Digital Gazetteer as a Knowledgebase for Open Data Science

Hara, S. and Sekino, T.
[1]Center for Southeast Asian Studies, Kyoto University, Kyoto, Japan, E-mail: shara@cseas.kyoto-u.ac.jp
[2]International Research Center for Japanese Studies, Kyoto, Japan, E-mail: sekino@nichibun.ac.jp

Abstract

Digital gazetteers are essential knowledge sources for the humanities in order to allow the association of heterogeneous data sets in the context of geographical proximity. There are some free gazetteer databases about Japanese place names but they contain only contemporary information. The Humanities' GIS Research Group (H-GIS) and the National Institutes for the Humanities (NIHU) have collaborated in creating "The Digital Gazetteer for Historical Japanese Place Names (DGHJ)" that is a free gazetteer database of Japanese historical place names. After more than ten years of work, H-GIS and NIHU finally succeeded in downloading the DGHJ data sets. This paper will describe sources, data compiling methods, organization of data structure, and current achievements of the DGHJ.

1. Introduction

Digital gazetteers are essential knowledge sources for the humanities in order to allow the association of heterogeneous data in the context of geographical proximity. This means that data from different sources and media are correlated with location information obtained from the conversion of place names into longitude and latitude pairs by referring to description and information from digital gazetteers. There are some free gazetteer databases for contemporary Japanese place names (e.g., the Getty Thesaurus of Geographic Names (Getty), and the Address Matching Service by the Center for Spatial Information Science at the University of Tokyo (CSIS)), but there is a lack of gazetteer databases which cover Japanese historical place names. On the other hand, there are commercially available dictionaries of Japanese historical place names (e.g., *Shinpan Kadokawa Nihon Chimei Daijiten* 新版角川日本地名大辞典 (The New Edition of the Kadokawa Geographical Dictionary of Japan) (Kadokawa, 2011)), but these are expensive, and the free use is limited by copyrights. Therefore, from 2005, the Humanities' GIS Research Group (H-GIS) and the National Institutes for the Humanities (NIHU) began the design and development of "The Digital Gazetteer for Historical Japanese Place Names (DGHJ)."

The DGHJ collects historical place names from four different printed materials, these are: *Dai Nihon Chimei Jisho* 大日本地名辞書 (The Dictionary of Geographical Names of Japan: DGJ), *Engishiki Jinmyōchō* 延喜式神名帳 (Register of Deities in Procedures of the Engi Era: EJM), *Nihon Ji'in Sōran* 日本寺院総鑑 (Directory of Japanese Temples: DJT), and *Kyū Go Manbun no Ichi Chikeizu* 旧5万分1地形図 (1:50,000 Old Topographic Maps: MAP). At the time of writing this paper, the DGHJ included a total of 377,471 historical place names, of which, 53,528 historical place names collected from *Dai Nihon Chimei Jisho* (DGJ); 2,842 historical place names from *Engishiki Jinmyōchō* (EJM); 78,557 historical place names from *Nihon Ji'in Sōran* (DJT), and 242,544 historical place names from *Kyū Go Manbun no Ichi Chikeizu* (MAP). This is the largest free digital gazetteer datasets for Japanese historical place names. After more than ten years of work, from March 2018, H-GIS and NIHU were finally able to start downloading the DGHJ data sets except *Nihon Ji'in Sōran* (DJT).

The DGHJ implements two interfaces, one which is an ordinary Web based graphical user interface (GUI), and the other which is an application programming interface (API) for Web services. A Web API is a convenient mashup framework to create new applications by combining extant applications and data sets. Recently, as Semantic Web becomes popular and offers easy and sophisticated ways to link data on the Web, the DGHJ also implements SPARQL Endpoint (W3C, 2013) as its third API. These interfaces will be released after a while. In this paper, Section 2 will describe the data sources, data structure and database system, Section 3 will show a new database system using RDF (Resource Description Framework) (W3C, 2004) repositories and SPARQL Endpoint, and lastly, the problems and final considerations will be discussed in Section 4.

2. Construction of Digital Gazetteer Data Sets and Database

This section will describe data sources, data structure and the database systems of the DGHJ. Data sources from printed dictionaries will be

Online ISSN 2673-0014/ © Geoinformatics International

described in section 2.1, printed maps in section 2.2, and the section 2.3 will discuss the database system.

2.1 Constructing Digital Gazetteers using Printed Dictionaries

The DGHJ uses three paper dictionaries: *Dai Nihon Chimei Jisho* (DGJ), *Engishiki Jinmyōchō* (EJM), and *Nihon Ji'in Sōran* (DJT) as the main sources for the collection of historical place names (Oketani, 2009).

The first source is *Dai Nihon Chimei Jisho* (DGJ) which was compiled and published by "Tohgo YOSHIDA (1864 - 1918)″in 1900 (Yoshida, 1900). This dictionary's edition consists of eight volumes and includes 53,676 historical place names used around the 19th Century not only inside the Japanese Archipelago but also in Taiwan and Karafuto (Sakhalin) as shown in Figure 1. *Dai Nihon Chimei Jisho* (DGJ) registers historical place names of countries, counties, cities, towns, villages, manors, temples, shrines, harbors, mountains, rivers, lakes, marshes, scenic spots, historic spots, and so on as indexes (Figure 2). Following a headword of each place name comes a description of details about historical transformations including change of names, locations and so on based on the philological study of each place name (Figure 3).

The second source is *Engishiki Jinmyōchō* (EJM) which was included as the 9th and 10th volume of *Engishiki* 延喜式 (Procedures of the Engi Era). *Engishiki* was originally compiled as a 50-volume edition in the year 927 containing information about laws and customs under use around the 10th century. As a part of *Engishiki*, *Engishiki Jinmyōchō* (EJM) registers 2,861 existing *Shinto* shrines (shrines from the traditional religion of Japan, *Shinto* 神道) and 3,132 officially recognized and enshrined *Kami* 神 (Japanese gods and spirits) at that time (Kuroita ed., 1979 and Shikinaisha Kenkyukai, 1979). *Engishiki Jinmyōchō* (EJM) lists every *Shinto* shrine according to the country name and county name, and includes detailed information about its enshrined deity and the shrine ranking (Figure 4).

The third source is *Nihon Ji'in Sōran* (DJT) which lists more than 78,000 contemporary temple names and includes information about the name of religious sect, addresse, phone number, chief priest name, vice priest name, and name of the principal object of worship at each temple (Ji'in Sōran Kankōkai, 2000).

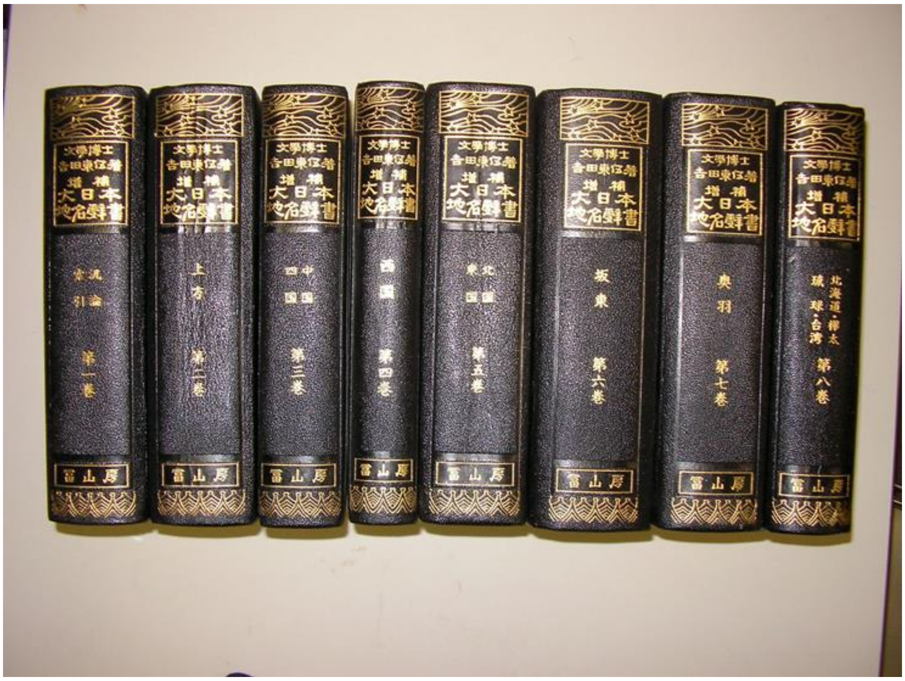
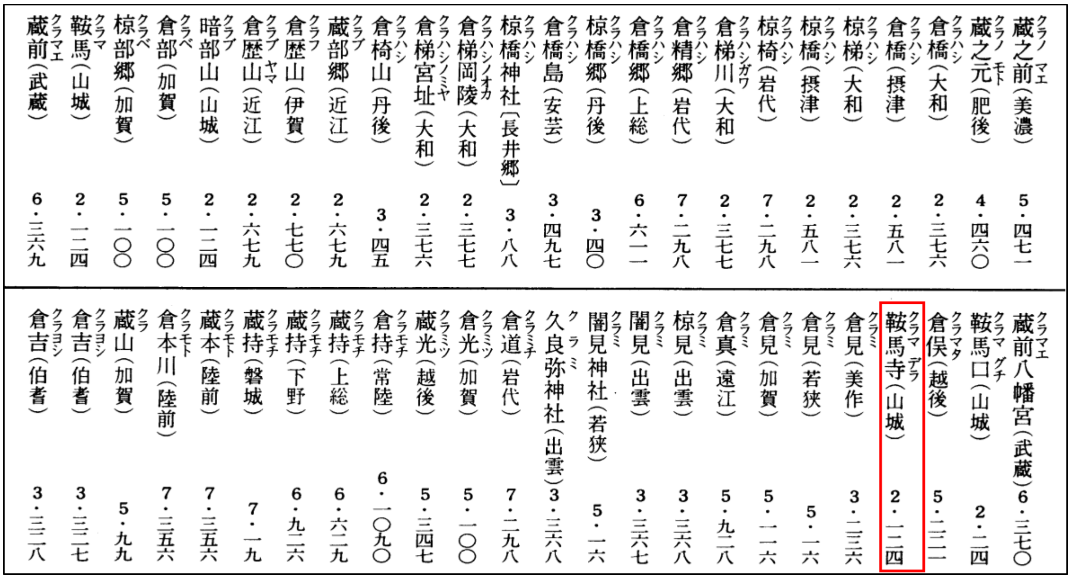
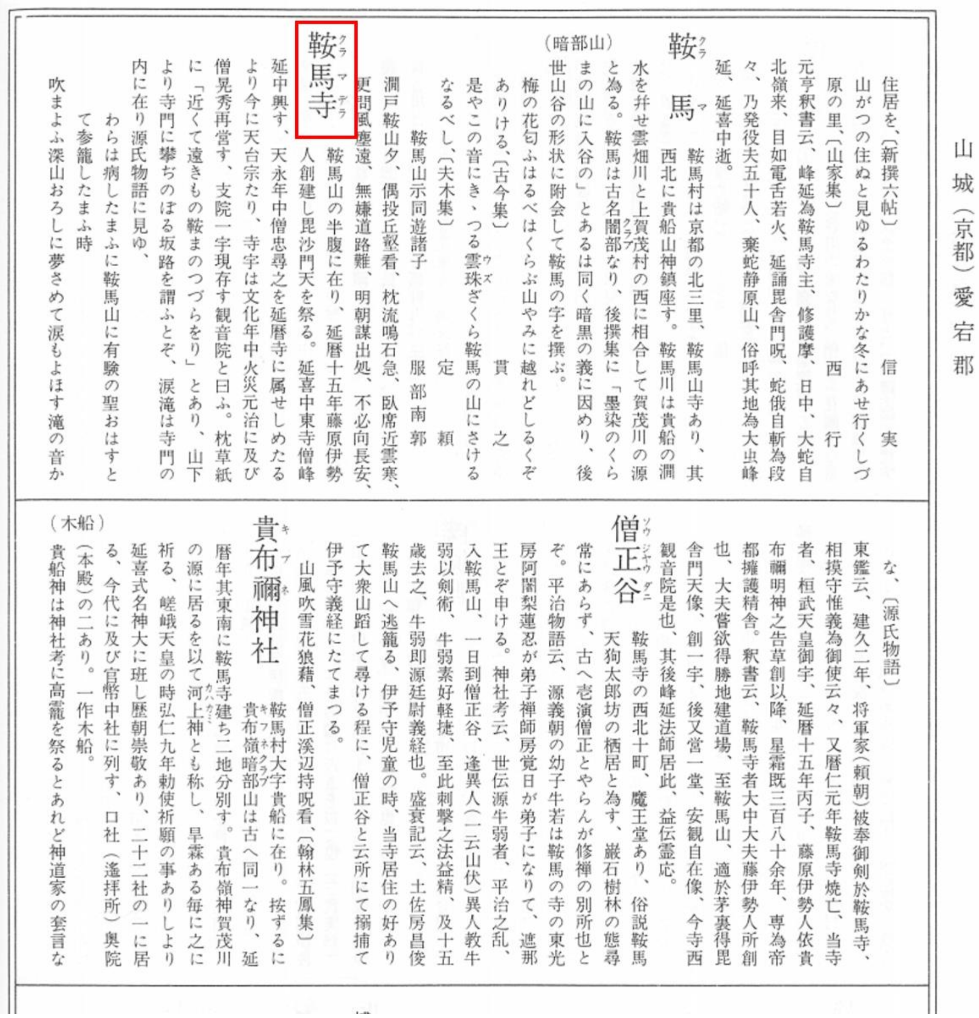


Figure 1: *Dai Nihon Chimei Jisho* 大日本地名辞書 (The Dictionary of Geographical Names of Japan: DGJ)



クラノマエ	蔵之前（美濃）	5・四七一
クラノモト	蔵之元（肥後）	4・四六〇
クラハシ	倉橋（大和）	2・三七六
クラハシ	倉橋（摂津）	2・五八一
クラハシ	椋梯（大和）	2・三七六
クラハシ	椋橋（摂津）	2・五八一
クラハシ	椋椅（岩代）	7・二九八
クラハシガワ	倉梯川（大和）	2・三七七
クラハシ	倉精郷（岩代）	7・二九八
クラハシ	倉橋郷（上総）	6・六一一
クラハシ	椋橋郷（丹後）	3・四〇
クラハシ	倉橋島（安芸）	3・四九七
クラハシ	椋橋神社〔長井郷〕	3・八八
クラハシノオカ	倉梯岡陵（大和）	2・三七七
クラハシノミヤ	倉梯宮址（大和）	2・三七六
クラハシ	倉椅山（丹後）	3・四五
クラブ	蔵部郷（近江）	2・六七九
クラフ	倉歴山（伊賀）	2・七七〇
クラブヤマ	倉歴山（近江）	2・六七九
クラブ	暗部山（山城）	2・一二四
クラベ	倉部（加賀）	5・一〇〇
クラベ	椋部郷（加賀）	5・一〇〇
クラマ	鞍馬（山城）	2・一二四
クラマエ	蔵前（武蔵）	6・三六九
クラマエ	蔵前八幡宮（武蔵）	6・三七〇
クラマグチ	鞍馬口（山城）	2・一二四
クラマタ	倉俣（越後）	5・二二一
クラマデラ	鞍馬寺（山城）	2・一二四
クラミ	倉見（美作）	3・二三六
クラミ	倉見（若狭）	5・一六
クラミ	倉見（加賀）	5・一一六
クラミ	倉真（遠江）	5・九二八
クラミ	椋見（出雲）	3・三六八
クラミ	闇見（出雲）	3・三六七
クラミ	闇見神社（若狭）	5・一六
クラミ	久良弥神社（出雲）	3・三六八
クラミチ	倉道（岩代）	7・二九八
クラミツ	倉光（加賀）	5・一〇〇
クラミツ	蔵光（越後）	5・三四七
クラモチ	倉持（常陸）	6・一〇九〇
クラモチ	蔵持（上総）	6・六二九
クラモチ	蔵持（下野）	6・九二六
クラモチ	蔵持（磐城）	7・一九
クラモト	蔵本（陸前）	7・三五六
クラモト	倉本川（陸前）	7・三五六
クラ	蔵山（加賀）	5・九九
クラヨシ	倉吉（伯耆）	3・三二七
クラヨシ	倉吉（伯耆）	3・三二八

Figure 2: Index of *Dai Nihon Chimei Jisho* (DGJ)



山城（京都）愛宕郡

住居を、〔新撰六帖〕　信実
山がつの住ぬと見ゆるわたりかな冬にあせ行くしづ
原の里、〔山家集〕　西行

元亨釈書云、峰延為鞍馬寺主、修護摩、日中、大蛇自北嶺来、目如電舌若火、延誦毘舎門呪、蛇俄自斬為段々、乃発役夫五十人、棄蛇静原山、俗呼其地為大虫峰延、延喜中逝。

鞍馬（クラマ）

鞍馬村は京都の北三里、鞍馬山寺あり、其西北に貴船山神鎮座す。鞍馬川は貴船の澗水を并せ雲畑川と上賀茂村の西に相合して賀茂川の源と為る。鞍馬は古名闇部（クラブ）なり、後撰集に「墨染のくらまの山に入谷の」とあるは同く暗黒の義に因めり、後世山谷の形状に附会して鞍馬の字を撰ぶ。

（暗部山）
梅の花匂ふはるべはくらぶ山やみに越れどしるくぞ
ありける、〔古今集〕　貫之
是やこの音にき、つる雲珠（ウズ）ざくら鞍馬の山にさける
なるべし、〔夫木集〕　定頼

鞍馬山示同遊諸子　服部南郭
澗戸鞍山夕、偶投丘壑看、枕流鳴石急、臥席近雲寒、
更問風塵遠、無嫌道路難、明朝謀出処、不必向長安、

鞍馬寺（クラマデラ）

鞍馬山の半腹に在り、延暦十五年藤原伊勢人創建し毘沙門天を祭る。延喜中東寺僧峰延中興す、天永年中僧忠尋之を延暦寺に属せしめたるより今に天台宗たり、寺宇は文化年中火災元治に及び僧晃秀再営す、支院一宇現存す観音院と曰ふ。枕草紙に「近くて遠きもの鞍まのつづらをり」とあり、山下より寺門に攀ぢのぼる坂路を謂ふとぞ、涙滝は寺門の内に在り源氏物語に見ゆ、

わらは病したまふに鞍馬山に有験の聖おはすと
て参籠したまふ時
吹まよふ深山おろしに夢さめて涙もよほす滝の音か
な、〔源氏物語〕

東鑑云、建久二年、将軍家（頼朝）被奉御剣於鞍馬寺、相摸守惟義為御使云々、又暦仁元年鞍馬寺焼亡、当寺者、桓武天皇御宇、延暦十五年丙子、藤原伊勢人依貴布禰明神之告草創以降、星霜既三百八十余年、専為帝都擁護精舎。釈書云、鞍馬寺者大中大夫藤伊勢人所創也、大夫営欲得勝地建道場、至鞍馬山、適於茅裏得毘舎門天像、創一宇、後又営一堂、安観自在像、今寺西観音院是也、其後峰延法師居此、益伝霊応。

僧正谷（ソウジヤウダニ）

鞍馬寺の西北十町、魔王堂あり、俗説鞍馬天狗太郎坊の栖居と為す、巌石樹林の態尋常にあらず、古へ壱演僧正とやらんが修禅の別所也とぞ。平治物語云、源義朝の幼子牛若は鞍馬の寺の東光房阿闍梨蓮忍が弟子禅師房覚日が弟子になりて、遮那王とぞ申ける。神社考云、世伝源牛弱者、平治之乱、入鞍馬山、一日到僧正谷、逢異人（一云山伏）異人教牛弱以剣術、牛弱素好軽捷、至此刺撃之法益精、及十五歳去之、牛弱即源廷尉義経也。盛衰記云、土佐房昌俊鞍馬山へ逃籠る、伊予守児童の時、当寺居住の好ありて大衆山蹈して尋ける程に、僧正谷と云所にて搦捕て伊予守義経にたてまつる。

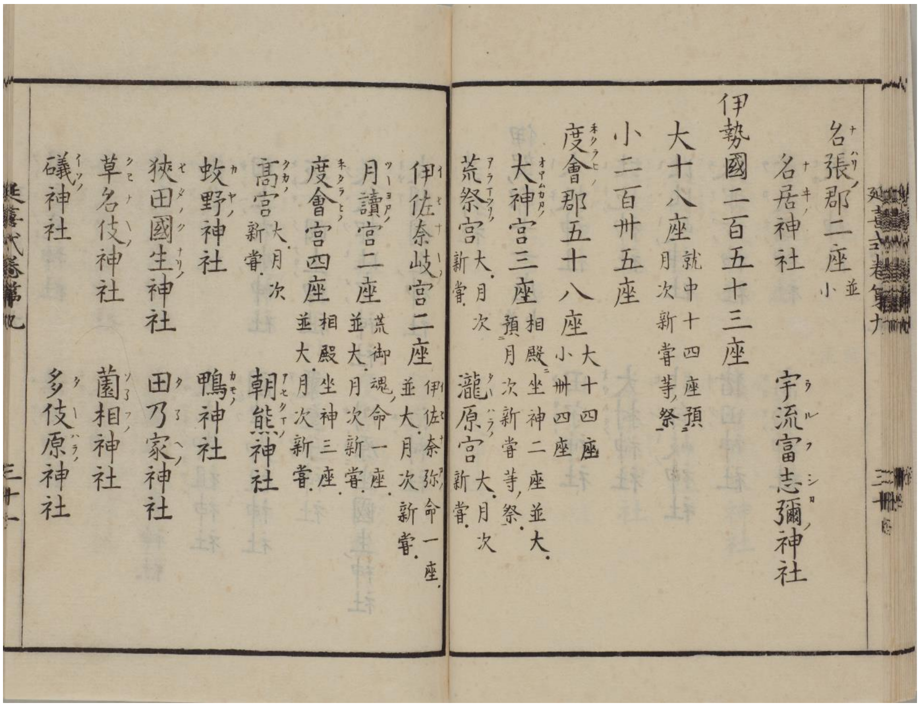
山風吹雪花狼藉、僧正渓辺持呪看、〔翰林五鳳集〕

貴布禰（キフネ）神社

鞍馬村大字貴船に在り。按ずるに貴布嶺（キフネ）暗部（クラブ）山は古へ同一なり、延暦年其東南に鞍馬寺建ち二地分別す。貴布嶺神賀茂川の源に居るを以て河上神（カハカミ）とも称し、旱霖ある毎に之に祈る、嵯峨天皇の時弘仁九年勅使祈願の事ありしより延喜式名神大に班し歴朝崇敬あり、二十二社の一に居る、今代に及び官幣中社に列す、口社（遙拝所）奥院（本殿）の二あり。一作木船。

（木船）
貴船神は神社考に高龗を祭るとあれど神道家の套言な

Figure 3: A detail description example of *Dai Nihon Chimei Jisho* (DGJ) concerning *Kurama Dera* 鞍馬寺 *(Kurama* Temple) enclosed by a square in the index of Figure 2



名張郡二座 並小
名居神社　宇流富志彌神社
伊勢國二百五十三座
大十八座 就中十四座預月次新嘗等祭
小二百卅五座
度會郡五十八座 大十四座 小卌四座
大神宮三座 相殿坐神二座並大 預月次新嘗等祭
荒祭宮 大 月次新嘗　瀧原宮 大 月次新嘗
伊佐奈岐宮二座 伊佐奈弥命一座 並大 月次新嘗
月讀宮二座 荒御魂命一座 並大 月次新嘗
度會宮四座 相殿坐神三座 並大 月次新嘗
高宮 大 月次新嘗　朝熊神社
蚊野神社　鴨神社
狹田國生神社　田乃家神社
草名伎神社　薗相神社
礒神社　多伎原神社
延喜式卷第九　三十一

Figure 4: Example of *Engishiki Jinmyōchō* 延喜式神名帳 (Register of Deities in Procedures of the Engi Era: EJM) (National Institute of Japanese Literature)

Dai Nihon Chimei Jisho (DGJ) and *Engishiki Jinmyōchō* (EJM) have indexes to link place names to their detailed descriptions. *Nihon Ji'in Sōran* (DJT) has a simple data table that describes temple names and their related data items as explained above. Though these dictionaries give us enough descriptive information about each place, they lack location information of longitude and latitude pairs. Thus, most of the time spent for creating the digital gazetteers was used to estimate the location of each place name.

Following is our procedure to identify the longitude and latitude pair of each historical place name collected from the above cited sources. First, we carefully read descriptions about each historical place name in the original sources and each place's longitude and latitude pair was identified on current maps by following the next guidelines:

- If the same place name can be found on current maps and its shape is represented by a point (e.g., a house, a mountain peak, a historic spot, and a monument), the location of the place name will be described with the longitude and latitude pair of that point.
- If the same place name (its shape is represented by a point) cannot be found on current maps but its place is identifiable using other evidences (e.g., related documents, ruins, and monuments), the evidences will be used to estimate the longitude and latitude pair of its location.
- If the place name refers to an administrative district (e.g., villages, counties, and countries) and the location of the government office at that time is identifiable on current maps, the location of the place name will be described as the longitude and latitude pair of the government office.
- If the place name refers to an administrative district and the location of the government office at that time is not identifiable on current maps, but the current administrative district has features that allow to identify the ancient administrative district, the location of the place name will be described as the longitude and latitude pair of the location of the current administrative office of the current district.
- If the place name can be found on current maps and its shape is represented by a plane (e.g., lake, marsh, and manor), the location of the place name will be described as the longitude and latitude pair at the center of the rectangle circumscribing the place.
- If the place name can be found on current maps and its shape is represented by a line (e.g., river and road), the location of the place name will be

described as two longitude and latitude pairs at the starting point and end point of the line.

- Otherwise, the identification of the place name on current maps will be consigned to researchers' resourcefulness based on their knowledge about that historical period, their experience with research on that geographical location, and so on. Especially for EJM, the identification is done in accordance with the opinions made by Shikinaisha Kenkyukai (Shikinaisha Kenkyukai, 1979). In the case of resourcefulness-based place name identification, its outline will be described in the memos. Consequently, the precision of locations is inconsistent. These inconsistencies about identified locations will be corrected by accepting experts' advices or opinions from the public.

The main data elements and description rules of the DGHJ are summarized as follows (element names printed in a bold type serve for the following explanation and are different from actual element names used in the database):

- **ID**: A unique number in the database.
- **Place Name**: The name of a historical place name found in original sources. It is described by KANJI (Chinese characters), KANA (Japanese syllabary characters) and Roman Characters within the ISO/IEC 10646 Character Sets.
- **County Name**: The name of the county in which a historical place name is included. It is described by KANJI (Chinese characters), KANA (Japanese syllabary characters) and Roman Characters within ISO/IEC 10646 Character Sets.
- **Country Name**: The name of the country in which a historical place name is included. It is described by KANJI (Chinese characters), KANA (Japanese syllabary characters) and Roman Characters within ISO/IEC 10646 Character Sets.
- **Identified Place Name**: The current place name corresponding to the historical place name. It is described by KANJI (Chinese characters), KANA (Japanese syllabary characters) and Roman Characters within ISO/IEC 10646 Character Sets.
- **Shape**: The shape of the historical place (point, line or polygon).
- **Locations**: Pair(s) of longitude and latitude. Format is "DDD.dddd" and datum is WGS84.
- **Attribute**: Type of place name (e.g., administrative district, structure, water area, land form). Each attribute is described by number (e.g., 2: country, 12: Shinto shrine, 22: river, 32: mountain, 55: gravy yard: *see* Table 1).
- **Reference**: Source name of the historical place name.
- **Memo**: Indication of place identification made according to resourcefulness.

"ID" and "Place Name" are required elements, and other elements are optional. For several reasons, including the fact that some data elements appear repeatedly, the DGHJ uses XML for its data description. Follows below an example of the DGHJ data using XML. Here, <attribute> shows the type of a historical place name as shown in Table 1 (Oketani, 2007).

```
<item pid="10026682">
  <country>山城</country>
  <county reading="カミキョウ">上京</county>
  <placename reading="ソウコクジ">相国寺</placename>
  <proma1>so^kokuji</proma1>
  <proma2>sokokuji</proma2>
  <proma3>sokokuji</proma3>
  <proma4>so^kokuji</proma4>
  ................
  <pname1>京都市上京区</pname1>
  <pname2/>
  <pname3/>
  <shp>1</shp>
  <loc>1</loc>
  <lat> 35.03333333 </lat>
  <long> 135.7627778 </long>
  ................
  <attribute>13</attribute>
  ...................
</item>
```

2.2 Constructing Digital Gazetteers using Printed Maps

The DGHJ uses *Kyū Go Manbun no Ichi Chikeizu* (MAP) which were created from 1890 to 1916 by the Japanese Imperial Army's Land Survey Bureau (Dai Nippon Teikoku Rikuchi Sokuryōbu, 1890-1916) (Figure 5). These are the oldest and most precise maps that were compiled by surveying the whole country based on a general standard. Digital gazetteer data using *Kyū Go Manbun no Ichi Chikeizu* (MAP) was completed by following the procedures below (Yotsui et al., 2010):

1. **Scanning**: *Kyū Go Manbun no Ichi Chikeizu* (MAP) were scanned with a precision of 600 dpi by a flatbed scanner.

2. **Geometric Correction**: Converting Old Tokyo Datum to WGS84.

Table 1: Attributes used in *Dai Nihon Chimei Jisho* (DGJ), *Engishiki Jinmyōchō* (EJM) and *Nihon Ji'in Sōran* (DJT)

ID	Group	Attribute
1	administrative name	local
2	administrative name	Country
3	administrative name	County
4	administrative name	Town
5	administrative name	Village
6	administrative name	Village
7	administrative name	City
8	administrative name	(Village) Section
81	administrative name	manore/
82	administrative name	ward/newly developed rice field
83	administrative name	prefecture
9	administrative name	Other
11	building	building
12	building	shrine
13	building	temple
14	building	bridge
15	building	checkpoint/barrier
16	building	castle
17	building	military
18	building	school
84	building	firm/factory
85	building	railway station
86	building	railway
19	building	other
21	hydrography	sea
22	hydrography	river/irrigation
23	hydrography	lake/marshy
24	hydrography	Bay/gulf/cove
25	hydrography	river mouth
26	hydrography	port
27	hydrography	lighthouse
29	hydrography	other
31	landform	mountain
32	landform	mountains
33	landform	mountain range
34	landform	highland
35	landform	hill
36	landform	basin
37	landform	plain
38	landform	island
41	landform	ridge
43	landform	slope
44	landform	valley
45	landform	road
46	landform	Beach/shore
47	landform	delta/fun
49	landform	other
51	sight/historic spots	sights/scenic spots
52	sight/historic spots	historic spots
53	sight/historic spots	waterfall
54	sight/historic spots	tot-spring/spa
55	sight/historic spots	graveyard
59	sight/historic spots	other
61	Other	volcanic zone
62	Other	fault
63	Other	mine
69	Other	other
99	Indistinguishable	Indistinguishable

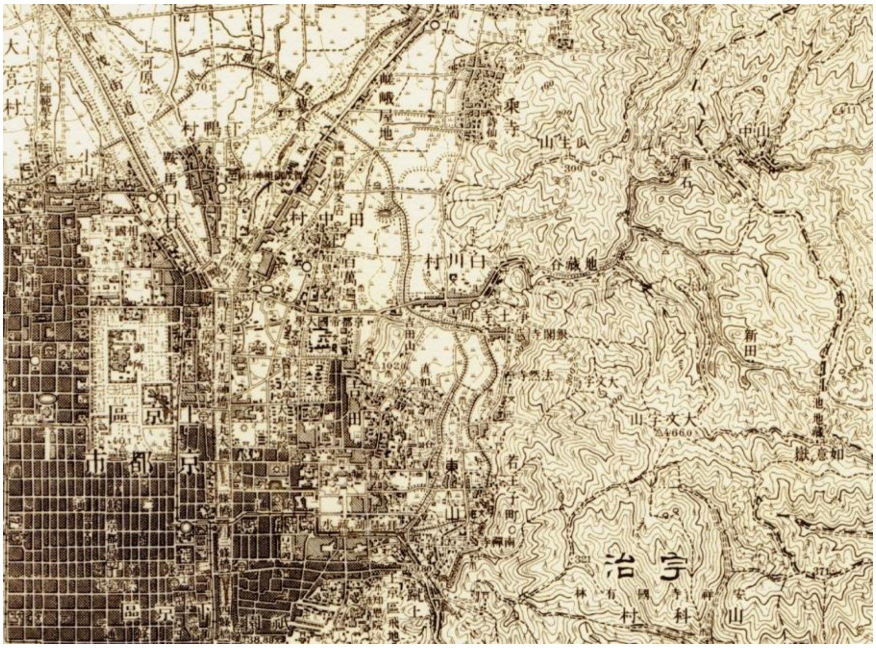


Figure 5: Example of *Kyū Go Manbun no Ichi Chikeizu* 旧５万分１地形図 (1:50,000 Old Topographic Maps: MAP in Kyoto area)

3. **Collecting Place Names**: All visual strings on the maps are collected as historical place names. A longitude and latitude pair of each place name was identified by following the guidelines below (Figure 6):
   - If the place name is depicted by a map symbol (e.g., town hall, bridge, station, and mountain summit), its location will be defined as the center of the symbol (Figure 6 upper left).
   - If the place name is depicted by an enclosed area (e.g., forest, lake, and town boundary), its location will be defined as the center of the area (Figure 6 upper right).
   - If the place name is not depicted by a map symbol but clearly specified (e.g., cape tip, mountain peak), its location will be defined as the point (Figure 6 lower left).
   - Otherwise, if the place name is depicted by only by a string, its location will be defined as the center of the string (Figure 6 lower right).
4. **Creation of Polygons**: Administrative boundaries, rivers, canals, lakes, marshes are described by polygons.

The main data elements and description rules of *Kyū Go Manbun no Ichi Chikeizu* (MAP) are compatible with those of *Dai Nihon Chimei Jisho* (DGJ), *Engishiki Jinmyōchō* (EJM), and *Nihon Ji'in Sōran* (DJT).

2.3 The DGHJ Database

At the time of writing this paper, the DGHJ includes total of 377,471 historical place names, of which, 53,528 historical place names from *Dai Nihon Chimei Jisho* (DGJ), 2,842 historical place names from *Engishiki Jinmyōchō* (EJM), 78,557 historical place names from Nihon *Ji'in Sōran* (DJT), and 242,544 historical place names from *Kyū Go Manbun no Ichi Chikeizu* (MAP). Figure 7 shows examples of the distributions of historical place names collected from *Dai Nihon Chimei Jisho* (DGJ), *Engishiki Jinmyōchō* (EJM), and *Kyū Go Manbun no Ichi Chikeizu* (MAP).

The DGHJ provides REST-like API which is a convenient tool for users to create mashup applications by combining the existing applications and data sets (Fielding, 2000). The outline of the DGHJ API specification is:

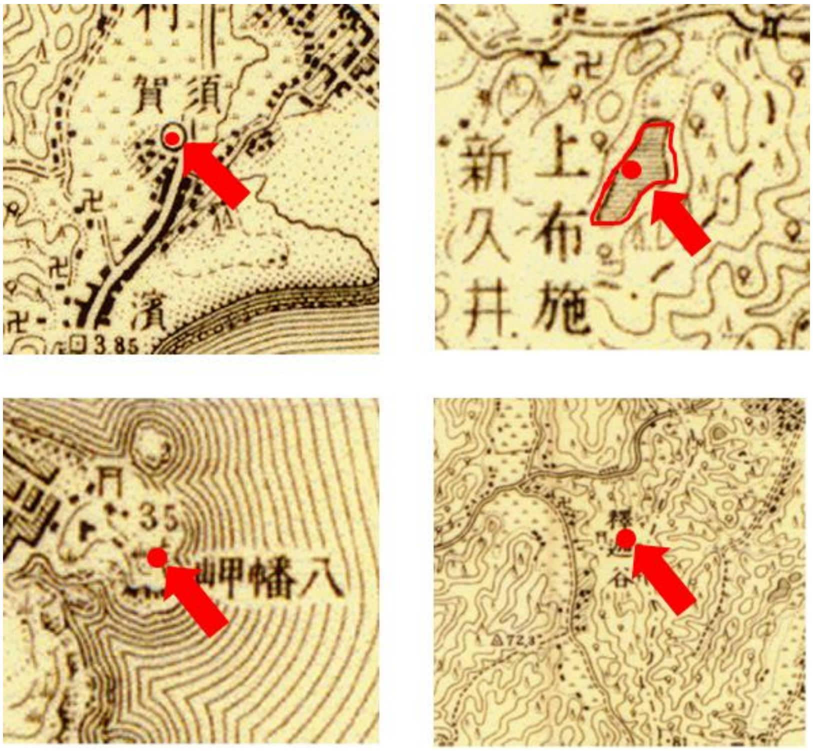
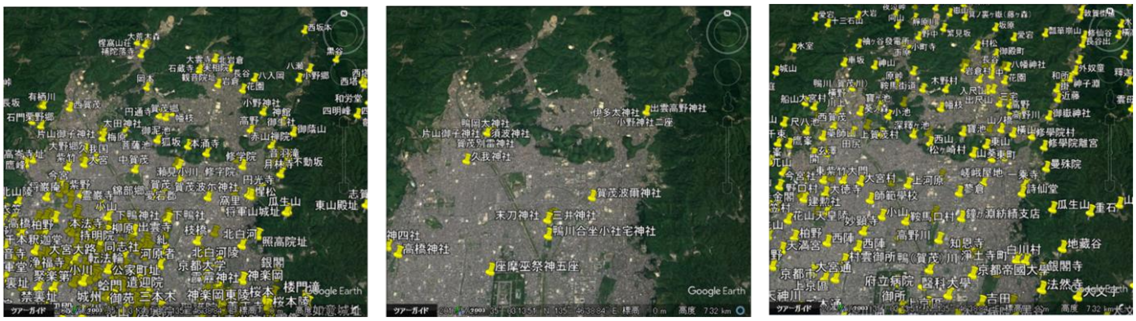


Figure 6: Identification of longitude and latitude pairs in *Kyū Go Manbun no Ichi Chikeizu* (MAP)



Dai Nihon Chimei Jisho (DGJ) *Engishiki Jinmyōchō* (EJM) *Kyū Go Manbun no Ichi Chikeizu* (MAP)

Figure 7: Example distributions of historical place names collected from *Dai Nihon Chimei Jisho* (DGJ), *Engishiki Jinmyōchō* (EJM) and *Kyū Go Manbun no Ichi Chikeizu* (MAP) (Base Map: Google, 2019)

```
AIP: api_base_url '/' Database_ID '?' Parameters
Parameters: 'operation'= ('searchRetrieval' | 'explain')
        '&' 'version' = '1.2' ('&'Query)?
        '&' 'recordSchema' = ('mods' | 'dc' | 'original')
        ('&' other)*
Query: 'query' = 'not'? (Term Op Keyword) (('and'|'or')
        'not'? (Term Op keyword))*
Term: Original_Field_Name|Mods_Terms|DC_Terms|
     'cql.anywhere'|
Op: '=' | '==' | '<>' | etc.
```

Following is a query example to search for Shōkokuji 相国寺 (Shōkoku Temple) and to get its result in JSON format (IETF, 2017). Here "%E7%9B%B8%E5%9B%BD%E5%AF%BA " is URL encode of "相国寺" in UTF-8.

```
http://API_Base_URI/Database_ID?operation=search
Retrieve&version=1.2&query=c1=%22%E5%A4%A7%E5
%92%8C%22&recordSchema=originaLxml
```

and its return is:

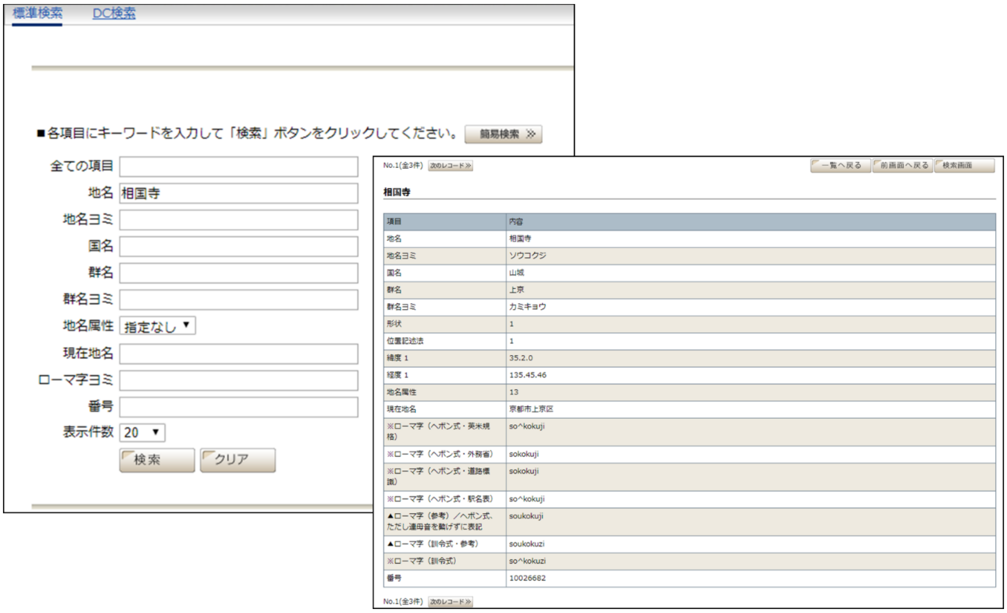
```
{"numberOfRecords":"3","recordData":
 [{"c1":"10026682","c2":"10026682","c3":"相国寺","c4":"
  ソウコクジ","c5":"35.03333333","c6":"135.7627778",
  "c7":[],"c8":[],"c9":[],"c10":[],"c11":[],"c12":[],"c13":"1",
  "c14":"1","c15":"13","c16":"山城","c17":[],
  "c18":"上京区","c19":"カミキョウ","c20":[],"c21":[],
  "c22":"京都市上京区////","c23":"京都市上京区////",
  "c24":[],"c25":"大日本地名辞書","c26":[],
  "c27":"30048501：相国寺と重複","c28":[],"c29":[],
  "30":[],"c31":[],"c32":[],"c33":[],"c34":[],"c35":[],"c36":[]},
  {"c1":"30027003","c2":"30027003","c3":"相国寺","c4":[],
  "c5":"37.18087159","c6":"138.6812922","c7":[],"c8":[],
  "c9":[],"c10":[],"c11":[],"c12":[],"c13":"1",c14":"1","c15":"92",
  "c16":"越後","c17":[],"c18":"中魚沼郡",
  "c19":"ナカウオヌマ",c20":[],"c21":[],"c22":"十日町市",
  "c23":"十日町市","c24":[],c25":"寺院名鑑",
  "c26":[],"c27":[],"c28":[],"c29":[],30":[],"c31":[],"c32":[],
  "c33":[],"c34":[],"c35":[],"c36":[]},
 {"c1":"30048501","c2":"30048501","c3":"相国寺","c4":[],
  "c5":"35.02958418",・・・・・・・・・・・・・}]
 "itemset":
 {"c1":"ID","c2":"OriginalID","c3":"地名","c4":"地名よみ",
  "c5":"緯度 tky",c6":"経度 tky","c7":"緯度 wgs",
  "c8":"経度 wgs","c9":"緯度 tky2","c10":"経度 tky2",
  "c11":"経度 wgs2","c12":"緯度 wgs2","c13":"形状",
  "c14":"位置記述法","c15":"地名属性",・・・・・・・
  "c35":"郡名ローマ字よみ","c36":"国"}
}
```

DGHJ Web GUI was built using this API. Figure 8 shows an example screenshot of the Web GUI.

A typical usage of historical gazetteers is "address matching" which is a data service used to convert a historical place name into the pair of longitude and latitude. Figure 9 is an example of address matching using the DGHJ which converted into pairs of longitude and latitude the place names found in historical documents about an earthquake occurrence. The screen shot shows the distributions of seismic intensities together with the faults in the surrounding area where the earthquake occurred.

3. Reconstruction of the DGHJ for Open Data Environment

Though the DGHJ is a large-scale historical gazetteer database in Japan, the DGHJ alone is unable to cover all historical place names. We consider that linking with other gazetteer databases is a possible solution to allow our digital gazetteers to offer more place names to users. Web APIs are convenient tools to create a mashup that is a single new Web service by linking contents from more than one database. However, as the usage of APIs is different among databases, programmers need to understand the specifications of all related databases' APIs before they write custom codes.



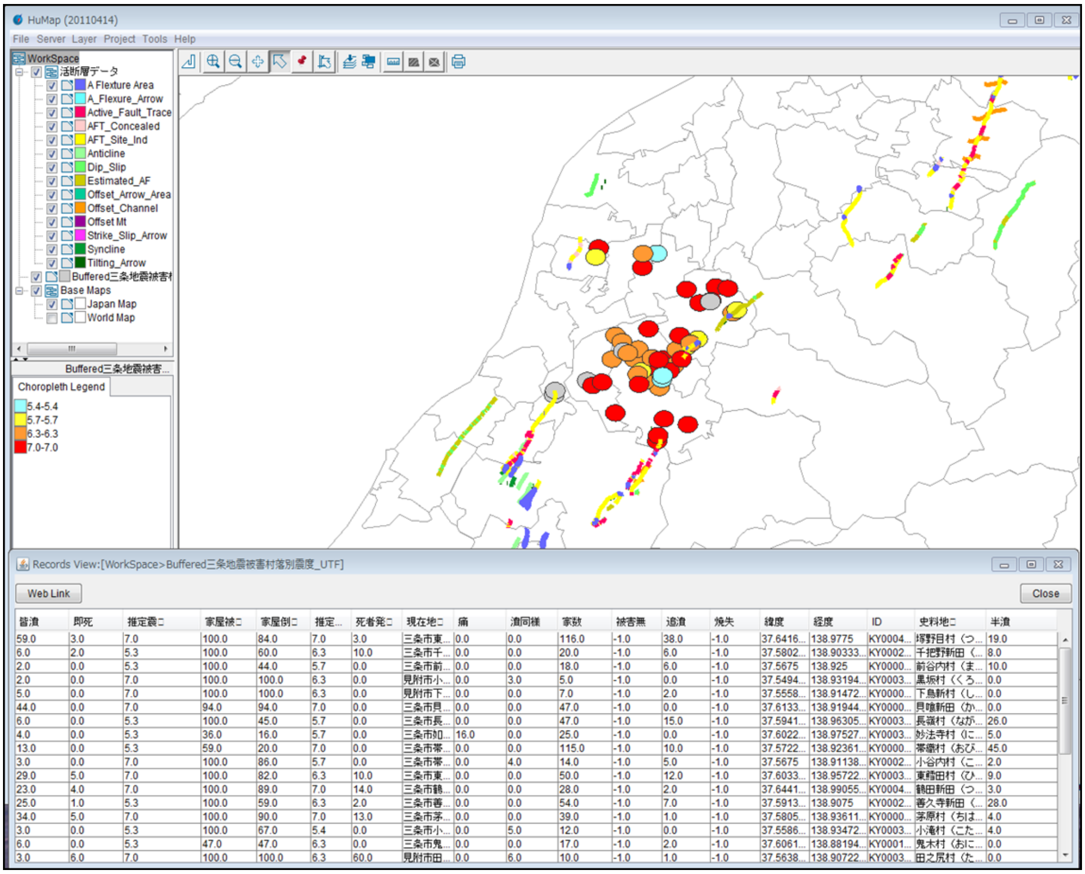
■各項目にキーワードを入力して「検索」ボタンをクリックしてください。

全ての項目	
地名	相国寺
地名ヨミ	
国名	
群名	
群名ヨミ	
地名属性	指定なし ▾
現在地名	
ローマ字ヨミ	
番号	
表示件数	20 ▾

相国寺

項目	内容
地名	相国寺
地名ヨミ	ソウコクジ
国名	山城
群名	上京
群名ヨミ	カミキョウ
形状	1
位置記述法	1
緯度 1	35.2.0
経度 1	135.45.46
地名属性	13
現在地名	京都市上京区
※ローマ字（ヘボン式・英米規格）	so^kokuji
※ローマ字（ヘボン式・外務省）	sokokuji
※ローマ字（ヘボン式・道路標識）	sokokuji
※ローマ字（ヘボン式・駅名表）	so^kokuji
▲ローマ字（参考）／ヘボン式、ただし連母音を繋げずに表記	soukokuji
▲ローマ字（訓令式・参考）	soukokuzi
※ローマ字（訓令式）	so^kokuzi
番号	10026682

Figure 8: A search screen example of the DGHJ (Retrieving Shōkokuji 相国寺 (Shōkoku Temple))



Records View:[WorkSpace>Buffered三条地震被害村落別震度_UTF]

皆潰	即死	推定震□	家屋被□	家屋倒□	推定...	死者発□	現在地□	痛	潰同様	家数	被害無	追潰	焼失	緯度	経度	ID	史料地□	半潰
59.0	3.0	7.0	100.0	84.0	7.0	3.0	三条市東...	0.0	0.0	116.0	-1.0	38.0	-1.0	37.6416...	138.9775	KY0004...	塚野目村（つ...	19.0
6.0	2.0	5.3	100.0	60.0	6.3	10.0	三条市千...	0.0	0.0	20.0	-1.0	6.0	-1.0	37.5802...	138.90333...	KY0002...	千把野新田（...	8.0
2.0	0.0	5.3	100.0	44.0	5.7	0.0	三条市前...	0.0	0.0	18.0	-1.0	6.0	-1.0	37.5675	138.925	KY0000...	前谷内村（ま...	10.0
2.0	0.0	7.0	100.0	100.0	6.3	0.0	見附市小...	0.0	3.0	5.0	-1.0	0.0	-1.0	37.5494...	138.93194...	KY0003...	黒坂村（くろ...	0.0
5.0	0.0	7.0	100.0	100.0	6.3	0.0	見附市下...	0.0	0.0	7.0	-1.0	2.0	-1.0	37.5558...	138.91472...	KY0000...	下鳥新村（し...	0.0
44.0	0.0	7.0	94.0	94.0	7.0	0.0	三条市貝...	0.0	0.0	47.0	-1.0	0.0	-1.0	37.6133...	138.91944...	KY0000...	貝喰新田（か...	0.0
6.0	0.0	5.3	100.0	45.0	5.7	0.0	三条市長...	0.0	0.0	47.0	-1.0	15.0	-1.0	37.5941...	138.96305...	KY0003...	長嶺村（なが...	26.0
4.0	0.0	5.3	36.0	16.0	5.7	0.0	三条市加...	16.0	0.0	25.0	-1.0	0.0	-1.0	37.6022...	138.97527...	KY0003...	妙法寺村（に...	5.0
13.0	0.0	5.3	59.0	20.0	7.0	0.0	三条市帯...	0.0	0.0	115.0	-1.0	10.0	-1.0	37.5722...	138.92361...	KY0000...	帯織村（おび...	45.0
3.0	0.0	7.0	100.0	86.0	5.7	0.0	三条市帯...	0.0	4.0	14.0	-1.0	5.0	-1.0	37.5675	138.91138...	KY0002...	小谷内村（こ...	2.0
29.0	5.0	7.0	100.0	82.0	6.3	10.0	三条市東...	0.0	0.0	50.0	-1.0	12.0	-1.0	37.6033...	138.95722...	KY0003...	東鱈田村（ひ...	9.0
23.0	4.0	7.0	100.0	89.0	7.0	14.0	三条市鶴...	0.0	0.0	28.0	-1.0	2.0	-1.0	37.6441...	138.99055...	KY0004...	鶴田新田（つ...	3.0
25.0	1.0	5.3	100.0	59.0	6.3	2.0	三条市善...	0.0	0.0	54.0	-1.0	7.0	-1.0	37.5913...	138.9075	KY0002...	善久寺新田（...	28.0
34.0	5.0	7.0	100.0	90.0	7.0	13.0	三条市茅...	0.0	0.0	39.0	-1.0	1.0	-1.0	37.5805...	138.93611...	KY0000...	茅原村（ちは...	4.0
3.0	0.0	5.3	100.0	67.0	5.4	0.0	三条市小...	0.0	5.0	12.0	-1.0	0.0	-1.0	37.5586...	138.93472...	KY0003...	小滝村（こた...	4.0
6.0	0.0	5.3	47.0	47.0	6.3	0.0	三条市鬼...	0.0	0.0	17.0	-1.0	2.0	-1.0	37.6061...	138.88194...	KY0001...	鬼木村（おに...	0.0
3.0	6.0	7.0	100.0	100.0	6.3	60.0	見附市田...	0.0	6.0	10.0	-1.0	1.0	-1.0	37.5638...	138.90722...	KY0003...	田之尻村（た...	0.0

Figure 9: An Example of the DGHJ application as address matching

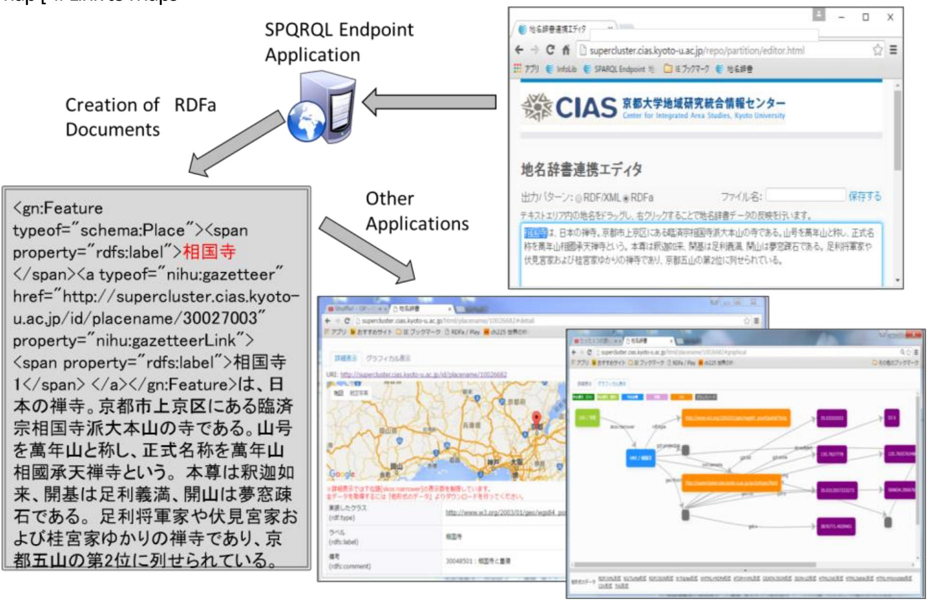
3.1 New DGHJ using RDF and SPARQL

Recent RDF (Resource Description Framework) (W3C, 2004) and related technologies (e.g., SPARQL (W3C, 2013)) are the basis of “linked open data” that is a sophisticated mechanism of developing APIs to integrate heterogeneous data so that fragmented data/knowledge can be linked and become more useful through semantic queries. These are appropriate technologies to link the DGHJ with other gazetteer databases in a standardized way. We have reconstructed the DGHJ according to RDF specifications and built a new application interface using SPARQLE Endpoint (RDF-DGHJ) (Hara, 2016 and 2017).

RDF-DGHJ data keeps original DGHJ data structure, and RDF-DGHJ data were created from original XML data almost automatically. Following is an example place name of RDF-DGHJ data in Turtle format (name spaces are ignored):

```
<http://base_uri/placename/id/245009>
  geo:lat "35.64444";
  geo:long "139.4545";
  gzt:lat "35°38′ 39.984″ ";
  gzt:long "139°27′ 16.2″ ";
  gzt:x "387203.042";
  gzt:y "3953623.327";
  gzt:zone "54";
  rdfs:label "千ケ崎";
  dc:subject [ # Link to Attributes
      rdfs:label "地方" ;
      dcq:subject <http://base_uri/placeattribute/id/02>
  ];
  gzt:country "*";
  rdfs:comment "Memo";
  gm:geomap [ # Link to Maps
      rdfs:label "大島";
      owl:sameAs <http://base_uri/map/id/000000>;
      gm:north <http://base_uri/map/id/000001>;
      gm:west <http://base_uri/map/id/000002>;
      gm:east <http://base_uri/map/id/000003>;
      gm:south <http://base_uri/map/id/000004>
  ];
  vcard:region "東京都".
```

The RDF-DGHJ implements SPARQL Endpoint as its basic query interface. Figure 10 shows an example of RDF-DGHJ tools using SPARQL Endpoint, which is a simple editor on Web viewers. Firstly, a text file including historical place names is prepared and uploaded onto the editor (upper right). Secondly, when a place name is selected by dragging a pointer, the tool retrieves the RDF-DGHJ and lists the candidate historical place names. When an appropriate place name is selected, the tool gets detailed information of the place name from the RDF-DGHJ and inserts it into the original text in the format of RDFa (Resource Description Framework in attributes (W3C, 2015) and creates a new HTML text which includes the link to the RDF-DGHJ (left). Using this link information in the HTML texts, users can access RDF-DGHJ data and further RDF repositories (lower right).



```
<gn:Feature
typeof="schema:Place"><span
property="rdfs:label">相国寺
</span><a typeof="nihu:gazetteer"
href="http://supercluster.cias.kyoto-u.ac.jp/id/placename/30027003"
property="nihu:gazetteerLink">
<span property="rdfs:label">相国寺
1</span> </a></gn:Feature>は、日本の禅寺。京都市上京区にある臨済宗相国寺派大本山の寺である。山号を萬年山と称し、正式名称を萬年山相國承天禅寺という。本尊は釈迦如来、開基は足利義満、開山は夢窓疎石である。足利将軍家や伏見宮家および桂宮家ゆかりの禅寺であり、京都五山の第2位に列せられている。
```

supercluster.cias.kyoto-u.ac.jp/repo/partition/editor.html

CIAS 京都大学地域研究統合情報センター Center for Integrated Area Studies, Kyoto University

地名辞書連携エディタ

出力パターン: ○ RDF/XML ● RDFa　ファイル名: 保存する

テキストエリア内の地名をドラッグし、右クリックすることで地名辞書データの反映を行います。

相国寺は、日本の禅寺。京都市上京区にある臨済宗相国寺派大本山の寺である。山号を萬年山と称し、正式名称を萬年山相國承天禅寺という。本尊は釈迦如来、開基は足利義満、開山は夢窓疎石である。足利将軍家や伏見宮家および桂宮家ゆかりの禅寺であり、京都五山の第2位に列せられている。

Figure 10: A tool using the RDF-DGHJ and SPARQL Endpoint

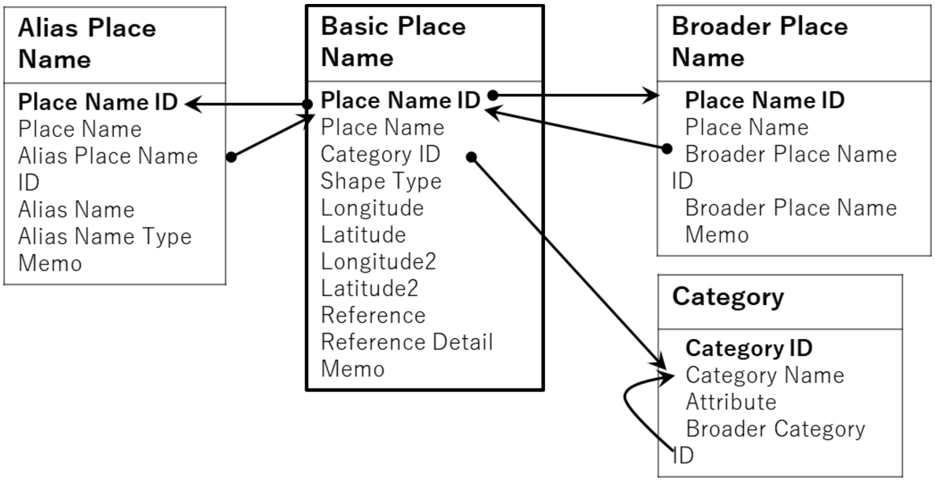


Figure 11: New data schema of the DGHJ

3.2 Standardization and Interoperability

The DGHJ uses intrinsic vocabularies to organize historical place names and their attributes. Thus, the compatibility with other place name databases has been disregarded. Even though the RDF-DGHJ has interlinking functions, this circumstance makes the RDF-DGHJ difficult to link related information on the Web. As a first step for the solution, we defined "Basic Place Names" which were selected as minimum data items to geocode DGHJ records. Other data items will be used for advanced applications. Figure 11 shows the new data schema of the DGHJ. The second step is to define relationships between our intrinsic vocabularies of the DGHJ and vocabularies used in well-known place name databases. As preliminary research, we surveyed GeoNames (GeoNames), and the Getty Thesaurus of Geographic Names (Getty) as well-known data sets of place names. The GeoNames is an open geographic database which comprises in large part the data from the United States Geological Survey (USGS which provides the geographic database for current places within the USA (SUGS)) and from the National Geospatial Intelligence Agency (NGA which provides the geographic database for current places of all nations outside the USA (NGA)). Names included in the GeoNames usually refer to current places. The Getty Thesaurus of Geographic Names includes names of current, historical and lost places, and it has diversified and rich information, such as, administrative hierarchy, main historical transitions, and general information about history, culture, art and architecture.

Table 2 shows tentative definitions of vocabulary mapping between the DGHJ, GeoNames, and the Getty Thesaurus of Geographic Names. This result implies the possibility of interoperability between the DGHJ and other well-known digital gazetteers.

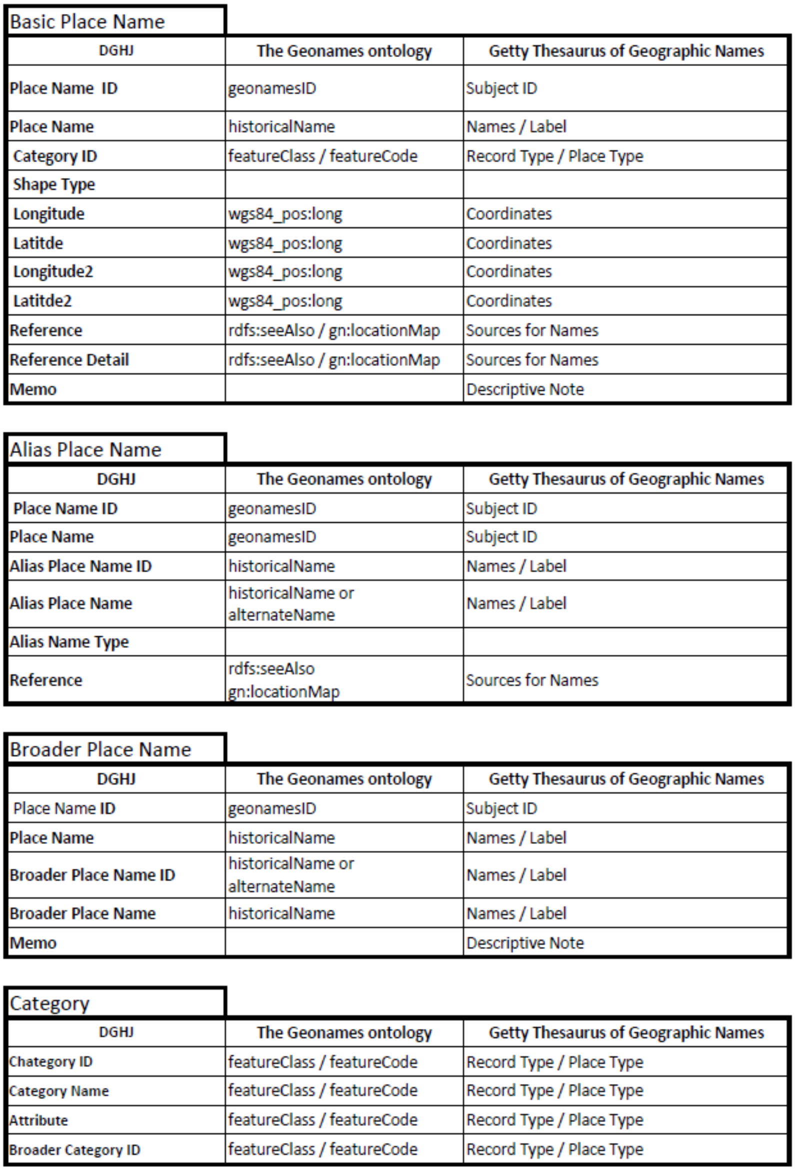
4. Problems and Considerations

Academic data, especially humanities' data, comprise small fragments of heterogeneous data, which have a possibility to generate big data through associations among each other. This is different feature from ordinary big data which usually have simple structures and are collected from sensors. Ontology has been thought to be an appropriate technology to associate heterogeneous information by referencing to varied dictionaries or thesaurus. We have developed the DGHJ and the RDF-DGHJ as knowledge bases for open data environment that allows to associate various information in the context of geographical proximity. As explaining in the preceding sections, our digital gazetteers show a potential to realize this objective. However, some problems emerged from the process and solutions for them must be sorted out.

4.1 Processing KANJI (Chinese characters)

The DGHJ includes many place names before the 19th century, that is, there are some KANJI (Chinese characters) that are not registered in Unicode character sets (ISO/IEC10646). We have tried to associate the old character forms or variant character forms with Unicode characters as faithfully as possible. However, when an appropriate KANJI (Chinese characters) was not found in Unicode character sets, "#" is used to indicate it as an external standard KANJI (Chinese characters). For effective and intelligent retrieval of historical gazetteer databases, we need dictionaries that serve to associate KANJI (Chinese characters) which have the same meaning with different forms. KANJI (Chinese characters) processing in databases has continuously been a problem since computer's dawn in Japan.

Table 2: Vocabulary mapping between the DGHJ, GeoNames, and the Getty Thesaurus of Geographic Names



Basic Place Name

DGHJ	The Geonames ontology	Getty Thesaurus of Geographic Names
Place Name ID	geonamesID	Subject ID
Place Name	historicalName	Names / Label
Category ID	featureClass / featureCode	Record Type / Place Type
Shape Type		
Longitude	wgs84_pos:long	Coordinates
Latitde	wgs84_pos:long	Coordinates
Longitude2	wgs84_pos:long	Coordinates
Latitde2	wgs84_pos:long	Coordinates
Reference	rdfs:seeAlso / gn:locationMap	Sources for Names
Reference Detail	rdfs:seeAlso / gn:locationMap	Sources for Names
Memo		Descriptive Note

Alias Place Name

DGHJ	The Geonames ontology	Getty Thesaurus of Geographic Names
Place Name ID	geonamesID	Subject ID
Place Name	geonamesID	Subject ID
Alias Place Name ID	historicalName	Names / Label
Alias Place Name	historicalName or alternateName	Names / Label
Alias Name Type		
Reference	rdfs:seeAlso gn:locationMap	Sources for Names

Broader Place Name

DGHJ	The Geonames ontology	Getty Thesaurus of Geographic Names
Place Name ID	geonamesID	Subject ID
Place Name	historicalName	Names / Label
Broader Place Name ID	historicalName or alternateName	Names / Label
Broader Place Name	historicalName	Names / Label
Memo		Descriptive Note

Category

DGHJ	The Geonames ontology	Getty Thesaurus of Geographic Names
Chategory ID	featureClass / featureCode	Record Type / Place Type
Category Name	featureClass / featureCode	Record Type / Place Type
Attribute	featureClass / featureCode	Record Type / Place Type
Broader Category ID	featureClass / featureCode	Record Type / Place Type

Moreover, the varied pronunciations of KANJI (Chinese characters) are important information in the process of retrieving Japanese historical gazetteer databases. However, in Japanese language, one KANJI (Chinese characters) has often more than two pronunciations (e.g., the same word 登戸 is pronounced "Nobuto" in one place but pronounced as "Noborito" in another place). Furthermore, even if the same place name is written by the same KANJI (Chinese characters), its pronunciation might change over time. Identification of the pronunciation of KANJI (Chinese characters) is also difficult and remains an unresolved problem.

4.2 Closed Gazetteer Data

The DGHJ includes many place names related to segregated areas according to social class and ethnic variations, which made it difficult to publish the DGHJ for a public use. We have exchanged opinions between related stakeholders and have been seeking ways to include this information for an academic use. At last, from March 2018, the National Institutes of Humanities (NIHU) and the Humanities' Research Group (H-GIS) have started the downloading service of DGHJ data (except *Nihon Ji'in Sōran* (DJT) (H-GIS and NIHU, 2018). Even if the data downloading service is under progress, careful handling of the information is necessary.

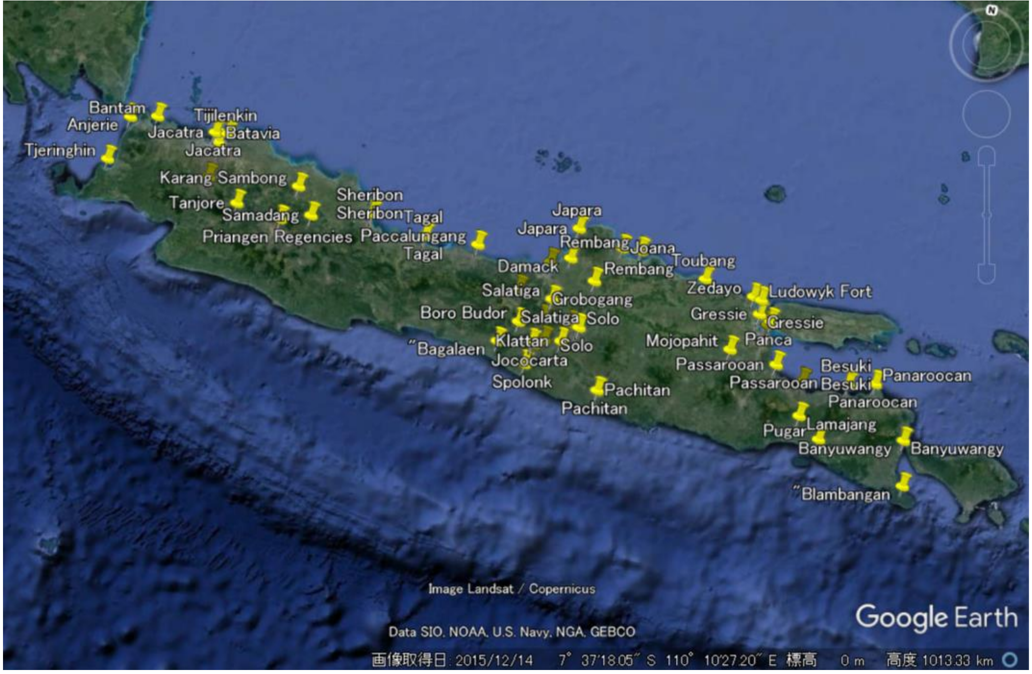


Figure 12: Example distribution of place names in Java collected from the East India Gazetteer

4.3 New Gazetteers

As members of research institutes of area studies, we have tried to apply the experiences of developing historical gazetteer databases and began creating a new historical gazetteer database about Southeast Asia and Southern Asia. In order to evaluate the possibility to use the methods developed so far, we use the East India Gazetteer edited by Walter Hamilton in 1815 (Hamilton, 1815), and try to extract historical place names, identify their corresponding contemporary place names, and determine the longitude and latitude pairs.

Currently, we have compiled and identified the current location of 665 historical place names (Borneo (29), Celebes (24), Ceylon (69), Indochina (24), Nanyang Islands (284), Java (75), Mindanao (11), Melaka (28), Malabar (48), Papua (4), Philippines (11), and Sumatra (58): including the place names that overlap spatially). Figure 12 shows an example of distribution of historical place names in Java collected from the East India Gazetteer.

4.4 Interoperability

Vocabularies and attributes used in the DGHJ are defined without considering the interoperability with other gazetteer databases. The RDF-DGHJ uses the same attributes and this plays a key role to define RDF data structure (Hara and Sekino, 2017), which may bring about an obstruction in linking to other gazetteer databases. Mapping vocabularies as discussed in the Section 3 is the first step to solve these problems. KANJI (Chinese characters) processing in databases is another problem that impedes the interoperability of historical gazetteer databases. Some research groups engaged in building historical gazetteer databases in East Asian areas encountered similar problems. We began discussing with these research groups in order to find mutual and appropriate solutions which will allow for the interoperability of gazetteer databases (Hara et al., 2018).

Acknowledgement

The authors would like to show our greatest appreciation to the late Prof. Ikuo OKETANI for his efforts on the early stages of this work. We would like to thank the members of H-GIS Research Group members for years of collaboration and advices. The work is supported by JSPS Grant-in-Aid for Scientific Research (A) Number 16H01897, and Glocal Information Platform of the Center for Southeast Asian Studies Kyoto University.

References

Center for Spatial Information Science, University of Tokyo (CSIS), Geocoding Tools and Utilities, http://newspat.csis.u-tokyo.ac.jp/geocode/

Dai Nippon Teikoku Rikuchi Sokuryōbu 大日本帝國陸地測量部 (The Japanese Imperial Army's Land Survey Bureau), 1890-1916, *Kyū Go Manbun no Ichi Chikeizu* 旧５万分１地形図 (1:50,000 Topographic Maps), *Kokudo Chiri'in* 国土地理院 (The Geospatial Information Authority of Japan).

Fielding, R. T., 2000, Architectural Styles and the Design of Network-based Software Architectures, https://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm

GeoNames, http://www.geonames.org/

Getty Research Institute, Getty Thesaurus of Geographic Names, http://www.getty.edu/research/tools/vocabularies/tgn/index.html

Hamilton, W., 1815, The East India Gazetteer: Containing Particular Descriptions of the Empires, Kingdoms, Principalities, Provinces, Cities, Towns, Districts, Fortresses, Harbours, Rivers, Lakes, &c. of Hindostan, and the Adjacent Countries, India Beyond the Ganges, and the Eastern Archipelago; Together with Sketches of the Manners, Customs, Institutions, Agriculture, Commerce, Manufactures, Revenues, Population, Castes, Religion, History, &c. of their Various Inhabitant, Printed for J. Murray by Dove, https://archive.org/details/eastindiagazette00hami

Hara, S., 2016, Open Platform for Academic Humanities Data, International Symposium on Grids and Clouds 2016 (ISGC 2016), https://indico4.twgrid.org/indico/event/1/session/18/contribution/69

Hara, S., 2017, Digital Gazetteer as a Knowledgebase for Open Data Science, Pacific Neighborhood Consortium Annual Conference and Joint Meetings (PNC) 2017, DOI: 10.23919/PNC.2017.8203524, IEEE Xplore.

Hara, S., Sekino, T. and Liao, H. M., 2018, International Workshop on Spatio-Temporal Knowledge, http://gis.rchss.sinica.edu.tw/documents/workshop_2018052425.pdf.

Humanities' GIS Research Group (H-GIS), http://www.h-gis.org/.

H-GIS and NIHU, 2018, Historical Gazetteer Data, http://www.nihu.jp/ja/publication/source_map.

Internet Engineering Task Force (IETF), 2017, The JavaScript Object Notation (JSON) Data Interchange Format, https://tools.ietf.org/html/std90.

Ji'in Sōran Kankōkai 寺院総鑑刊行会 (Publishing committee of the Directory of Japanese Temples), 2000, *Nihon Ji'in Sōran* 日本寺院総鑑 (Directory of Japan Temples: DJT), Kotobuki Kikaku (寿企画).

Kadokawa Nihon Chimei Daijiten Henshū Iinkai 角川日本地名大辞典編纂委員会 (Kadokawa editing committee of The New Edition of the Kadokawa Geographical Dictionary of Japan), 2011, *Nihon Chimei Daijiten DVD-ROM* 新版角川日本地名大辞典 DVD-ROM (The New Edition of the Kadokawa Geographical Dictionary of Japan DVD-ROM), Kadokawa Publishing (角川書店).

Kuroita, K., ed., 1979, *Shintei Zōho Kokushi Taikei* 新訂増補国史大系 (Newly revised and enlarged survey of Japanese history), Yoshikawa Kobunkan(吉川弘文館).

National Geospatial Intelligence Agency （NGA), https://www.nga.mil/Pages/Default.aspx.

National Institutes for the Humanities (NIHU), https://www.nihu.jp/en.

Oketani, I., 2007, The Development the Gazetteer of Japanese Place Names based on Humanities and the Feature Analysis of Place Name Attribute, *IPSJ Symposium Series*, Vol. 2017, No.15, 79-86.

Oketani, I., 2009, The Database of Topographical Maps and Place Names, *IPSJSIG Technical Report*, Vol. 2009-CH-83, No. 3, 1-8.

Shikinaisha Kenkyukai 式内社研究会 (Research Group of Shrines Listed in the Engishiki) , 1979, *Shikinaisha Chōsa Hokoku* 式内社調査報告 (The Report of the Shikinaisha Survey), Kogakkann Daigaku Shuppanbu 皇学館大学出版部 (Kogakkann University Press).

The National Institutes for the Humanities (NIHU), https://www.nihu.jp/ja.

The United States Geological Survey （USGS), https://www.usgs.gov/.

W3C, 2004, Resource Description Framework (RDF): Concepts and Abstract Syntax, https://www.w3.org/TR/2004/REC-rdf-concepts-20040210/.

W3C, 2013, SPARQL 1.1 Query Language, https://www.w3.org/TR/sparql11-query/.

W3C, 2015, RDFa 1.1 Primer - Third Edition, https://www.w3.org/TR/rdfa-primer/.

Yoshida, T., 1900, *Dai Nihon Chimei Jisho* 大日本地名辞書 (The Dictionary of Geographical Names of Japan: DGJ), Fuzambo (冨山房).

Yotsui, K., Sekino, T., Hara, S., Oketani, I. and Shibayama, M., 2010, Construction of an Historical Gazetteer Based on 1:50000 Maps from the Meiji and Taisho Eras, *I IPSJ Symposium Series*, Vol. 2010, No.15, 211-216.