

文化資源特別講義b 文化資源特論III

大阪市立大学・大阪市立大学大学院講義資料

福田一史

fukudakz@gmail.com

<https://scrapbox.io/fukudakz/文化資源学特別講義b・文化資源特論III>



図. 講義ウェブサイトリンク (QRコード)
※LMSの資料のページにもリンクがあります

前回講義のコメントより

- 「（PCについて）正直少し敏感になりすぎている人も多い気がしてしまいました。しかし、実際私も、これは流石にダメでしょと思うCMやポスターに出会ったこともありますし、その「常識的にアウトかセーフか」というラインは本当に難しいなと感じました。」
- 「そのゲームの媒体がカセットなのかソフトなのかによりますが、もし後者なのだとすれば随時アップデートすることができるのはデジタルの良さである」
- 「（ルックバックについて）広く読まれたことによってさまざまな意見が出て、それに対して作者側が応答して変更するというかたちは新しいのではないかとも思います。」
- 「人権を軽んじるような描写があったとして、その原因を作者の思想に求めることもできるが、キャラクター自体の思想として作者のものとは別個に考えることもできる。」

前回講義のコメントより

- 「（ゲームやアニメに）社会的な意義 = 付加価値 が今の時代には求められていて、そうした表現の幅の広さゆえに文化的な翻訳や差別問題に人々が敏感になってしまう」
- 「歴史的に男性目線の、消費されるような女性性が強調される部分が多かったため、そうした表現が現状においても続いてしまっていて、フェミニズムといった思想と衝突を起こしているのだと思われる。」
- 「SNSなどでもよく話題に上がる分、人々の意見をリアルタイムに取り入れながら"適切"な表現を模索しなければならない」

前回講義のコメントより

- 「...ゲームでも文化的翻訳が必要なんやと思っておもしろかったです。ディズニーの例だと、『インサイドヘッド』で子どもが嫌いな食べ物として、アメリカの原作ではブロッコリーやったけど、日本バージョンではピーマンになってる、とか。」
- 「（MODはゲームを最適化したり、運営が改善できていないバグですら修正できる、良い方向に向けてのものもあれば、いわゆるチートと呼ばれるものもあり、対人ゲームに関しては、明らかにアウトだが、アンチチートは常にいたちごっこなので、それも課題である。」
- 「二次創作はグレーとされていますが、もしこれから規制の対象になっていくとすれば、二次創作が盛んに行われる日本は特に打撃を受けそう」
- 「FPSゲームをやる時に女性が差別視されるのも感じますが、近年e-sportsには女性選手の姿もだんだん増えています。」
- 「「漫画村」が違法に漫画を転載して問題となっていたのは記憶に新しく、アニメ（というか映画も含めた映像作品全般）も少し探せば違法転載サイトがあるイメージ...技術の発達により、データが複雑化・大容量化して漫画やアニメと比べるとコピーが難しくなった」

前回講義のコメントより

- 「個人的に高校時代世界史と地理の勉強の際、国のイメージがしやすいという点でヘタリアには大変お世話になったので、「勉強になる漫画」というコンテンツが着目されていることは嬉しかったです。」
- 「（社会応用について）私の友人に、アーサー王が女性であると思い込んだまま授業の研究発表に用いようとした人がいた。他にも三蔵法師を女性であると思っていた人がいたので、そういう危うさがあると思う。」
- 「（ゲームさんぽについて）『アサシンクリードオデッセイ』を用いて、考古学者とともにギリシャ観光をしていたのは個人的にとっても興味深かったです。マンガでもゲームでもあくまで現実とは異なる創作にとよんだ内容だとは思いますが、ゲームをすることを通じて新しい発見や知的好奇心を刺激する点が重要」

文化資源としてのアニメ・マンガ・ゲーム（まとめ）

- ビジネス、産業、経営学
- インディーズ
- メディアミックス、トランスメディア・ストーリーテリング
- グローバリズム・文化翻訳
- ポリティカリー・コレクトネス
- プラットフォーム・スタディーズ
- 社会応用（学習マンガ、シリアスゲーム）

コンテンツ

1. デジタル技術を用いた文化資源の分析

デジタル技術を用いた 文化資源の分析

計量的に文化を記述し、分析する

既存の文化研究

- 文化を研究するための理論・方法論は多様
 - クローズ・リーディング
 - 美学・分析哲学
 - 文化人類学・民族誌
 - 経済学・社会学
 - 心理学
 - メディア論

デジタル技術による文化研究

- 文化資源のデジタル化と、分析のためのコンピュータ技術を用いた方法論の開発・発展により、それまでにできなかった研究が可能に。
- さらに、人文学以外の領域、社会科学、自然科学、芸術などとの融合、越境が促進される。
- その代表的・中心的な動向が「**デジタル・ヒューマニティーズ**」。

デジタル・ヒューマニティーズ (DH)

- Digital Humanities (DH) ≡ Humanities Computing ≡ 人文情報学
- デジタル技術を用いた人文学研究。主に文化資源が研究対象になる。
- 「デジタル・ヒューマニティーズとは、共同で、分野横断的に、コンピュータを用いて取り組まれる、研究、教育、出版のための学問と組織の新しいあり方のことである」
 - Digital Humanities, by Anne Burdick, Johann Drucker, Peter Lunenfeld, Todd Presner, Jeffrey Schnapp, MIT Press, 2012
 - デジタル・ヒューマニティーズ入門. <https://www.dhii.jp/nagasaki/blog/sg2dh>.

DHの起源

- 1949年、Roberto BusaがIBMと共同で、トマス・アキナスの著作の用語索引を機械的に操作する研究に着手。
 - 単語検索、ソート、表作成などを自動化する
- 1966年、専門誌 “Computer and the Humanities” が刊行。
- 1973年 “the Association for Literary and Linguistic Computing” が、1978年には “the Association for Computers and the Humanities” が設立。
 - 「デジタル・ヒューマニティーズ入門」より

多様なDHの方法論・視覚

- The Digital Humanities Literacy Guidebook
 - <https://cmu-lib.github.io/dhlg/>



TABLE OF CONTENTS

- 3D Modeling
- Black Digital Humanities
- Computational History
- Computational Linguistics
- Corpus Linguistics
- Critical Code Studies
- Cultural Analytics
- DH Feminism
- DH Postcolonialism
- Data Visualization
- Digital Art History
- Digital Editions
- Digital Forensics
- Digital History
- Digital Pedagogy

- Digital Public History
- Digitization
- Distant Reading
- Global Digital Humanities
- HGIS
- Humanities Visualization
- Latinx Digital Humanities
- Machine Learning
- Multilingual DH
- Network Analysis
- Open Data
- Oral History
- Stylometry
- Text Encoding Initiative
- Text Mining and Analytics
- Virtual Reality
- Web Archives

☒. Table of Contents | Topics in DH. <https://cmu-lib.github.io/dhlg/topics/>

テキストマイニング

- テキスト（文章・文字データ・文献など）を対象とするデータマイニング。
- 文章を単語や文節で区切り（形態素解析）、それら単語や文節の出現頻度や共起関係などを計量的に分析する。自然言語処理技術の一種。
- KH Coder
 - テキストマイニングソフトの一つ。
<http://khcoder.net/>



#	抽出語	品詞/活用	頻度
1	先生	名詞	595
2	K	タグ	411
3	奥さん	名詞	388
4	思う	動詞	296
5	父	名詞C	269
6	自分	名詞	264
7	見る	動詞	225
8	聞く	動詞	219
9	出る	動詞	185
10	人	名詞C	182
11	母	名詞C	170
12	お嬢さん	名詞	168
13	前	副詞可能	163
14	帰る	動詞	155
15	今	副詞可能	139
16	顔	名詞C	135
17	来る	動詞	131
18	考える	動詞	130
19	言葉	名詞	126
20	眼	名詞C	123

図. 夏目漱石「こころ」の抽出語リスト
<http://khcoder.net/> スクリーンショット集より

Voyant Tools

- <https://voyant-tools.org/>
- “Voyant Tools is a web-based text reading and analysis environment. It is a scholarly project that is designed to facilitate reading and interpretive practices for digital humanities students and scholars as well as for the general public.” – About – Voyant Tools Help.
<https://voyant-tools.org/docs/#!/guide/about> (accessed 2021-04-15)



ゲームタイトルの計量的分析

- 1980年代前半から2000年代前半までのビデオゲームのタイトルについての計量的分析。
- ゲームタイトルは経時的に長大化傾向。また、トレンドの変化が影響する。
 - 福田一史 et al.. 2017. “家庭用ゲームソフトのタイトルに関する研究：DBを活用した文字数・文字種の観点からみたマクロ的分析.” アート・リサーチ 17 (November): 29–44.
<http://doi.org/10.34382/00007765>.

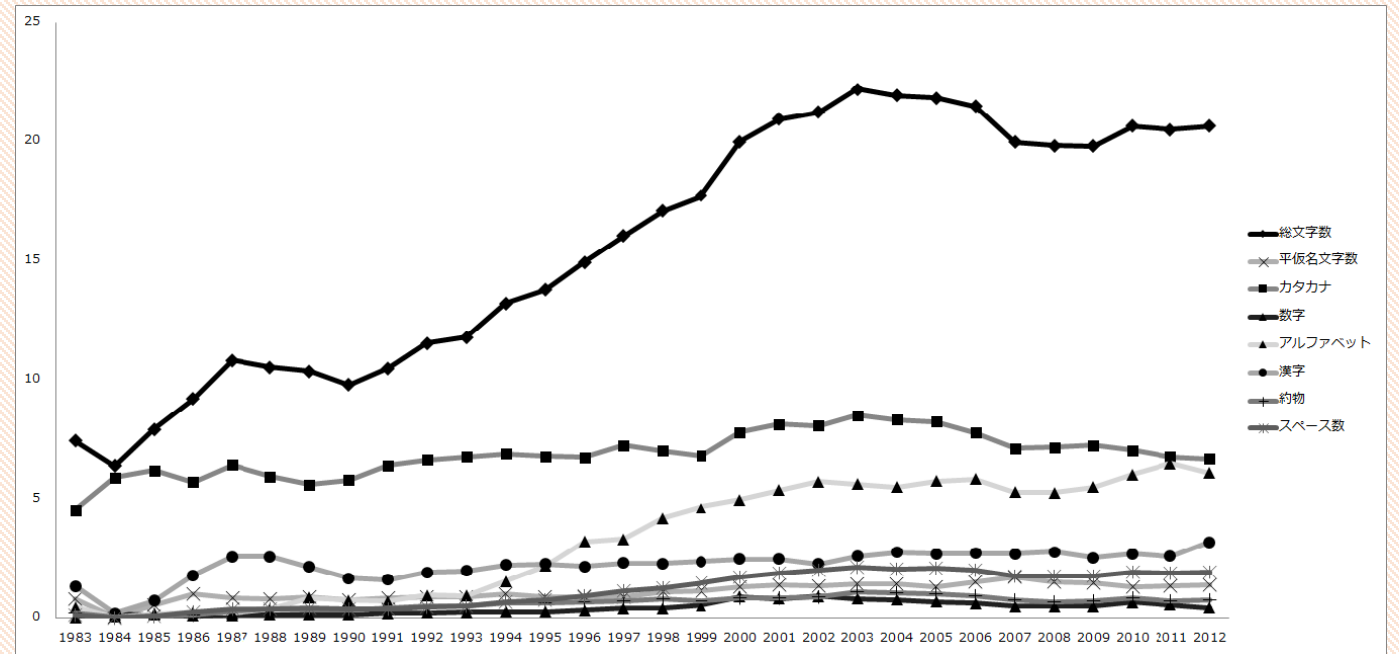


図. 文字種ごと平均文字数の推移

```

version="1.0" encoding="UTF-8"?>
model href="tei_all_ja_20191012_4.rnc" type="application/relax-ng-compact-syntax"
xmlns="http://www.tei-c.org/ns/1.0" xmlns:eaj="http://www.example.org/ns/ejaTEI"
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>走れメロス</title>
      <author>太宰治</author>
    </titleStmt>
    <publicationStmt>
      <istributor>青空文庫</istributor>
      <authority>金川一之</authority>
      <authority>高橋美奈子</authority>
      <date when="2011-01-17"> 2011年1月17日</date>
    </publicationStmt>
    <sourceDesc>
      <bibl>
        <author>太宰治</author>
        <title>走れメロス</title>
        <publisher>筑摩書房</publisher> 「太宰治全集3」 ちくま文庫、 <date when="
          >1988 (昭和63) 年10月25日</date>初版発行 <date when="
          >1998 (平成10) 年6月15日</date>第2刷 </bibl>
      </sourceDesc>
    </fileDesc>
    <revisionDesc>
      <list>
        <item>
          <date when="2011-01-17">2011年1月17日</date>修正 </item>
        <item> 入力<persName>金川一之</persName> 校正<persName>高橋美奈子</persName>
          <date when="2000-12-04"> 2000年12月4日</date>作成 </item>
        </list>
      </revisionDesc>
    </teiHeader>
    <text>
      <front>
        <p>
          <listPerson>
            <person xml:id="メロス">
              <persName>メロス</persName>
              <occupation>牧人</occupation>
            </person>
          </listPerson>
        </p>
      </front>
    </text>
  </teiHeader>

```

TEI

- Text Encoding Initiatives
- 人文学資料の適切なマークアップ（XML）のための仕様・標準、それを構築するためのプロジェクト。
- The TEI Guidelines
 - <https://www.tei-c.org/release/doc/tei-p5-doc/ja/html/index.html>
- aozora_tei wiki
 - https://github.com/TEI-EAJ/aozora_tei/wiki

図. マークアップサンプルファイル「走れメロス」.
https://www.dhii.jp/nagasaki/dazai_all_20191012.xml

XML: Extensive Markup Language

- 任意の目的に用いるための汎用性が高いマークアップ言語。
- 比較的人間にとって読みやすいデータ形式。
- APIのレスポンスや、MS Word、TEIなどのテキスト文書の構造化などに用いられる。
- 参照：“Extensible Markup Language (XML) 1.1 (Second Edition)”.
<https://www.w3.org/TR/xml11/>.

IIIF: International Image Interoperability Framework

- <https://iiif.io/>
- 「**画像へのアクセスを標準化し相互運用性を確保するための国際的なコミュニティ活動**」
 - IIIFを用いた高品質／高精細の画像公開と利用事例 | 人文学オープンデータ共同利用センター. <http://codh.rois.ac.jp/iiif/>
- もしくはそれらにより定義された**API**の仕様
 - [IIIF Image API](#)
 - [IIIF Presentation API](#)
 - など



図. IIIFロゴ

API

- Application Programming Interface
- プログラムコンポーネント同士を接続するためのインターフェイス、およびその仕様
- バックエンドの一部の機能の公開や共有
- ワールド・ワイド・ウェブを介するAPIは「ウェブAPI」と呼ばれ、HTTP要求とJSONなどによる応答で構成される。

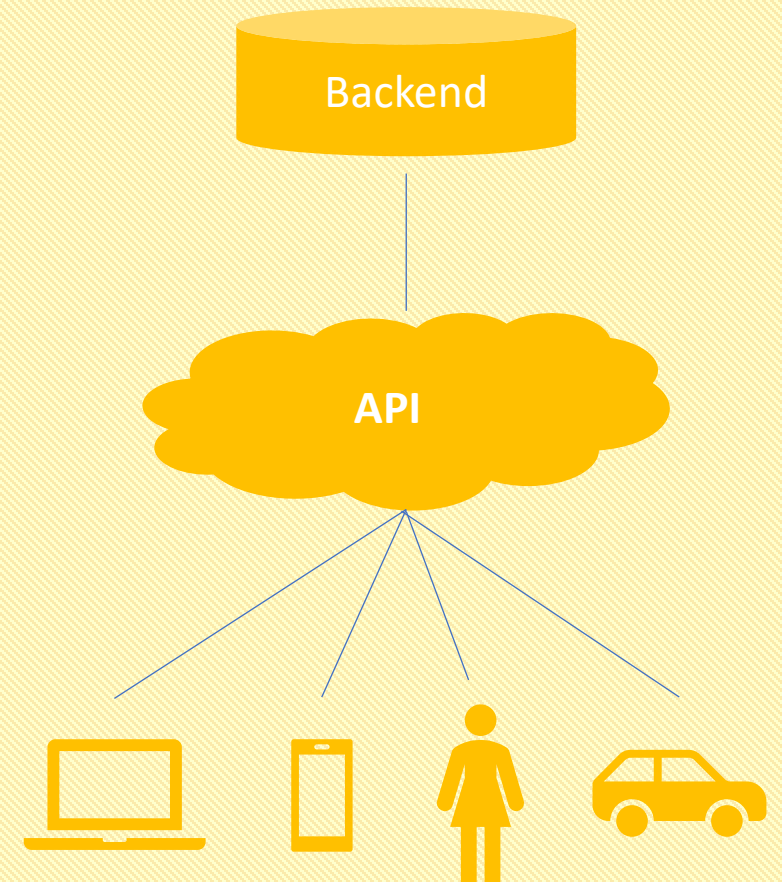


図. APIイメージ

API

- 「Web 2.0」 (2005～) の基礎要素としても注目される
- **APIファースト**
 - ウェブサービス開発において、まずAPIの設計や開発を進めてから、ビューや他のシステムの接続などの開発を進めるという開発法
- APIを提供することで、それを軸とするエコシステム (≒APIエコノミー) が形成される。それにより、サービスの価値向上につながる。
 - **ネットワーク外部性**が強く影響する

APIを提供するサービスの例

- 小売 : Amazon MWS API
- 金融 : 各銀行API
- 気象 : Open Weather Map API, Dark Sky API
- 地理 : Google Maps API, 電子国土web
- SNS : Twitter API, Facebook API
- 図書 : NDL (国立国会図書館) , OpenBD

JSON (JavaScript Object Notation)

- データ記述言語。
- 機械可読性の高い「軽量のデータ交換フォーマット」。
 - JSONの紹介. <https://www.json.org/json-ja.html>.
- 1) 名前と値 (Key-Value) の組み合わせ、2) 配列 (Array) 、 の構造で記録される。
- ウェブ上のデータ交換のための形式としてデファクトスタンダードのように広く用いられる。

Gettyと「どうぶつの森」

Getty

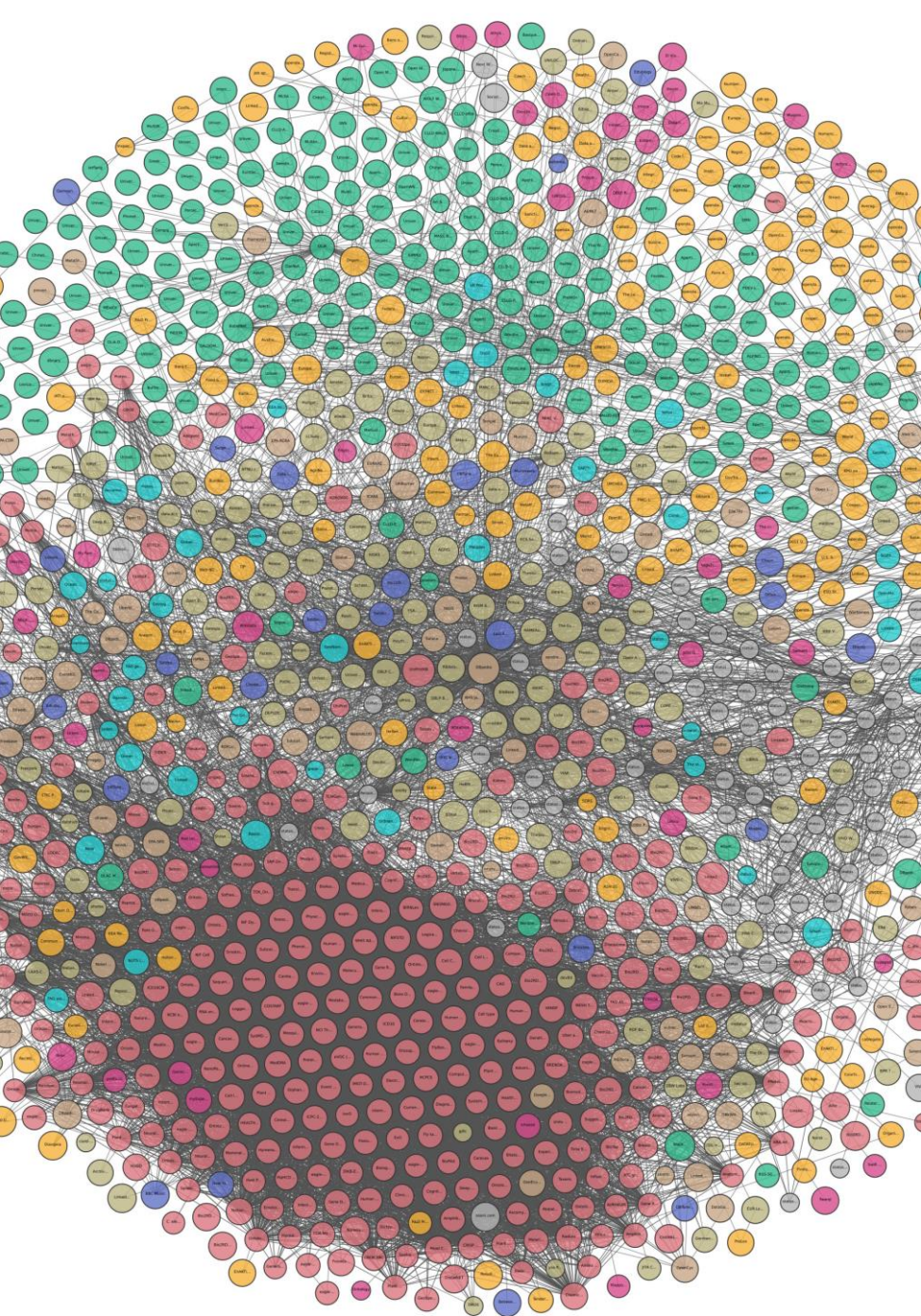
Animal Crossing Art Generator

Museum Research Institute Conservation Institute Foundation Support Us

Visit What's On Explore Art Resources Our Work About 



 Animal Crossing Art Generator <https://experiments.getty.edu/ac-art-generator#iiifloader> (accessed 2020-06-29)



Linked Data

- リンクト・データ
- ワールド・ワイド・ウェブを通じて機械可読性の高いデータを公開するための技術。
- 特にオープンデータとして公開するものを**Linked Open Data (LOD)**と呼ぶ。
- セマンティックウェブを構成するための主要技術の一つ。

Linked Data

- 文書のウェブからデータのウェブへの転換
- Tim Berners-Leeの提案するLinked Dataの原則
 1. Use **URIs** as names for things
 2. Use **HTTP URIs** so that people can look up those names.
 3. When someone looks up a URI, **provide useful information**, using the **standards** (RDF*, SPARQL)
 4. **Include links to other URIs**. so that they can discover more things.
 - Linked Data – Design Issues. <https://www.w3.org/DesignIssues/LinkedData.html>.
- より詳しくは以下論文など参照のこと
 - 大向一輝. 2013. “オープンデータ活用：1. オープンデータとLinked Open Data.” 情報処理 54 (12): 1204–10. <http://id.nii.ac.jp/1001/00095922/>.

機械学習

- KuroNetくずし字認識ビューア
- 人文学オープンデータ共同利用センターが提供する、ディープラーニングを用いたくずし字のOCRサービス。
 - <http://codh.rois.ac.jp/kuronet/>

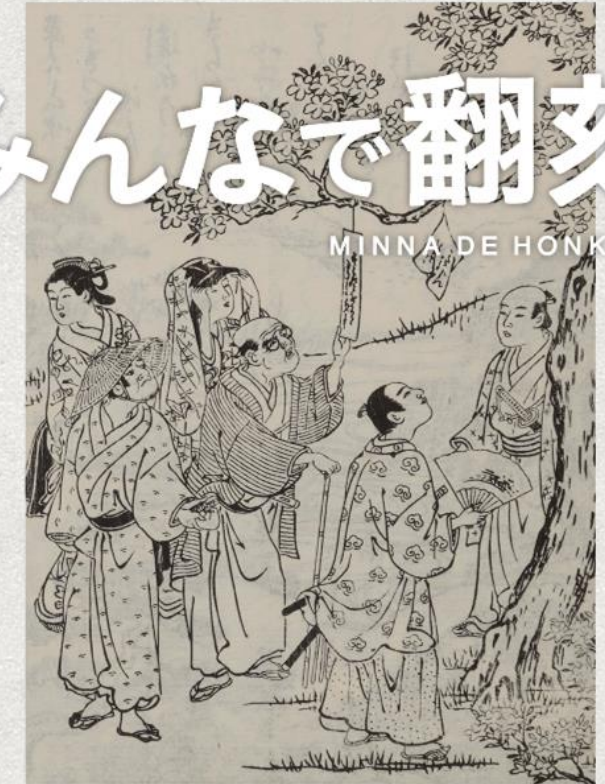
集合知・クラウドソーシング

- みんなで翻刻
- <https://honkoku.org/>
- 「すでに『みんなで翻刻』では、5000人の人々により600万文字もの史料が翻刻されています。」
 - <https://honkoku.org/>
- コミュニティ維持のために**ゲーミフィケーション**の方法論を採用。



みんなで翻刻

MINNA DE HONKOKU



と、AIと協力して、
書解読に参加しよう！

JOIN
参加する

※現在新バージョンの試験運用中です。
安定稼働するまで旧バージョンも併せてご利用ください。

いいね! ツイート

B!ブックマーク LINEで送る

DHの方向性

- 既存の文化研究の**ブラックボックス化を危惧**し、デジタル技術を用いて**横串を指す** / **透明化**をすすめる。
- **オープン化**や**標準化**により、文化資源へのアクセシビリティを向上させ、持続的に提供可能とし、**知識基盤**を確立する。
 - 特にオープンな標準や、コミュニティで維持するオープンソースを指向する。**方法論の共有地**（Methodological Commons）。
- アクセシビリティが向上することにより、これまで想定していなかった**新しい方法論や知見**が登場。
 - 「デジタル・ヒューマニティーズ入門」
 - デジタル・ヒューマニティーズー人文学と情報学の接点が導く新たな知識の世界。
<https://www.youtube.com/watch?v=aemYqznZM-8> (accessed 2021-04-15)