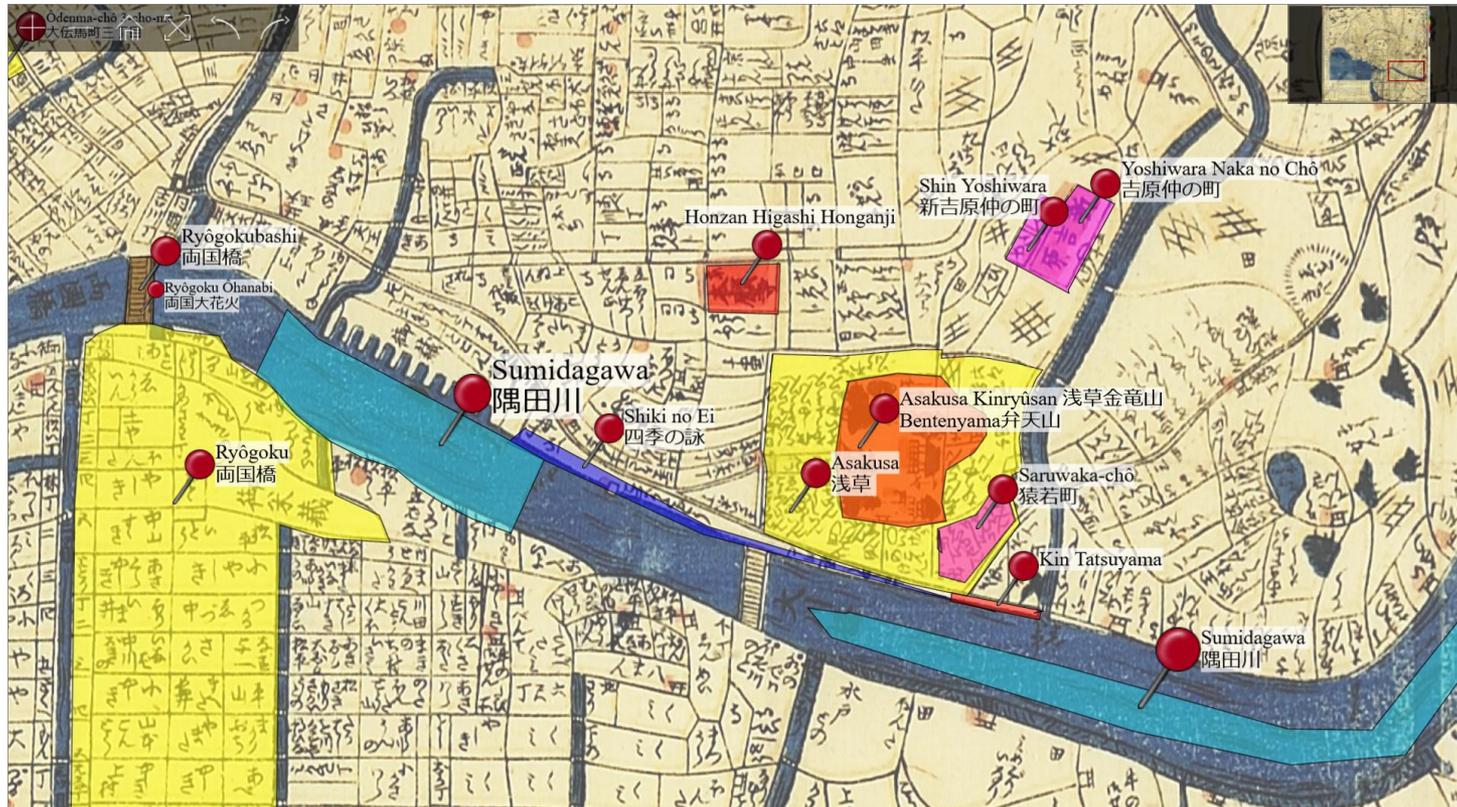


ドイツにおけるデジタル・ヒューマニティーズ と日本学における応用

Michael Kinski, Frankfurt



デジタル・ヒューマニティーズ / デジタル人文科学 (DH) とは何か

- 人文科学は、人間の社会、文化、言語、歴史、思考、コミュニケーションのあらゆる側面を扱って、個々の分野で構成されています。
- デジタル人文科学は、これらの関心領域を共有し、デジタル化されつつある労働環境をおよびメディアの世界の条件下で、新しい知識を獲得して伝えるプロセスをさらに発展させようとしています。
- デジタル人文科学は知識と文化遺産のデジタル化、分析用のデジタルツールを使用し、デジタル世の方法論的および理論的基盤の考察を行っています。

自慢話

- 2012年、江戸時代の思想史に関する学会で、ベティーナ・グラムリッヒ・オカ氏（上智大学教授）が、インタラクティブな人名録データベースを作成するプロジェクトの第一段階を紹介しました。彼女はこれを、**ソーシャルネットワーク分析**の一環として捉えていました。
- 当時、私は「デジタル・ヒューマニティーズ」や「DH」という言葉を聞いたことがなかったので、私も他の参加者も、グラムリッヒ・オカ先生が提案したことをよく理解できませんでした。
- それ以来、私はDHに関心を持ち、コンピュータサイエンスやコンピュータ言語学のバックグラウンドはありませんが、**テキストマイニング**や**トピックモデリング**、**セマンティックネットワーク分析**などを江戸時代の思想史研究に取り入れたいと考えています。

- 2017年から、同僚のB・Voigtmannと共に、フランクフルト大学の日本研究の学生たちの間、DHを屈指するアプローチへの関心を高め、文献分析に応用させようとしています。私たちは、特に重要な2つの分野、すなわち

テキストマイニングと 自動文字認識

に関心を持っています。前者は（私たちのささやかな能力の範囲で）授業でも研究でもある程度まで生かすことができましたが、後者はコンピュータサイエンスやAI、Deep Learningの知識がない私たちが貢献できる範囲をはるかに超えています。

- しかし、私たちは、編集されていない江戸時代の資料（フランクフルト大学は1650年から1850年の間の印刷された本の小さなコレクション、いわゆる「江戸文庫」を所有しています）を読めるようにするために、別のアプローチを取りました。
- コンピュータによる分析に備えるために必要な最初のステップとして、より大量のテキストを解読し、書き写すための共同グループ作業です。これは、「みんなで翻刻」プロジェクトがクラウドソーシングを利用して試みたことと似ています。しかし、私たちの場合、テキスト解読の作業は教室内で行われます。

クラスルーム・プロジェクト

- 解読の結果を発表し、原文と翻刻の両方を揃えて公開するために、あらゆるニーズに対応できるツールを求めました。その結果生まれたのがDemiScriptというアプリです。プログラミングを担当したのは、当時は学部生であったK. Birenheideで、現在は修士論文を書こうとしている方です。
- DemiScript: 文献を表示し、直接その上にコメントや説明文の列を追加するなど、様々な操作が可能です。このツールは、前近代・近代の文字に書かれた文献、画像を問わず、あらゆるソースを扱う上で多くの可能性を持っているので、学生の小グループが独立してプロジェクトを進める、より大規模なクラスルーム・プロジェクトのプラットフォームとして利用できるかと確信しました。
- しかしツールがいくらよくても、ほかに欠かせない側面が色々あります。

学生とIT

- 私がDHに興味を持ち始めたとき、デジタル・ヒューマニティーズを導入すれば、学生の間で好意的な反響が得られるだろうと考えていました。
- しかし、デジタル時代だと言われても、コンピュータやコンピュータを使った仕事のやり方を効率的に使いこなすことや、これまでにない方法で学習することに、学生たちが親近感を持っているわけではないことに気がつきました。
- 学生の大半は、江戸時代の日本を勉強するためにDHを利用するという私の突然の熱意を共有していないだけでなく、MS Wordやその他の一般的なアプリケーションを使用することに関しては、私よりも想像力・意欲に欠けている人が多いようです。

それでも

- 立命館大学のARCを強力をえて、2020年の冬学期に江戸時代の地図・広重の江戸名所を示す錦絵・そしてDemiScriptを使って学習プロジェクトを行いました。
- 「御江戸絵図」プロジェクトでは、コンピュータを使う江戸時代の日本の探求に対する抵抗感を克服し、学生たちが現代の日本ではなく、過去の日本についてもっと学びたいという内発的な動機を強化したいと考えていました。私の想像では、教官である私あまり介入することなく、また厳しすぎるタスクスケジュールに従うことなく、明確に定義されたプロジェクトに自主的に取り組むことで、
 - (1)通常の教室で経験するよりも創造的なエネルギーを解放することができ、
 - (2)江戸時代の一次資料を自分で探求することへの興味を育むことができる
- と思いました。

- 確かに、学生たちが成し遂げたことは、私の期待や希望とほぼ一致していました。もちろん、中にチャレンジしなかった人やおざなりになってしまった人もいました。しかし、多くの学生にとって、このプロジェクトは、楽しさ、洞察力、そして探求と発見の感覚をもたらした。
- その結果、この授業の期末評価では、ほとんどの学生が「前近代の日本について学ぶことは、自分にとってまったく新しい経験であり、非常に満足のいくものだった」と述べています。

- このプロジェクトは、研究プロジェクトとしてではなく、学生に一次資料に触れてもらうことを目的とした学習プロジェクトである。参加者のほとんどは、古文の授業で習うような古典文法や江戸時代の日本語を学んだことがありませんでした。また、江戸時代の文字、いわゆる「ヘンタイガナ」にも触れたことがありませんでした。
- 学生にとって地図を扱うことはそれほど大きな問題ではなく、最終的には地図に使われている平仮名の異体字（と数少ない漢字）を識別することに慣れていたと思います。同時に、このプロジェクトのもう一つの目的は、江戸時代の都市の概要、都市のインフラ、二次元の地図を使った地理の表示、そしてそのために編集者が選んだ方法を学生に伝えることができたと思います。

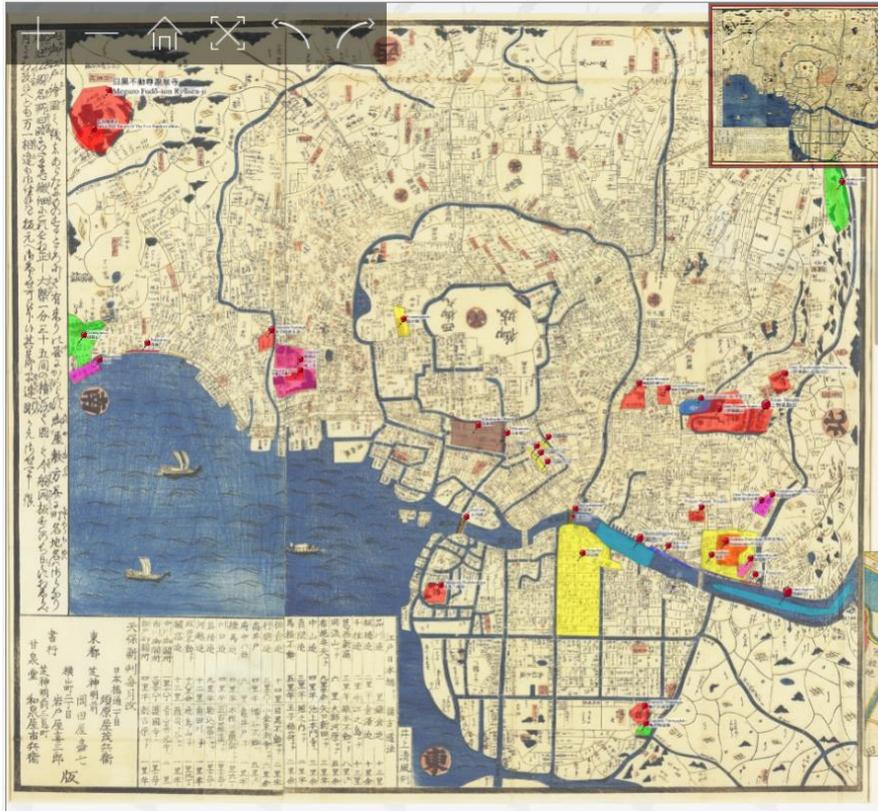
グループ内の作業の流れ

- 名所を絵図の上で確認
- 目印をつける
- 歴史、現状などを検索
- 説明文を書く
- 説明を大学のオンライン学習システムのOLATでWikiを作成しアップロード
- 編集
- 作文をDemiScriptの作文ツールに入れ名所の目印と連動

White Overlay Slider:

Map Rotation Slider:

Norton Password Manager



How to Use

Welcome to DemiScript!

You have the following options available to you when browsing the interactive *On Edo ezu*:

Layout

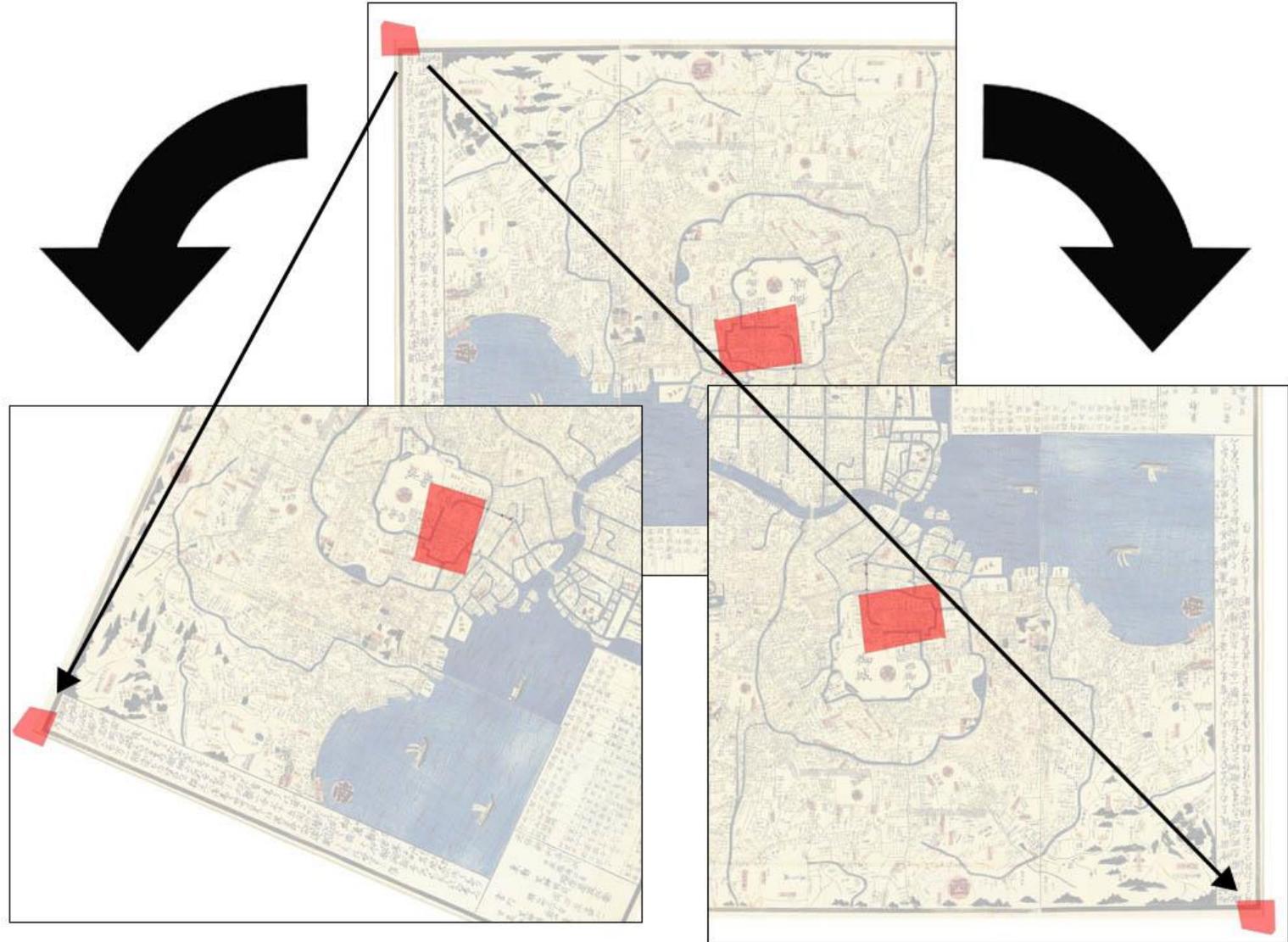
- The center left area is the interactive map.
- The center right area is the article and image document area.
- Located above the map are the overlay and rotation sliders.
- Located below the map is the map key.
- When hovering over the map, a navigation bar appears on the top left.
- When hovering over the map, a small full view appears on the top right.

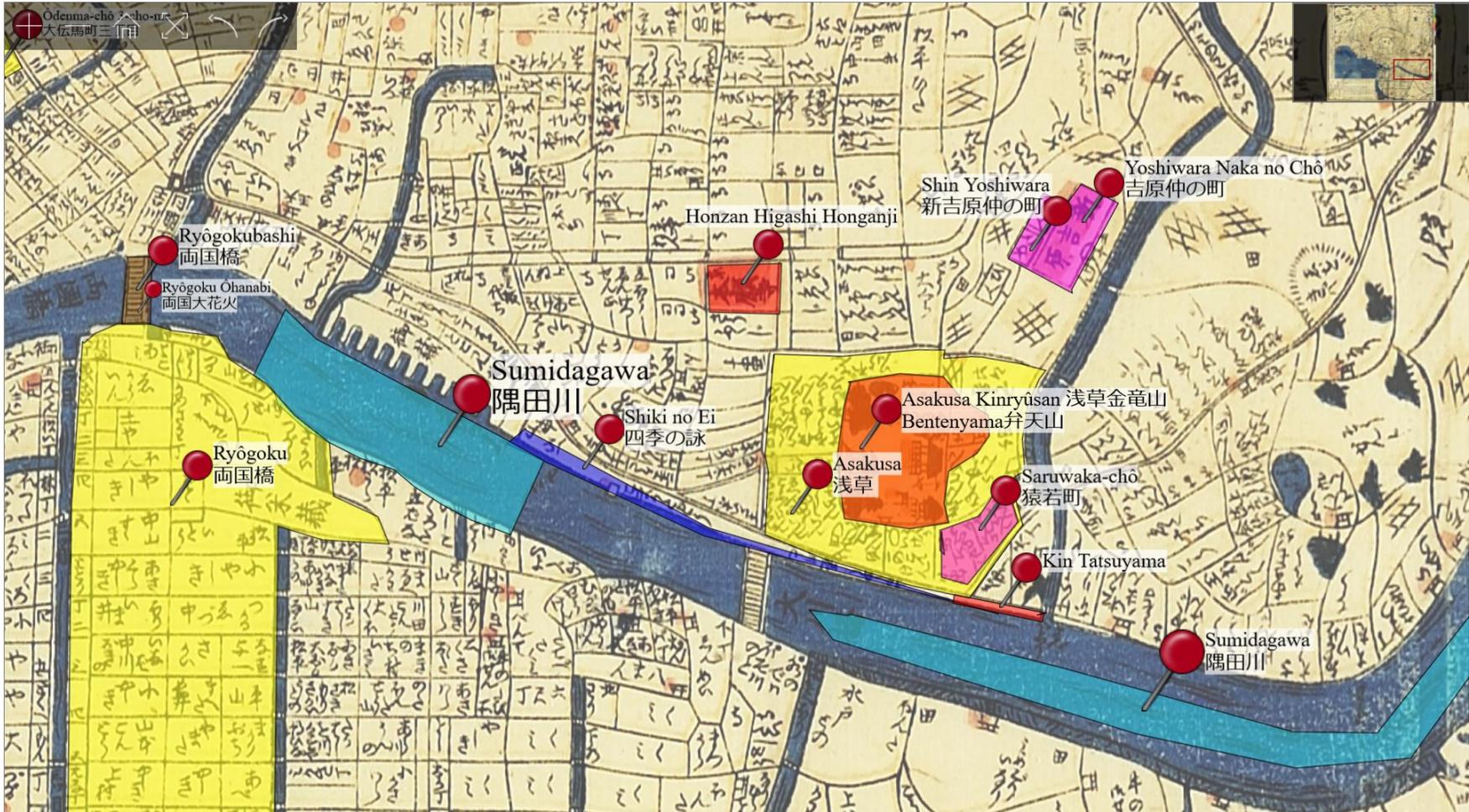
Controls

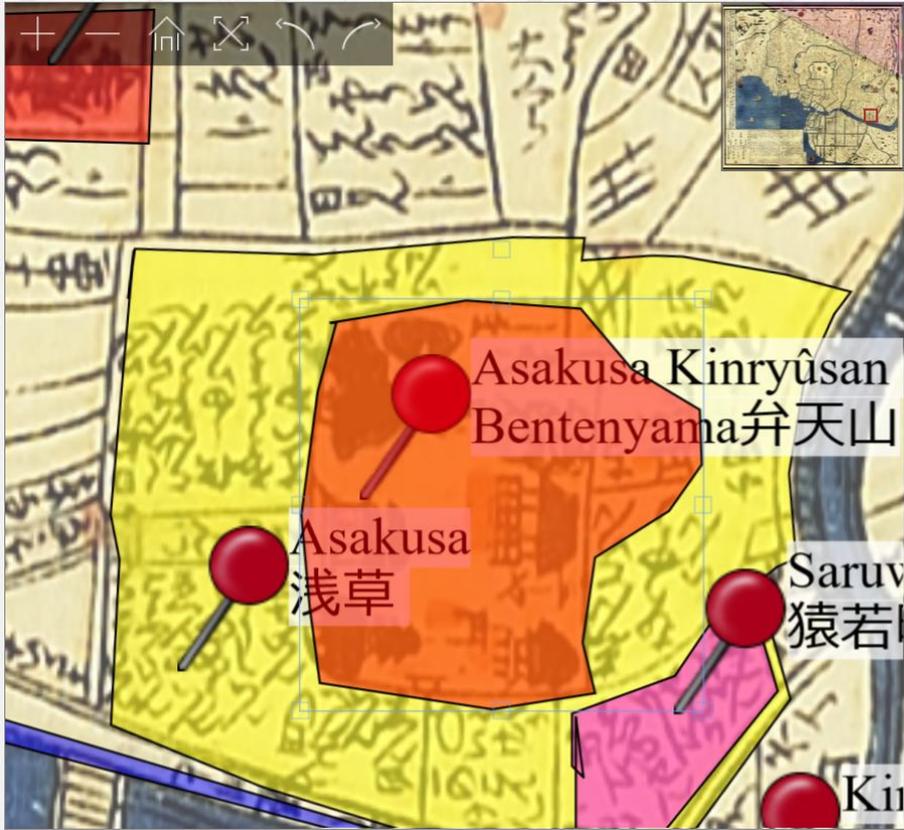
- Clicking on the map, pressing the "+" key, or clicking on the plus sign on the navigation bar will zoom into an area.
- Pressing the "-" key, or clicking the minus sign on the navigation bar will zoom out again.
- The "+" and "-" keys can be held down to continually zoom in or out.
- The mouse wheel can also be used to zoom in or out.
- Clicking and holding the left mouse button down allows the dragging around of the map.
- Using the arrow keys likewise moves the map around.
- Clicking on the rotation buttons in the navigation bar will rotate the map 90°.
- Clicking the home icon in the navigation bar will return the zoom level to default.
- Clicking the full-screen icon in the navigation bar will open the map in full-screen mode.
- Moving the overlay slider above the map to the left will make the overlay between map and map objects more transparent and hide all objects when set to 0.

Color Key: ■ Religious site ■ Park, garden, mountain, or hill ■ River or river bank ■ Bridge ■ Entertainment district ■ City ward ■ Other

Rotating Features



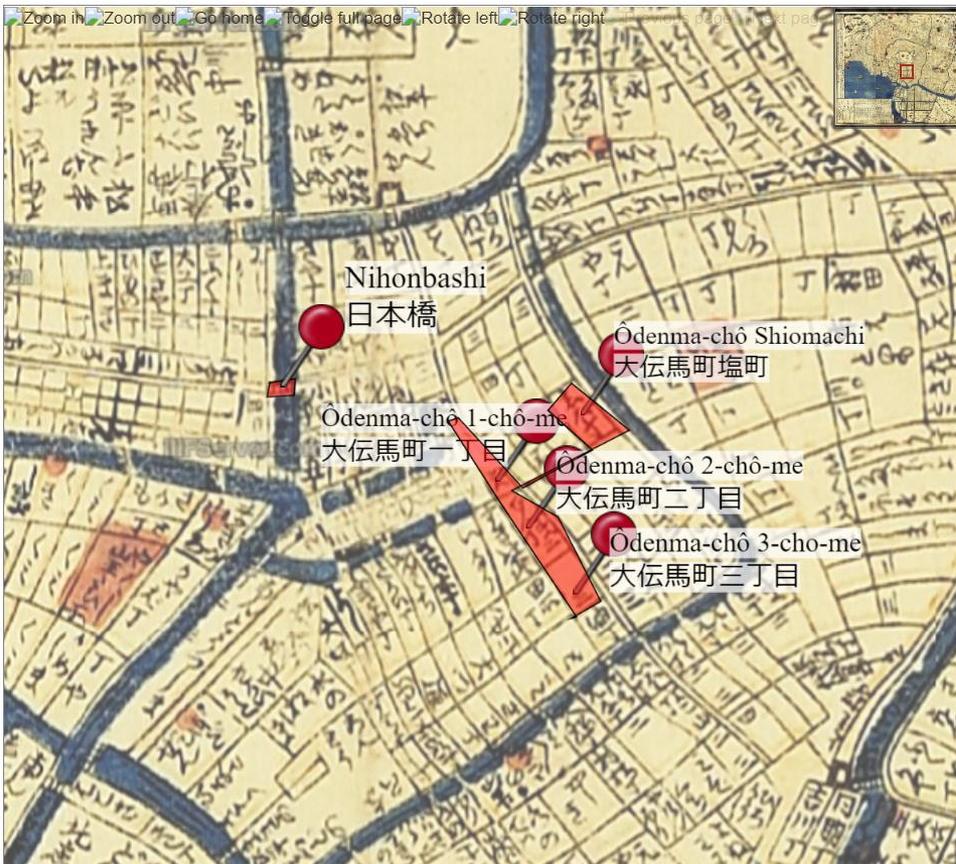




ARC Resource:
[「新撰江戸名所」, 「浅草金龍山観世音之図」](#)
Zoomable IIIF-Image:



セクション図と名所の紹介



Ôdenma-chô 大伝馬町

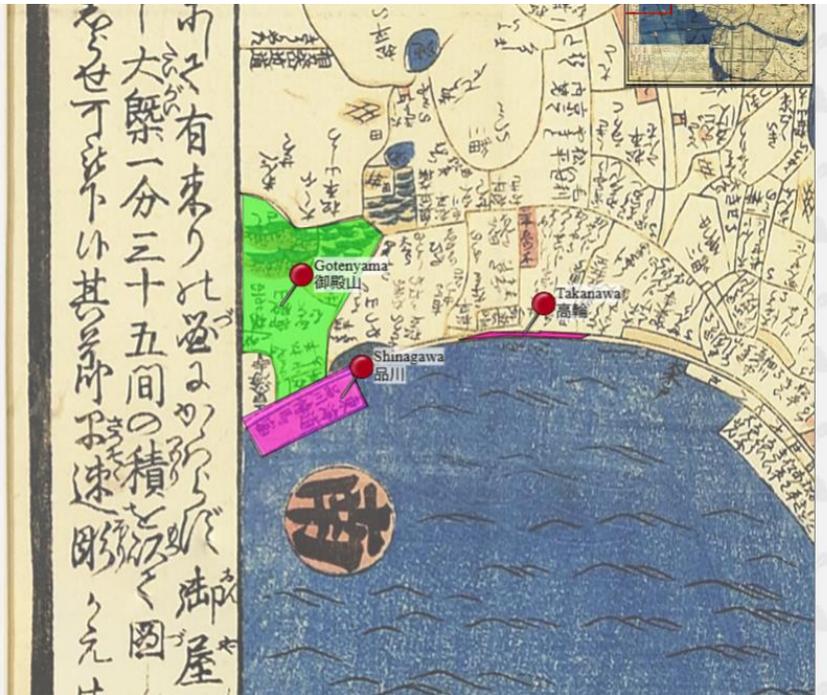
Ôdenma-chô 大伝馬町

• [Edit](#)

Written by Gruppe A (J11.1)
Category: [J11.1 - O Edo ezu](#)
Last Updated: 01 February 2021



Ôdenma-chô 大伝馬町 Ôdenma-chô1 war ein Stadtteile in Edo in der Nähe der Nihonbashi (日本橋), dem während der Edo-Zeit bis zu vier weitere Stadtteile untergeordnet waren. Bekannt war er hauptsächlich wegen der hohen Anzahl an Baumwollgroßhändlern, die sich im Stadtteil 1-chô-me2 versammelt haben. Neben anderen Großhändlern, befanden sich auch Musikgeschäfte, sogenannte Jôruri honya (浄瑠璃本屋), also Buchläden, die sich auf den Verkauf von Jôruri Stücken spezialisiert haben und viele weiteren Läden in Ôdenma-chô. Ein Teil des ehemaligen Stadtgebiets ist im heutigen Tokio im Distrikt Nihonbashi ôtenma-chô eingegliedert. Da viele Teile des ehemaligen Bezirks 1932 durch eine Stadtumplanung neu verteilt wurden, ist dieser Stadtteil jedoch nicht deckungsgleich mit Ôdenma-chô zur Edo-Zeit.



Gotenyama 御殿山

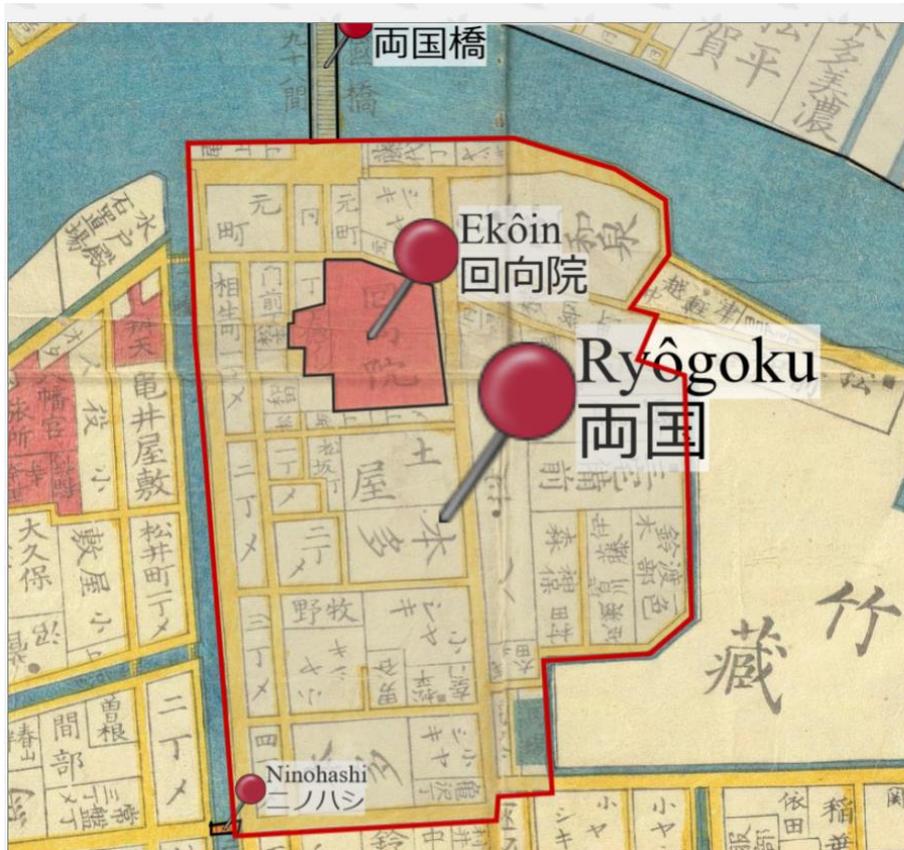
Gotenyama 御殿山 is an area and former popular public place in the northern Shinagawa District (北品川 *kita-shinagawa*) in Tōkyō. During the Edo Period it was known as one of the famous places of Edo (江戸名所 *Edo meisho*) because of its natural beauty and popularity for Hanami.

A Digital Project by Koray Birenheide (B.A.), 2021

Supervisor: Prof. Dr. Michael Kinski, Goethe University Frankfurt

Table of Contents

- [Digital Bunken Edo ezu](#)
- [Introduction](#)
- [About Saitô Gesshin](#)
- [The Lunisolar Calendar](#)
- [DemiScript](#)
- [Methods of Localizing Places](#)
- [Results and Conclusions](#)
- [Sources](#)
- [Contact Author](#)



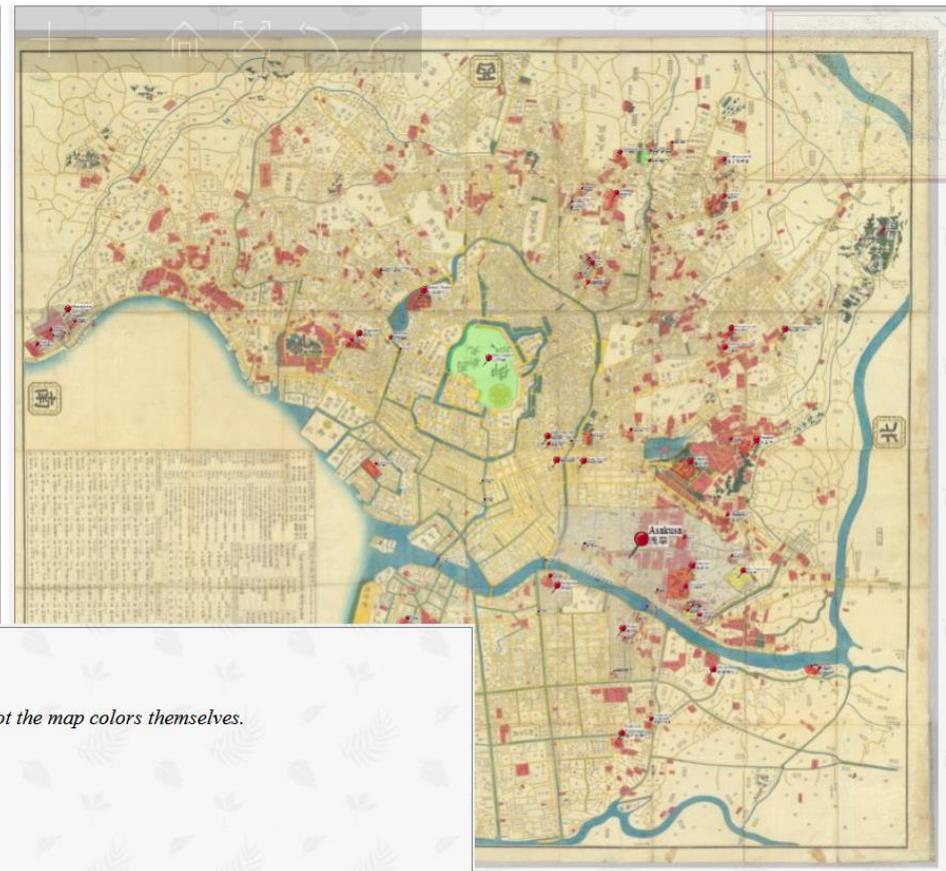
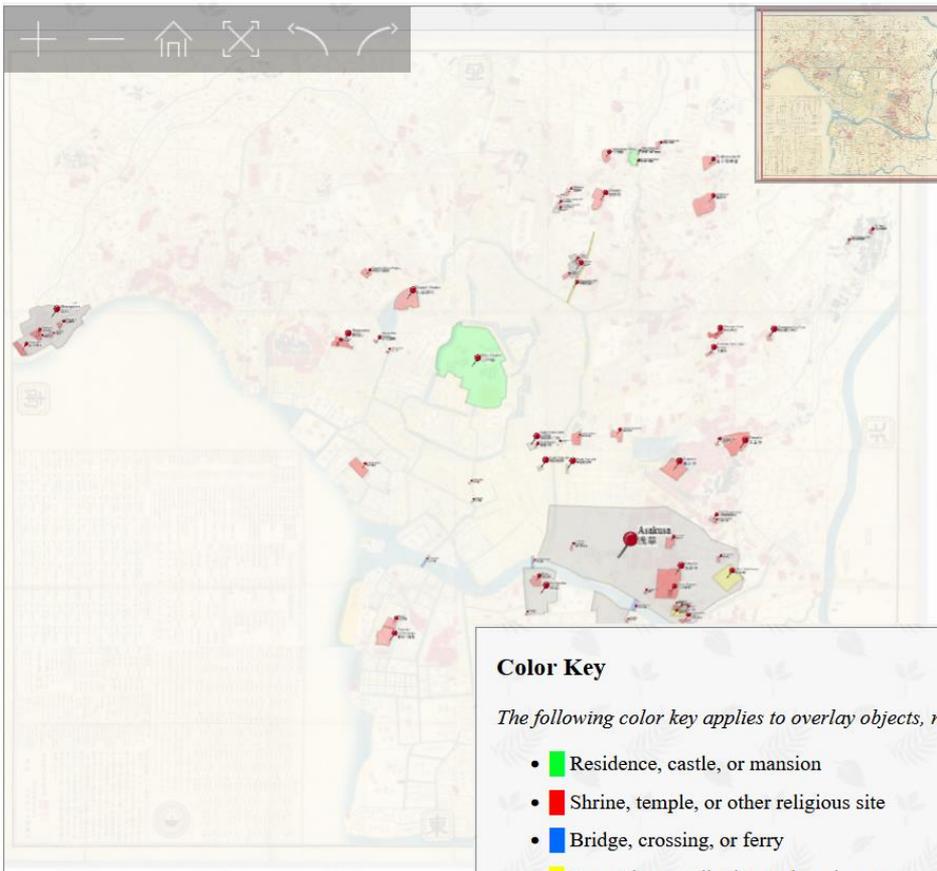
– [Tokyo NDL](#)

On the fourth day of the sixth month he visited fairs at Honjo; and on the sixteenth, twentieth, twenty-seventh, and twenty-ninth of the same month he went to such events at Ryôgoku.

– Nishiyama, Kazuo (1997): *Edo Culture : Daily Life and Diversions in Urban Japan, 1600-1868*, University of Hawaii Press. (p.85)

東都両国ばし夏景色 (Toto Ryogoku-bashi natsugeshiki)





Color Key

The following color key applies to overlay objects, not the map colors themselves.

- Residence, castle, or mansion
- Shrine, temple, or other religious site
- Bridge, crossing, or ferry
- Entertainment district or shopping street
- City ward or neighborhood

ドイツにおけるDH

- 一般的様子

理解度の問題

- 概していえばドイツの学术界ではデジタル・ヒューマニティーズについて誤解や偏見が根付いています
- その1 デジタル・ヒューマニティーズはつまりデータバンクを作ることの意味している
- その2 デジタル・ヒューマニティーズはズルい

つまり本業を怠けた偽りをする作業

- Franco Morettiのいう „distant reading vs. close reading“ という役割分担を認識せずに、ほとんどの知識なくデジタル・ヒューマニティーズを軽んじる傾向が受け止められます

...その反面

- 連邦政府による支援（科研費）
- 州政府のレベルでの支援（科研費）
- DFG (Deutsche Forschungsgemeinschaft = 科学研究費を提供する公の機関) による支援

Bundesministerium für Bildung und Forschung (BMBF)
ドイツ連邦の文部科学省

„Digital humanities: the rise of a new academic discipline“ (2018)

Research in
Germany



Land of Ideas

支援プログラムの3つの側面

- **Computerised access to sources:** Manuscripts, images and archaeological artefacts are scanned and prepared for online access. The Cuneiform Digital Library for example contains tens of thousands of cuneiform scripts from collections worldwide, complete with transcriptions of the inscriptions and information about where they were found and their dating. This is a project that was initiated back in the 1990s by the Max Planck Institute for the History of Science in Berlin and the University of California in Los Angeles.
- 資料・研究対象へのコンピュータ化されたアクセス:原稿、画像、考古学的遺物がスキャンされ、オンラインアクセス用に準備
- 例えば楔文字のデジタルライブラリ作成

Digital humanities: the rise of a new academic discipline

- **Data mining and visualisation:** data analysis and data processing are other fields in the digital humanities. Mannheim & Heidelberg
- **データマイニングと視覚化: データ分析とデータ処理**
- **例** Leibniz Science Campus Empirical Linguistics & Computational Language Modeling: 自然言語処理モデルの導入を通じてドイツ語の大規模コーパスリソースの高品質な自動注釈をサポートする革新的な研究を実施し、言語学における高度な実証的研究を可能にすることを目的とする

BMBF 2018:

Digital humanities: the rise of a new academic discipline

- **Digital research environments for humanities and social science scholars:** The [TextGrid](#) platform was the first virtual research environment to be set up for the humanities in Germany. Funded by Germany's Federal Ministry of Education and Research (BMBF), the platform has been assisting researchers, particularly with the creation of digital editions, since the summer of 2011.
- 人文科学および社会科学の学者のためのデジタル研究環境: デジタル版の作成で研究者を支援するためのバーチャル研究環境を設置 例 TextGrid platform (2011より支援)

BMBF 2018:

Digital humanities: the rise of a new academic discipline

**DARIAH-DE Digital Forschungsstruktur für die
Geistes- und Kulturwissenschaften**

人文科学および文化研究のためのデジタル研究構造

- Text Grid  DARIAH-DE
- <https://de.dariah.eu/en/web/guest/dienste-und-werkzeuge>

- **CosmoTool**は、さまざまな資料からの伝記情報を歴史的人物の国内および外国における移動プロフィールに統合するデジタルツールです。
CosmoToolは「DARIAH-DEフェデレーションアーキテクチャ」に基づいており、非構造化テキストからデータを抽出できます。現在、CosmoToolは開発段階にあり、まだ機能が制限されています。
- <https://cosmotool.de.dariah.eu/cosmotool/personsearch/>



Johann Wolfgang von Goethe

* 28. August 1749

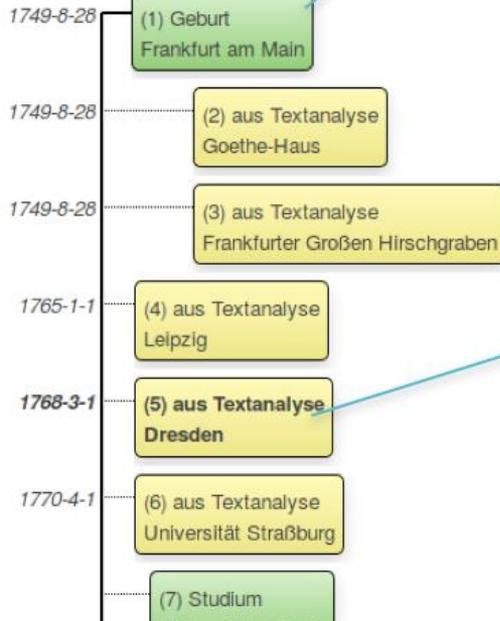
† 22. März 1832

cosmotool / Biographische Daten / Johann Wolfgang von Goethe

Biographische Daten

from structured data

Zeitleiste



Ereignis-Details

Ereignis 5

aus Textanalyse: Dresden

Textstelle: "Auf Oesers Empfehlung besuchte er im März 1768 Dresden und die Gemäldegalerie."

Quelle: <http://de.wikipedia.org>

Extracted from full-text



連邦政府の科研費機構 DFG

- Beispiel: Transkribus



宣伝 Transkribus = 文字起こしのツール

- 「歴史的文献の文字起こし・翻刻を行いますか？中世からの文献でも、20世紀の文献でも高度に標準化された、柔軟で信頼性の高い方法で翻刻を実行したいのですか？また、手書きのテキスト認識やレイアウト分析などの自動ツールのサポートを利用したいのですかか？」
- 「それならTranskribusはぴったり」

Transkribusのインターフェイス

The screenshot displays the Transkribus software interface. At the top, there is a toolbar with various icons for navigation and editing. Below the toolbar, the document viewer shows a handwritten page of text. To the left of the document viewer, there is a metadata table and a tags table. The metadata table lists various tags and their corresponding values. The tags table lists the tags used in the document and their colors and shortcuts. At the bottom of the interface, there is a transcription output area showing the transcribed text with red arrows indicating corrections or annotations.

Tag	Value	Text
1	date	17.1.6. [242] 17.1.6. Rilke Lieber
2	person	Rainer M. Rilke Lieber Herr Rainer M
3	abbrev	R. Herr Rainer Maria R. Ich h
4	place	Ham- gleich an die beste Ham-
5	place	burger die beste Ham- burger St
6	organization	Firma die Antwort der Firma. Es
7	place	Lokal ein ähnliches Lokal wie da
8	organization	Keller & Lokal wie das von Keller &
9	place	Berlin Keller & Reiner in Berlin, <
10	organization	Kunstge Ihren Brief an das Kunstg
11	organization	haus Hu Kunstgewerbe- haus Hult
12	textStyle	zu Hänc Kunstgewerbe- haus Hult
13	textStyle	des Her Hulbe zu Händen des Her
14	person	Herrn A zu Händen des Herrn Agt
15	textStyle	Matinée den Vortrag als Matinée u
16	textStyle	Sonntag und auf einen Sonntag ar

Tag	Color	Shortcut
abbrev	Red	Ⓜ
add	Green	Ⓜ
Address	Purple	Ⓜ
blackening	Black	Ⓜ
date	Blue	Ⓜ
div	Dark Blue	Ⓜ
gap	Light Blue	Ⓜ
organization	Pink	Ⓜ
person	Light Green	Ⓜ
place	Purple	Ⓜ
etc	Yellow	Ⓜ

12 Doch möchte ich Ihnen in keiner Weise vorgreifen, teile ↵

13 also der Firma Ihre Adresse nicht mit, sondern über ↵

14 lasse Ihnen selbst die Anknüpfung der Verhandlungen; ↵

15 Sie würden dann gut tun, Ihren Brief an das Kunstgewerbe ↵

16 haus Hulbe zu Händen des Herrn Agte zu adressieren. ↵

17 Ich empfehle Ihnen noch, den Vortrag als Matinée und ↵

18 auf einen Sonntag anzusetzen; das entspricht am besten ↵

... 定期的に行われる学会

学会: ドイツ語圏における デジタル・ヒューマニティーズ

Programm Tagung Tickets Standorte Call for Papers

DHd 2019



multimedial &
multimodal



6. Jahrestagung
des Verbands
Digital Humanities im
deutschsprachigen Raum

Frankfurt &
Mainz



25.03. —
29.03.2019



DhD 学会、2019年

- 会議はメディアの概念とデジタル人文科学におけるその内訳に取り組むことに焦点を当てることを意図しています。
- 目的は、デジタル研究対象への多面的アクセスと分析の可能性と制限について議論することです。
- 焦点は、モデリング（エンコーディング、インターフェイス、注釈）、分析（分類、セグメンテーション、ネットワーキング）の理論的および方法論的基盤（インター/クロスモダリティ、メディア理論）、合成（インタラクティブ3D再構築、設計主導型アプローチへの統合、インフラストラクチャの概念）に関する問題に置く
- もう一つの主題は、調査方法のアルゴリズムの基礎を議論・検討することです。

DARIAH-DEがまとめたデータによると14ヶ所の大学がDHを専攻にする学士課程・修士課程を提供する

- Bamberg
 - Bielefeld
 - Darmstadt
 - Erlangen
 - Frankfurt
 - Gießen
 - Göttingen
 - そしてTübingen
- Graz
 - Groningen
 - Hamburg
 - Köln
 - Lüneburg
 - Saarbrücken
 - Würzburg

日本学とDHは？

- 概していえば日本学におけるDHの受容は研究・学習両方面で遅れています

フランクフルト大学の日本学科におけるデジタル・ ヒューマニティーズワークショップ・シリーズ 第4号

TALKS + WORKSHOP
SOCIAL NETWORK ANALYSIS

KINSKI
HAERPFER
BINGENHEIMER
VOIGTMANN
BEERENS
BORN
ITO

DIGITAL
APPROACHES
4000

20./21.08.
NEUE MENSA KIII

01000100
011010010110011
011010101100110001
01101111000110101
1000000100101001100001011
1011011011011011011011
01101001011001010010000001
0001111010110001
10111011011010001
0011101011110010
01110100

DIGIJAPPOFFM
UNIVERSITÄT
FRANKFURT AM MAIN

GOETHE
UNIVERSITÄT
FRANKFURT AM MAIN

日本学

REGISTRATION REQUIRED @ DIGIJAPOFFM@EM.UNI-FRANKFURT.DE

一人の教授のコメントの要点

- 日本学の分野でも研究や授業において「デジタルなアプローチ」を使う試みがあるよう
- フ大日本学科のワークショップはその一つの表れ
- 目的は本来日本学の使命であるべき文献の解読や解釈ではなく、デジタルな技法による social network analysis
- 危険の1 “deep reading”による文献との取り組みが蔑ろにされるという恐れ
- 危険の2 若い研究者が自分の努力を以って文献を解明する能力が伸びないという恐れ
- 結論 将来の日本学にどれが大事だろう？
コンピューターによる social network analysis？それとも「文献解読の努力」？

日本学におけるDH

- Heidelberg大学 > Judit Árokay教授
- 近世日本の旅行記を使用して、文学的空間と場所の地理的視覚化

このプロジェクトでは、文化研究の地形的変化(topographical turn)を取り上げ、テキスト内の空間と場所を視覚化するための表現方法を開発します

日本学におけるDH

- Tübingen大学 > Robert Horres教授
- 日本の伝統的色彩

辞書式の視点と歴史的な色の問題の紹介から出
発して、日本の伝統的な色の分野を考察する。
伝統的な色彩システムの象徴性、使用状況、社
会的エンコーディングは参照システムとして理
解されており、今日までの日本社会のさまざま
な状況で効果的

フランクフルト大学・日本学科

- 最初に紹介した教育プロジェクトが成功したからといって、一般の学生がDHのアプローチを使って勉強することに興味を持つようになったり、DHに特化した授業に多く参加するようになったわけではありません。
- ビジュアル的にも面白く、遊び半分の学習クラスはうまくいきますが

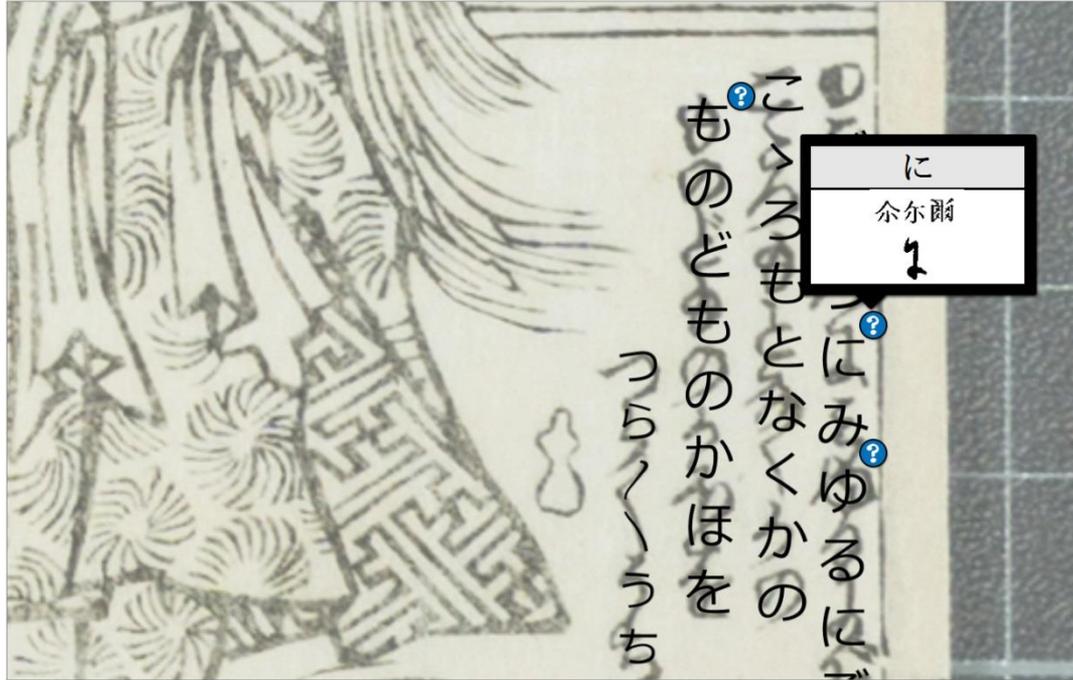
文字起こし

Add Line or Edit Line Directly.

Export Image

セリフ

セリフ



Sortable Text Segmented Text TEI Formatted



3 つまききょうに
5 獅子しし
6 ぶこつにみゆるにぞ
7 こころもたなくかの
8 ものどものかほを
9 つらくうち
10 ながむるにこほこれ
11 まへのやくしやに
12 あらでちほとゆうきの
13 両にんなるにぞうち
14 おどろきたるち氏が
15 たちまちこゑをふり
16 たて、「りよぐわい
17 なり氏とちさだ
18 たねなんぢら
19 ふたりはさきたたつて
20 みさほうてながへとより
21 してひやうきといひ
22 たてしゆつしさへいた
23 さざる、身であり
24 ながらでんがとど
25 にうちまぎれわが
26 めどほりへすい
27 さんせしは
28 しゆゑんの
29 さまたけ
30 なさんが
31 ためか
32 きりく
33 そこそ
34 たてばよしさなくは
35 めにものみすべき
36 そといわれてくだんの
37 両にんはかぶりしもの
38 とりのけつ、いしやうの
39 うちにかくしもつたる
40 大小とつてこゝにたほ
41 さみおんまへちかく
42 めかづきたる、なかに
43 氏とちかうべを
44 ちたげ
45 いたる
46 二のまきへ
47 一の巻より
48 こと
49 ながら
50 いかにも
51 さま
52 たげ

- しかし、本来の目的である文献分析の環境づくりになると
- ちなみに ...
- word2vec algorithm is one of the most commonly used algorithms for creating word vectors
- ... the word-context structure, which includes word vectors, uses the **distributional hypothesis** that '**words in similar contexts tend to have similar meanings**' (Turney & Pantel 2010: 143).
- ... the algorithm takes in a corpus of texts and represents words as points in a multi-dimensional space, and word meanings and relationships between words are encoded as distances and paths in that space, through the creation of an artificial neural network (2013).
- The created model is a simple, shallow neural network which encodes not only **syntactic but also semantic relationships between words**. Like topic modelling, **word vectors aim to reveal the underlying structure of a text or corpus of texts**. However, unlike topic modelling, they allow us to ask 'what does this corpus say about this theme?'. (S. Kerr 2016: 42)

Scraping



[Startseite](#) | [Beschreibung](#) | [Gesamtliste](#) | [Suche](#)

Lexikon Japans Studierende - Beschreibung

Japans Studierende in Deutschland 1868–1914

Vorbemerkungen

Inhalt

1. Vorbemerkungen
2. Hauptstudiengebiete



Notebook Basics

This lesson assumes that the user has Jupyter [installed](#) and that the notebook server can be started by running:

```
jupyter notebook
```

For more details on how to run the notebook server, see [Running the Notebook Server](#).

The Dashboard

When the notebook server is first started, a browser will be opened to the notebook dashboard. The dashboard serves as a home page for the notebook. Its main purpose is to display the portion of the filesystem accessible by the user, and to provide an overview of the running kernels, terminals, and parallel clusters.

Files Tab

The files tab provides an interactive view of the portion of the filesystem which is accessible by the user. This is typically rooted by the directory in which the notebook server was started.

The top of the files list displays clickable breadcrumbs of the current directory. It is possible to navigate the filesystem by clicking on these breadcrumbs or on the directories displayed in the notebook list.

A new notebook can be created by clicking on the **New** dropdown button at the top of the list, and selecting the desired language kernel.

Notebooks can also be uploaded to the current directory by dragging a notebook file onto the list or by clicking the **Upload** button at the top of the list.

@@ -29,7 +29,7 @@

```

29 29     "cell_type": "markdown",
30 30     "metadata": {},
31 31     "source": [
32 32 -     "![wideview](../figs/wideView400.png)"
32 32 +     "![wideview](../images/wideView400.png)"
33 33     ]
34 34 },
35 35 {

```

@@ -58,7 +58,7 @@

```

58 58     "source": [
59 59     "This leads to a usage model where you can create as many engines as you want per compute node, and then control them all from your clients via a c
60 60     "\n",
61 61 -     "![cluster](../figs/ipcluster-kernels.png)"
61 61 +     "![cluster](../images/ipcluster-kernels.png)"
62 62     ]
63 63 },
64 64 {

```

@@ -186,7 +186,7 @@

```

186 186     "cell_type": "code",
187 187     "execution_count": 1,
188 188     "metadata": {
189 189 -     "collapsed": false
189 189 +     "collapsed": false
190 190     },
191 191     "outputs": [],
192 192     "source": [

```

@@ -199,7 +199,7 @@

```

199 199     "cell_type": "code",
200 200     "execution_count": 2,
201 201     "metadata": {
202 202 -     "collapsed": false
202 202 +     "collapsed": false

```

- 明治時代の日本人留学生の生没年の抽出

```
for student in data:
```

```
    dates = student['Daten'].strip()
```

```
    print("Birth and death dates:", dates)
```

```
    date_list = dates.split("-")
```

```
    if len(date_list) == 2:
```

```
        first_date = date_list[0]
```

```
        second_date = date_list[1]
```

```
        if first_date != "":
```

```
            try:
```

```
                birth_date = dateparser.parse(first_date)
```

```
                print("Birth date:", birth_date)
```

```
                print("Birth year:", birth_date.year)
```

```
                birth_year_list.append(birth_date.year)
```

```
            except:
```

```
                print(first_date, "cannot be parsed!")
```

```
                unparseable_dates += 1
```

```
                unparseable_dates_list.append(first_date)
```

```
        if second_date != "":
```

```
            try:
```

```
                death_date = dateparser.parse(second_date)
```

```
                print("Death date:", death_date)
```

```
                print("Death year:", death_date.year)
```

```
            except:
```

```
                print(second_date, "cannot be parsed!")
```

```
                unparseable_dates += 1
```

```
                unparseable_dates_list.append(second_date)
```



データベース

人物

イベント

出典

プロジェクト



< 前

題名 -

責任者 (PI) -

JBDB

Bettina Gramlich-Oka, Sophia University

Hartmann Lexicon (seminar project)

Bastian Voigtmann, Frankfurt University

2件中1-2件目

<< < 1 > >>



阿部磯雄 / Abe Isoo* [ID: 10270] の情報

基本情報

経歴 (0)

親族関係 (0)

非親族関係 (0)

関連するイベント (0)

名前: 阿部磯雄

ふりがな:

ローマ字: Abe Isoo

性別: 男

生年: 元治二年二月四日

没年: 昭和二十四年二月十日



プロジェクト:

- Hartmann Lexicon (seminar project)

出典:

- Japanische Studenten an deutschen Universitäten und Hochschulen 1868-1914

[← データベースに戻る](#)

- 今学期は
 - Text mining
 - Rstudio
 - Quanteda

Replication: Text Analysis with R for Students of Literature

Kenneth Benoit, Stefan Müller, and Paul Nulty

Source: `vignettes/pkgdown/replication/digital-humanities.Rmd`

In this vignette we show how the **quanteda** package can be used to replicate the analysis from Matthew Jockers' book *Text Analysis with R for Students of Literature* (London: Springer, 2014). Most of the Jockers book consists of loading, transforming, and analyzing quantities derived from text and data from text. Because **quanteda** has built in most of the code to perform these data transformations and analyses, it makes it possible to replicate the results from the book with far less code. Throughout this vignette, we name objects based on Jockers' book, but follow the **quanteda** [style guide](#).

In what follows, each section corresponds to the respective chapter in the book.

1 R Basics

Our closest equivalent is simply:

```
install.packages("quanteda")
install.packages("readtext")
```

But if you are reading this vignette, than chances are that you have already completed this step.

2 First Foray

2.1 Loading the first text file

```
library(quanteda)
```

We can load the text from *Moby Dick* using the **readtext** package, directly from the [Project Gutenberg website](#).

```
library(readtext)
data_char_mobydick <- texts(readtext("http://www.gutenberg.org/cache/epub/2701/pg2701.txt"))
names(data_char_mobydick) <- "Moby Dick"
```

意思表示

- It is my dream to help bring about a generation of young scholars of Japanese Studies to take the problem and hurdle hiding behind the expression „the Great Unread“ seriously and who will endeavour to develop a methodological foundation for new approaches to the study of Japanese literature and history of thought ... thereby overcoming the restrictions that necessarily follow on focusing on too small a sample of source materials.
- 私の夢は、「偉大なる未読者」という表現に隠された問題やハードルを真剣に受け止め、日本文学や思想史研究の新しいアプローチのための方法論的基盤を開発しようとする若い日本研究者の世代をもたらすことです。それによって、少なすぎる資料のサンプルに焦点を当てることで必然的に生じる制約を克服することができる。