

# edomi構築を通じた 歴史資料と実世界の結合



ROIS-DS人文学オープンデータ共同利用センター

鈴木親彦  
特任助教

2022年03月28日 第16回CODHセミナー

# アジェンダ

全体の内容：

一つ目の講演で北本センター長が示した概念の実践の一例

発表の流れ：

1. 「都市」江戸に関する情報の歴史資料からの抽出
2. 資料情報と実体情報の間にあるギャップ
3. 「文書空間」と「実体空間」の双方向結合
4. edomi 江戸をみる／みせるデータポータル
5. 歴史ビッグデータ構想に向けての展望

# 「都市」江戸に関する情報の抽出

- 国文学研究資料館、国立国会図書館他の様々な機関から人文学研究・歴史研究に利用できる資料がデジタル化されて公開
- 再利用・再編集が容易になった＝デジタルの大きな利点の一つ
- デジタル化された人文学資料から、「部分」を特定の粒度で抽出・整理し直して研究活用可能にする

→ 「**人文学資料マイクロコンテンツ化**」

※電子書籍の「マイクロコンテンツ」を転用した造語



# 江戸観光案内

Q : 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) / 俳優似顔東錦絵 (62) / 絵本江戸桜 (23) / 絵本江戸の見図 (21) / 絵本江戸錦 (17)

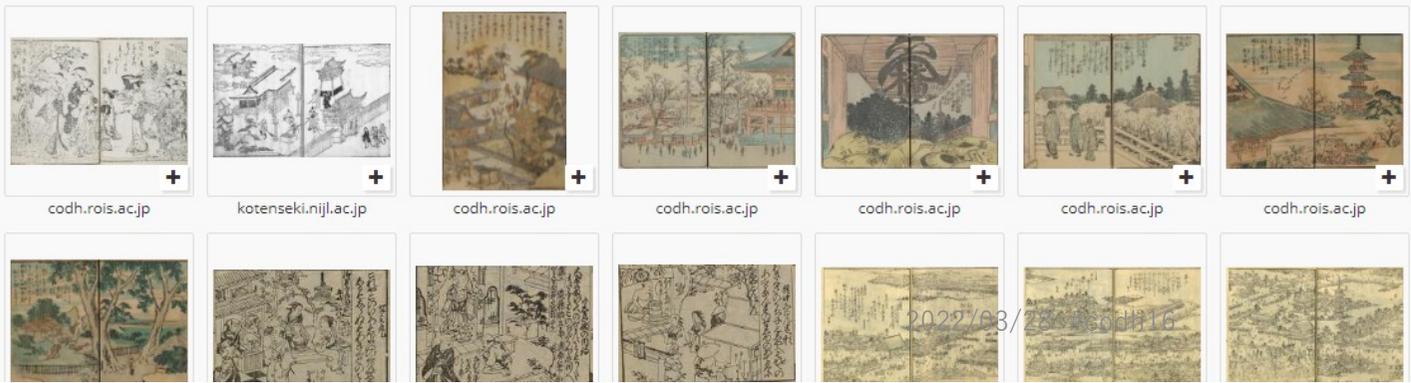
キーワードへ  
寺社 (775) / 水 (456) / 鳥居 (262) / 稲藪 (153) / 橋 (104) / 歌舞伎 (67) / 坂 (63) / 富士山 (59) / 松 (51) / 桜 (48) / 祭 (48) / 花見 (43) / 伝説 (42) / 市場 (32) / 漁業 (30) / 食事 (29) / 墳墓 (22) / 宿場 (19) / 雲龍 (18) / 武家屋敷 (17) / 農業 (17) / 門 (16) / 広場 (14) / 梅 (13) / 遊郭 (12) / 工芸 (12) / 武芸 (11) / 城 (10) / 紅葉 (8) / 五重塔 (6) / 富士塚 (6) / 藤棚 (5) / 衣料 (5) / 医療 (5) / 濡仏 (5) / 納涼 (4) / 榎木屋 (4) / 御旅所 (4) / 塩 (4) / 花火 (4) / 海苔 (3) / 塩業 (3) / 虫 (3) / 月見 (3) / 祭礼 (3) / 竹 (3) / 酒 (2) / 火除地 (2) / 潮干狩 (2) / 蕎麦 (2) / 炭 (2) / 御簀 (2) / 蓮 (2) / 筑波山 (2) / 榎 (2) / 材木 (2) / 椎 (2) / 松浦家 (2) / 菅蒲 (2) / 料理 (2) / 柳 (2) / 祇園会 (2) / 紅葉 (1) / 米 (1) / 大久保家 (1) / 出版 (1) / 鶴屋喜右衛門 (1) / 山王祭 (1) / 海鼠 (1) / 鮎 (1) / 鷹場 (1) / 菅 (1) / 林業 (1) / 鮎 (1) / 三重塔 (1) / 桃 (1) / 杉 (1) / 卯花 (1) / 焼き米 (1) / 蓮見 (1) / 両大師運座 (1) / 大黒詣 (1) / 聴虫 (1) / 鏡物 (1) / 櫻枝 (1) / 報恩講 (1) / 芝 (1) / 蕎麦 (1) / 糠 (1) / 萩 (1) / 瓦 (1) / 鱈 (1) / うどん (1) / 梨 (1) / 神楽 (1) / 鯉 (1) / 牡丹 (1) / 二十六夜 (1) / 楓 (1) / 湯屋 (1) / 黒田家 (1) / 鯉 (1) / 菊 (1) / 桜 (1)

名所 (統一地名) へ  
浅草寺 (26) / 隅田川 (26) / 寛永寺 (23) / 増上寺 (11) / 羅漢寺 (11) / 日本橋 (10) / 両国橋 (10) / 新吉原 (10) / 神田神社 (10) / 亀戸天満宮 (10) / 富岡八幡宮 (10) / 待乳山 (9) / 愛宕山 (9) / 飯倉神明宮 (9) / 高輪 (9) / 飛鳥山 (9) / 氷川神社 (9) / 護国寺 (9) / 不忍池 (8) / 品川 (8)

<http://codh.rois.ac.jp/edo-spots/>

江戸時代に出版されたガイドブック「名所案内」から「名所記」「名所図会」の挿絵を抽出したアイコン

名称やキーワードを付与すると、江戸の観光地を詳しく紹介する動画が作成できる



2022/03/28 ©codh16

# 資料情報と実体情報の間にあるギャップ

- マイクロコンテンツ化  
= **文書空間の情報**（資料に内在する情報）を取り出し、メタデータとして整理してきた
- より広範な研究や利用には**文書空間の情報**のみでは限界
  - 資料についての情報は、深掘して詳細化できるが……
- **実世界・実体空間の情報**（資料に内在しない情報）との紐づけがなければ、実世界の活動での利活用は難しい
- そこで……

## 花小金井駅の「花」って何？ 実は西武の名称戦略、いまも幅広ホームに名残あり

3/27(日) 9:42 配信 259



### 小金井5駅 頭に新、武蔵、東…「花」？



花小金井駅の高田馬場寄り。かつては中線がホームへと入っていた(2022年3月、内田家治撮影)。

西武新宿線の花小金井駅ホーム(東京都小平市)に降り立つと、妙な違和感に襲われます。ホームの幅がやたら広いのです。ひとつのホームの両側に上下線の線路がある1面2線の島式ホームなのですが、そのホームは通常の2倍くらいの幅があります。「花小金井」という駅名も何やらいわくありげです。

【幅広ホームの秘密】「中線」があった昭和の花小金井駅写真で見る

同駅の南約1kmの地に都立小金井公園が広がっています。小金井公園は、都内でも屈指の花見の名所。園内にはヤマザクラ、サトザクラ、ソメイヨシノなど約1700本の桜が植えられていて、開花時期には大賑わいとなります。

花小金井の「花」は同園のことを指していると思うかもしれませんが、同園の開園は1954(昭和29)年。その前身をたどっても、東京緑地計画を基に1940(昭和15)年の「紀元二千六百年記念事業」として計画された小金井緑地です。一方、花小金井駅の開業は西武新宿線の高田馬場～東村山間が開通した1927(昭和2)年。小金井公園ができるよりずっと前のことです。

ここで、ほかに「小金井」の名が付く駅の開業年を列挙しておきましょう。

- ・小金井駅(栃木県下野市)：1893(明治26)年、JR東北本線(当時は私鉄の日本鉄道)
- ・新小金井駅(東京都小金井市)：1917(大正6)年、西武多摩川線(当時は多摩鉄道)
- ・武蔵小金井駅(同上)：1924(大正13)年仮開業、JR中央線。本開業は1926年
- ・東小金井駅(同上)：1964(昭和39)年開業、JR中央線

小金井の前に付く文字に関して、新小金井は、すでにあつた東北本線の小金井より後にできた新しい駅が由来です。すぐ近くの武蔵小金井と区別するためと誤解されがちですが、武蔵小金井より早く開業しています。武蔵小金井は旧武蔵国にあるため、東小金井はその東隣に位置するため。では花小金井の「花」とは何でしょうか。

<https://news.yahoo.co.jp/articles/29a8c0834eae949397b8457c8cc3e038388ba24e>

## 江戸観光案内

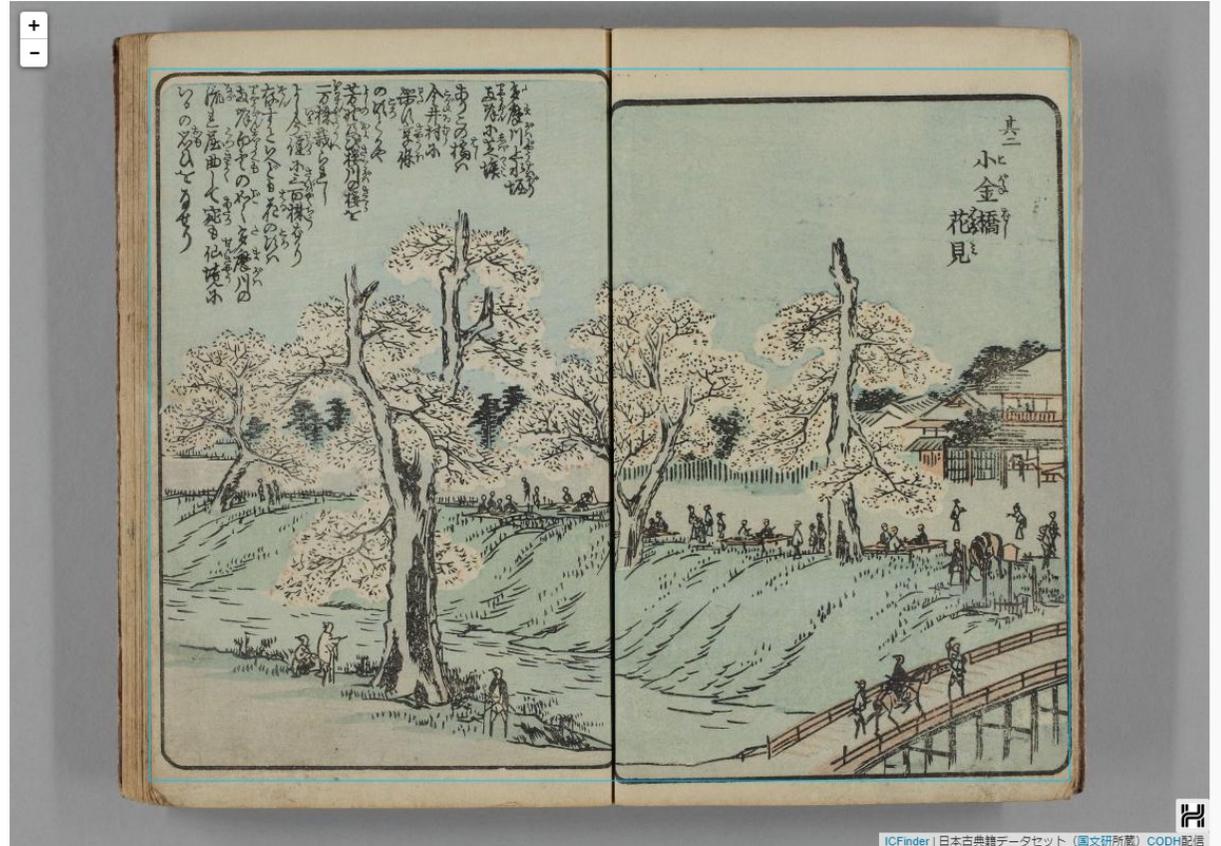
日本語 / English ログイン ヘルプ リスト

絵本／江戸土産

戻る + リストに入れる

歴史地名データID：40243098

名所(統一地名)：小金井橋 歴史地名データID：40243098 出典：絵本江戸土産 名所(原本表記)：小金井橋 キーワード：花見 キーワード：植栽  
キーワード：桜 江戸観光ID：338



ICFinder | 日本古典籍データセット (国文研所蔵) CODH配信

2022/03/28 #codh16



2022/03/28



0116

# 解決方法：識別子の双方向結合

情報に識別子を与えて双方向結合

= 「ある文書が実世界のある実体に言及している」という関係を表現できる

二つの結合の方向性

- エンティティリンクング
  - 文書空間側から実体空間側に対するリンク
  - 例：「ある挿絵」→「実体空間のどこに当たるか」へリンク
- エビデンスハンティング
  - 実体空間側から文書空間側に対するリンク
  - 例：「ある現実の名所」→「様々な資料に登場する記述」すべてへリンク

# 「文書空間」と「実体空間」の 双方向結合（エンティティリンキング）

- マイクロコンテンツとして整理した**文書空間の情報**と、実世界の活動で利用したい**実体空間の情報**の、双方に識別子を与える
- 識別子同士の紐づけ（リンク・リレーション）を行えば、それを通じて文書空間と実体空間を相互に行き来できる
- 「江戸観光案内」に登場する名所を例とすると
  - 文書空間の情報 = 原典、資料に記載されている名称、資料内の位置…
  - 実体空間の情報 = 現在の一般的呼称、観光ジャンル、緯度経度…

## 文書空間識別子

Canvas ID :

<https://kotenseki.nijl.ac.jp/biblio/100000011/canvas/23#xywh=848,917,2408,3160>

原典表記：神田明神

原典：江戸名所記

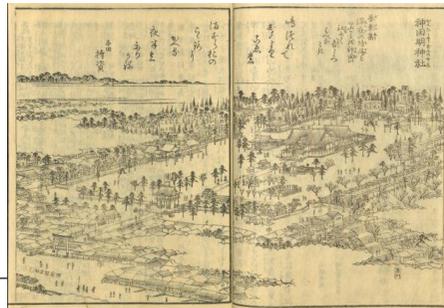


Canvas ID :

<http://codh.rois.ac.jp/pmjt/iiif/100249896/canvas/00839#xywh=944,901,4928,3384>

原典表記：神田明神社

原典：江戸名所図会



## 実体空間識別子と分類体系

江戸観光ID : 395

名称：神田神社

解説文：神田神社（神田明神）は、今日でも江戸総鎮守として……

江戸観光分類 : 3

分類名：寺社

GeoLOD : lKa0o2

緯度 : 35.702222

経度 : 139.767777

# edomiによる双方向結合の試み

- edomi 江戸をみる／みせるデータポータル
  - <http://codh.rois.ac.jp/edomi/>
  - 江戸の街を一望できることで有名だった「江戸見坂」に由来
  - 都市としての江戸／時代としての江戸に関する各種のデータを一望できるデータポータルを目指す
- edomiのコンテンツとして「江戸買物案内」「江戸観光案内」を実体空間と紐づけて展開した
  - 江戸観光案内→トラベル
  - 江戸買物案内→ショッピング 総務省標準産業分類も活用



テーマから探す



カテゴリ一覧



トラベル



ショッピング



マップ



グルメ



政治



経済

# 江戸見坂



GeoLOD	<a href="#">nusq0U</a>
名所分類	地名
名所 (原本表記)	江戸見坂
キーワード	坂
出典	絵本江戸土産
江戸マップID	4-140
名所 (統一地名)	江戸見坂
IIIF Manifest	<a href="http://codh.rois.ac.jp/iiif/Book/200093687/manifest.json">http://codh.rois.ac.jp/iiif/Book/200093687/manifest.json</a>



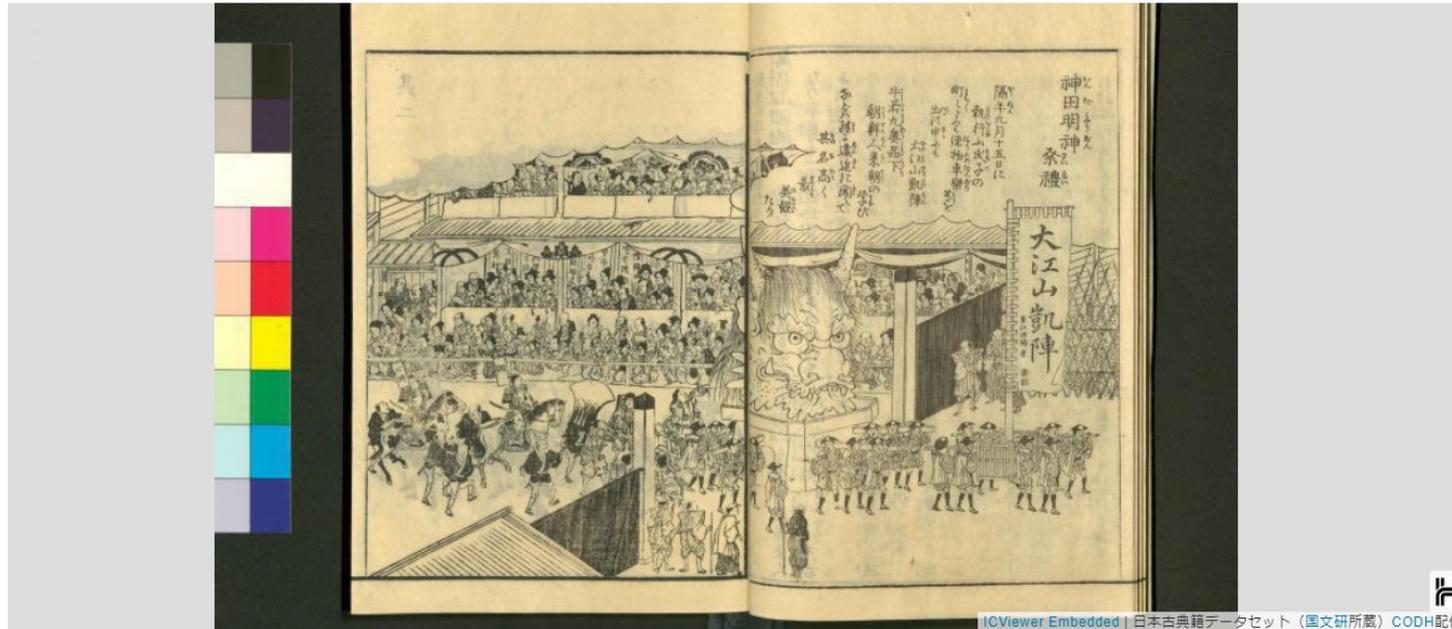
# 神田神社



ICViewer Embedded | 日本古典籍データセット (国文研所蔵) CODH配信

位置情報	GeoLOD	<a href="#">lKa0o2</a>	
	名所分類	寺社	観光ジャンル
外部典拠	歴史地名データID	10015082	
	名所 (原本表記)	神田明神の社	文書空間内の名称
文書空間に描かれている要素	キーワード	鳥居	
	出典	絵本江戸土産	文書空間の原典
	江戸マップID	<a href="#">20-051</a>	

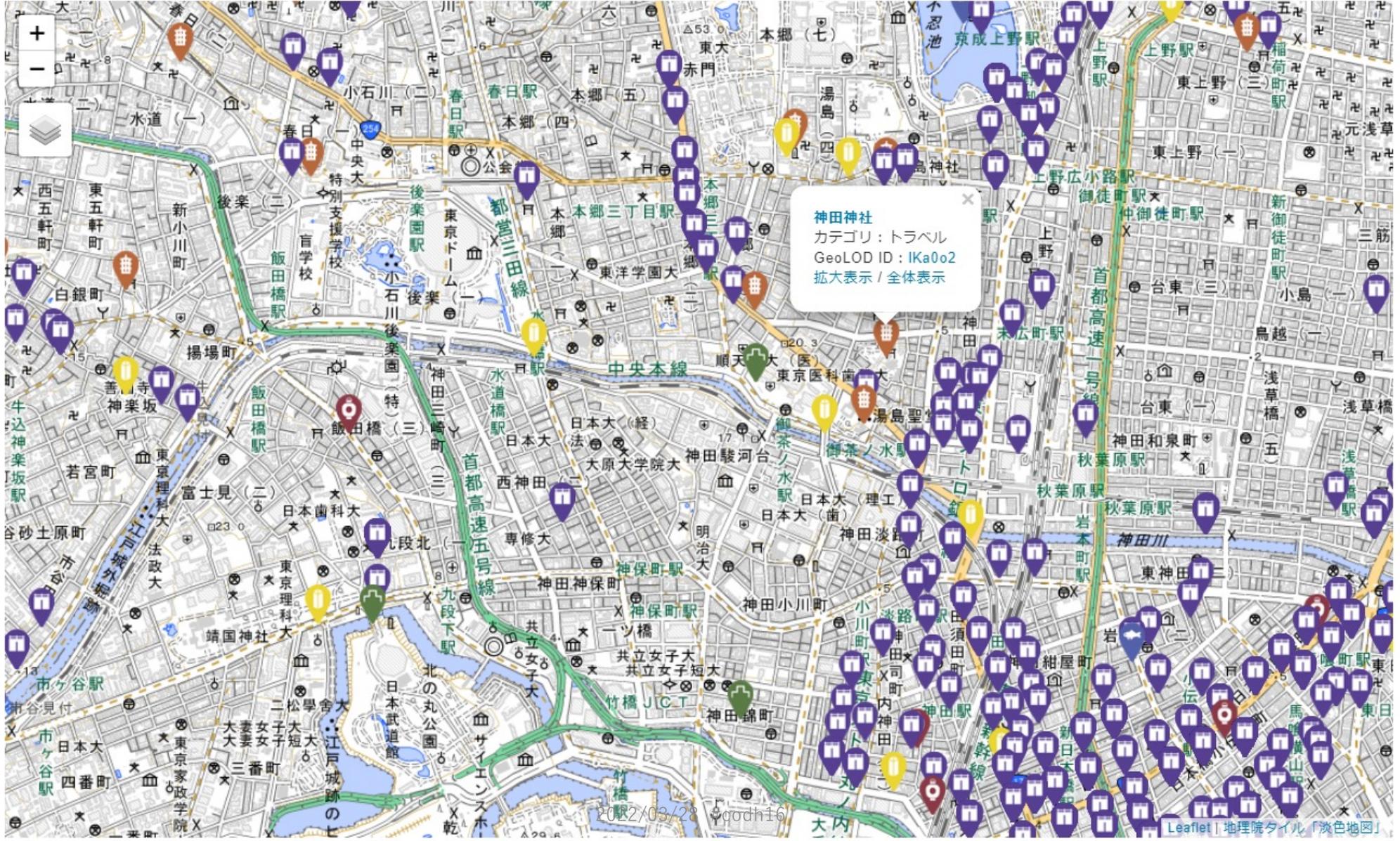
# 神田神社



ICViewer Embedded | 日本古典籍データセット (国文研所蔵) CODH配信

GeoLOD	lKa0o2
名所分類	寺社
歴史地名データID	<a href="#">10015082</a>
名所 (原本表記)	神田明神祭礼
キーワード	祭
出典	江戸名所図会
名所 (統一地名)	神田神社

2022/03/28 #codh16

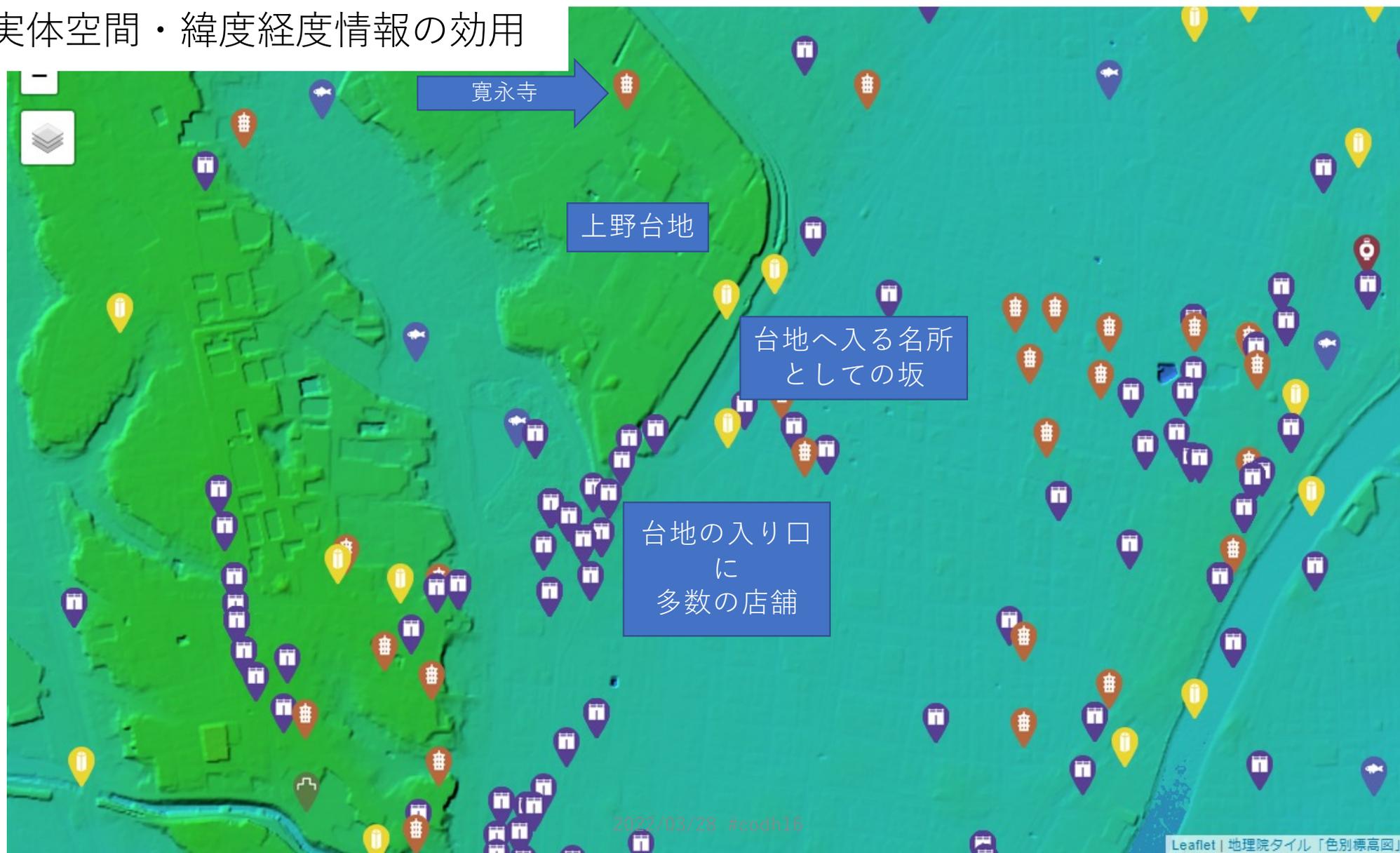


神田神社  
カテゴリ: トラベル  
GeoLOD ID: 1Ka0o2  
拡大表示 / 全体表示

# 実体空間・緯度経度情報の効用



# 実体空間・緯度経度情報の効用





# 歴史ビッグデータ構想に向けて

- 歴史ビッグデータ構想
  - 過去の間人が書き記した実世界に関する記録を構造化したデータを「歴史的記録」と定義（歴史的状況記録・歴史的行動記録）
  - 5W1Hなどの形式で過去の事実に関する構造化データを収集
  - 各種研究のエビデンスとして活用し、過去の世界を復元・分析
  - 類例：<https://www.timemachine.eu/>
- 文書を出発点とする半構造化データと、実体を出発点とする構造化データを双方向から接続し検索可能にする必要性
- 今回の双方向結合はそのための方法の一つ

# 今後の展開

- 「歴史的行動記録」と名所・商人情報の紐づけ
- 人々の行動を記した歴史資料には5W1H情報が含まれている
  - 例：清河八郎による旅日記『西遊草』
  - 清河が（Who）江戸のどの名所に（Where）いつ滞在し（When）何をどうしたか（What How Why）
- 様々な日記の情報を抽出し、edomiにすでに載せている文書空間識別子・実体空間識別子と紐づけ
  - 清河八郎、シュリーマン、泉光院（薩摩藩支藩の修験者）……
  - DH2022で発表予定