International Cartographic Association
Association Cartographique Internationale

Development of a Japanese 'GIS place name dictionary' as a basis for interdisciplinary research

Keiji Yano [a, *], Kazumasa Hanaoka [a], Ryo Kamata [a], Narushige Shiode [b], Shino Shiode [c]

[a] *Ritsumeikan University, yano@lt.ritsumei.ac.jp, kht27176@fc.ritsumei.ac.jp, kamata.y.ryo@gmail.com*
[b] *Kingston University, n.shiode@kingston.ac.uk*
[c] *Birkbeck, University of London, s.shiode@bbk.ac.uk*

* Corresponding author

Keywords: place name, gazetteer, GIS, Japanese old maps, digital humanities

Abstract:

The aim of this research is to develop a “GIS Place Name Dictionary”, an interactive place name dictionary (Gazetteer) that can search geospatial information from all regions of Japan, from past historical designations to current place names and addresses, as well as identify longitude and latitude and map representation. Area statistics and historical data used in the humanities and social sciences often include place names and addresses, which are geospatial information, in addition to changes in names over time. By identifying a place by its name or address and visualising various phenomena that occurred there in the past on a map, the spatial location of these phenomena can be clarified, and by superimposing them on other maps, new knowledge in fields such as literature, history, sociology and economics, which have not sufficiently incorporated a spatial perspective, can be discovered. New knowledge is expected to be discovered. The database that will serve as the basis for this is the GIS Place Name Dictionary.

Furthermore, in recent years, the field of Digital Humanities (DH) has expanded in Europe and the US, and new knowledge is being discovered from large-scale digitised academic data through data-driven machine analysis, which could not be achieved in traditional humanities research conducted by human hands. For example, if the digitisation, reprinting and textualisation of travel diaries and other documents from a past period progresses, it will be possible to easily map the places visited on a journey using the 'GIS place name dictionary' to be constructed in this research. Furthermore, by superimposing the location of place names in past documents related to historical disasters with current hazard maps, it will be possible to estimate disaster risks from historical place names in the past. Such research is positioned in historical GIS, which deals with spatial perspectives in DH research, and collaboration with traditional literature, history, archaeology and other humanities fields is being promoted, with the expectation of creating new knowledge by utilising comprehensive knowledge in the humanities and social sciences.

Geocoding (address matching) is a GIS process that identifies the latitude and longitude of representative points from the textual information of place names and addresses and displays the locations on digital maps as point and area data, and the place name dictionary used in this process is called a geocoder. The geocoding system is an information processing process that normalises fluctuations in place name and address notation, but a geocoder is a database of place names, addresses and their location coordinates, which someone has to create. The basic geocoders have been created by national and local governments in the course of administrative digitalisation, and many of them are open to the public as open data. The private sector has also created its own geocoders based on them, and there are many commercially based geocoders. However, existing geocoders have fluctuations in address notation between the organisations that created them (e.g. in the national government, the Ministry of Land, Infrastructure, Transport and Tourism and the Ministry of Internal Affairs and Communications use different town and street names), and new addresses are frequently set and changed even after creation, which means that geocoders need to be continuously updated and address notation normalisation is necessary.

Such a 'GIS place name dictionary' is also indispensable for the provision of administrative services to residents and business establishments at the national and local level, and the Digital Agency of Japan, which will be established in 2021, will position the address and location master data as a social infrastructure database that is held by public institutions and referred to in a variety of situations. The Digital Agency, which will be established in 2021, has begun to position address and location master data as a base registry, a database held by public authorities and others that will serve as the foundation of society and be referred to in various situations. In collaboration with such trends, a 'GIS place name dictionary' using open data will be constructed and released as open data.

Abstracts of the International Cartographic Association, 6, 283, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-283-2023

This research proposes a system that enables the development of a 'GIS place name dictionary' for current place names and addresses for the whole of Japan, and for the past, for priority areas such as Kyoto (*Kyo*) and Tokyo (*Edo*), where there are many resources since they are historical cities from the past. In the future by enabling crowdsourcing, the dictionary can be continuously expanded with input from users in areas other than the priority areas.

(a) 'Integration of existing GIS-based 'GIS place name dictionaries'.

At present, GIS data on place names and addresses provided by the State as open data include location reference information at the Oaza/cho/chome level and street level (Ministry of Land, Infrastructure, Transport and Tourism), electronic national land basic maps (place name information) and residential indication addresses (GSI), and (statistical GIS) boundary data from the National Census tract (Cho-cho) boundaries (Statistics Bureau, Ministry of Internal Affairs and Communications) data, etc. However, those at the land number level are limited. There is also address data including place numbers from the private-sector Zmap-TOWN II of Zenrin CO., LTD. and NTT's telephone directories, but they are not compatible. The integration of all these GIS data and the construction of a 'GIS place name dictionary' is significant not only in making historical place names and addresses available over time, but also in integrating address data developed independently in both the public and private sectors.

(b) Development of a 'GIS place name dictionary' of past place names

Next, a 'GIS place name dictionary' of past place names will be developed by referring to the current one developed in (a) and going back from the present to the past, using the following data: 1) basic unit wards from the 1990 census, towns, streets and characters from the 1995 census, 2) towns, streets and characters from 1982 onwards. The 'GIS place name dictionary' will be created and analysed using the following data: 1) basic unit wards from the 1990 census, town, street and character codes from the Geographical Survey Institute's 11-digit codes from the 1995 census, 2) local government address information from the 1962 law on residential indication, 3) land registers and maps attached thereto (now the registry office) from the 1884 land tax article, 4) place name notes on Japanese old maps and old edition topographical maps, 5) place names in various place name dictionaries. 1) and 2) have been digitised and will be considered for purchase, while the others are basically in paper form, so a programme to acquire textual information and paper map location information by image recognition using AI technology (machines-reading-maps/map-kurator https://github.com/machines-reading-maps/map-kurator) will be developed for their digitisation and textualization (Chiang *et al.*, 2022). With regard to the latter use of paper maps (including old maps, etc.), we will use the "Japan's Old Maps Portal Site (https://www.dh-jac.net/db/maps/search_portal.php)", "Japanese Old Maps Online (https://japanese-old-maps-online-rstgis.hub.arcgis.com/)" and the "Japanese MapWarper (https://mapwarper.h-gis.jp/)", which corrects distortions in paper maps on the web (Yano *et al.*, 2019, 2021, 2022).

This presentation will report on the current status of 'GIS place name dictionaries' as open data, and will report on the status of the creation of past 'GIS place name dictionaries' and their challenges.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Numbers 22K18254, 22H00245, 20H00040.

References

Yano, K., Satoshi Imamura1, S., Kamata, R., 2019. Japanese Map Warper for Spatial Humanities: The Japanese old maps portal site. Abstr. Int. Cartogr. Assoc., 1, 418.

Yano, K., Imamura, S., Kamata. R., 2021. Construction of an old-map framework for promoting historical GIS research and education. Abstr. Int. Cartogr. Assoc., 3:1-2

Yano, K., Natsume, M., Imamura, S. and Kamata, R., 2022. Japanese Old Maps Online for Promoting Digital Humanities. Digital Humanities 2022 Conference Abstracts, 402-403.

Chiang, Y.-Y., Holmes-Wong, D., Kim, J., Li, Z., McDonough, K., Simon, R., Vitale, V.. 2022. Machines Reading Maps: from text on maps to linked spatial data. Digital Humanities 2022 Conference Abstracts, 443-445.