。
- できるだけひらがなで入力してください。
- 変換できない文字もあります。変換できない場合は空白になります。
- 濁点などなしにすると出やすくなります。(例 ず→す)

くずし字に変換！
(我春集ver.)

くずし字に変換！
(昔)

「くずしじとええあい」
を古文書に記録しました。

くずしじとええあい

『日本古書籍字形データセット』(国文研ほか所蔵/ CODH加工)
提供: 人文学オープンデータ共同利用センター

くずしじとええあい

くずし字、いろいろ。

<https://lab.nijl.ac.jp/kuzusu/>



くずし字学習用

- ・ 誠心堂書店 変体仮名を調べる
<http://www.book-seishindo.jp/kana/>

花の色はつりにけりな
 わがみよになる
 ながめせしまに
 小野小町



テスト版アプリケーション

古典かな、歴史的仮名書体

変体仮名を調べる

「ひらがな」は現代では一音一字ですが、平安時代以来明治の初めまで、さまざまな種類の字体が用いられてきました。この歴史的書体である「変体仮名」を知らなければ、江戸時代以前の和本や文書を読むことも調査することもできません。本サイトは変体仮名を調べやすいように工夫してみました。まだ未完のテスト版です。[ご意見を下されば幸いです。](#)



くずし字判読ひとりまなび

♡ 20


 静大生錦画深読み
 2021/01/27 13:03

- ・ 小二田誠二（静岡大教授）くずし字判読ひとりまなび
https://note.com/shizu_nishiki/n/nb329d58687a0



読めないとせず、読める文字を探してみる（練習）



百人一首古今狂歌袋



それではこのセッションを終わりにします