The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume L-4/W1-2026
Free and Open Source Software for Geospatial (FOSS4G) 2026 – Academic Track "Bridging Geospatial Technology and Humanity",
30 August–5 September 2026, Hiroshima, Japan

From Point Anchors to Geospatial Support: A Resolution-Aware Representation of Toponymic and Non-Toponymic Place Evidence in a Japanese Yokai Archive

Kosuke Shimizu[1], Hiroki Ichikura[1], Riri Ikebe[2], Mirai Hoshikawa[1]

[1] University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki 305-8577, Japan
[2] Japan Women's University, 2-8-1 Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan
Kosuke Shimizu: shimizu@ai.iit.tsukuba.ac.jp; Hiroki Ichikura: icchyworks@gmail.com;
Riri Ikebe: asagilabo11@gmail.com; Mirai Hoshikawa: hoshikawa@golem.iit.tsukuba.ac.jp

Keywords: Folklore geography, Geospatial support, Spatial humanities, Historical geocoding, Geographic resolution

Abstract

Folklore archives often identify a broad administrative region while their summaries mix named places with non-toponymic descriptions of rivers, roads, dwellings, ritual sites, and boundaries. Treating an administrative representative point as an event location can therefore make the mapped evidence appear more precise than the source permits. We present a resolution-aware representation for 33,378 records from the Database of Folktales of Mysterious Phenomena and Yokai. It separates administrative support areas, candidate toponyms, non-toponymic place descriptions, geographic proximity, human-condition terms, derived interfaces, display anchors, and provenance. Rule-based coverage is 28.0% for candidate toponyms and 75.0% for place descriptions; these values are not extraction accuracy because no human gold labels are available. Rule-constrained local matching assigns municipality support to 1,231 records across 452 units. Relative to prefecture support, municipality polygons reduce support area by a median 97.5% and change sampled median coastline distance by 9.2 km. Prefecture-stratified language associations recover expected Kappa–water and Yurei–death constructions, but these are construct checks rather than new folkloristic findings. A separate Kappa distance diagnostic does not support closer physical proximity to mapped major water. The contribution is an evidence-preserving, failure-aware baseline that distinguishes display anchors from the support areas warranted by the archive.

1. Introduction

Japanese yokai traditions have long been widely acknowledged through local narratives, collected sources, and changing cultural classifications (Foster, 2024). Digital archives make these materials searchable at scale, but a map adds claims that a catalogue does not: a point can suggest an event location, and a derived environmental label can suggest physical conditions. For short historical summaries, neither implication is safe unless the evidence supporting the geometry is retained.

The Database of Folktales of Mysterious Phenomena and Yokai, maintained by the International Research Center for Japanese Studies, provides bibliographic and regional metadata with short Japanese summaries (International Research Center for Japanese Studies, 2025). A record may identify only a prefecture, mention one or more candidate toponyms, or describe a riverbank, route, dwelling, grave, shrine, livelihood, or boundary without naming a resolvable place. These signals differ in kind and resolution.

Historical geocoding compounds this problem because present administrative boundaries do not necessarily represent the places named in older sources. A representative point can be useful as a display anchor, but it is not evidence that the narrated event occurred there. We therefore treat resolution, support area, matching method, ambiguity, and provenance as explicit fields rather than as a calibrated probability.

This work represents vague folklore locations through separate evidence channels for administrative support areas, candidate toponyms, non-toponymic place descriptions, geographic context, human-condition terms, interface evidence, and support provenance. It shows how geographic summaries change when administrative support is refined, without treating representative coordinates as exact event locations. Three questions guide the study:

RQ1 (Representation): What evidence channels can be derived from vague folklore records while preserving the geographic support warranted by each record?

RQ2 (Resolution sensitivity): For records with uniquely matched municipality support, how do support area and coordinate-derived geographic features change relative to the prefecture-level baseline?

RQ3 (Archive-language association): Which category–place-function associations remain after controlling for prefectural composition?

The contributions are (i) an evidence-channel representation that separates textual and geographic support; (ii) an application to 33,378 archive records; (iii) a rule-constrained municipality-support subset; (iv) polygon-based resolution-sensitivity diagnostics; and (v) prefecture-stratified archive-language association analyses with corrected inference and robustness checks.

2. Related Work

2.1 Folklore, Spatial Humanities, and Archive Bias

Yokai categories are historically produced and transmitted rather than fixed biological taxa (Foster, 2024). Spatial humanities likewise combines distant views with source criticism and close reading, while warning that mapped corpora inherit the coverage and classification decisions of their archives (Gregory and Geddes, 2014). Accordingly, record density in the present database is not interpreted as population prevalence or as a complete geography of Japanese folklore.

2.2 Toponyms and Non-Toponymic Spatial Language

Geoparsers generally detect toponyms and link them to gazetteer entries, as exemplified by HeidelPlace (Richter et al., 2017). Japanese document-level work has developed richly annotated geographic-entity resources (Higashiyama et al., 2024). However, locations can also be expressed through relations and common-noun descriptions that are not reducible to a named point (Stock et al., 2022). Our representation retains both candidate toponyms and non-toponymic place descriptions; neither channel is treated as a validated extractor in the absence of a corpus-specific gold standard.

2.3 Historical Geocoding and Region-Based Support

Historical geocoding must accommodate temporal, textual, positional, and source uncertainty (Cura et al., 2018). Gazetteer-linked annotation systems also preserve editorial intervention and source links (Simon et al., 2017). The present study takes a more limited step: current administrative polygons provide support areas for records, and interior representative points provide display anchors. A matched current municipality is therefore a resolution refinement, not historical ground truth.

2.4 Visualizing Evidential Limits

Geospatial uncertainty research distinguishes sources of uncertainty and their visual encodings (MacEachren et al., 2005). Digital-humanities visualizations can appear more objective than the underlying evidence and often omit uncertainty introduced during datafication (Panagiotidou et al., 2023). Because the present pipeline has no calibrated probability, we describe it as "resolution-aware" and expose support basis, geometry, and ambiguity directly.

3. Data

The analysis uses a local export of 35,305 database rows. Cleaning retains 33,378 and excludes 1,927: 141 summaries shorter than 10 characters and 1,786 records whose region is outside the Japanese prefecture set. A record is one database row, not a unique narrative event. Repeated narratives were not deduplicated because a defensible source-level equivalence rule is not available; their dependence is a limitation. Empty evidence channels are retained as missing rather than converted to negative evidence.

Prefecture names are normalised to the 47 contemporary Japanese prefectures. Category labels are analysed as supplied in the local export; the residual Other label is not assigned a new substantive interpretation. The public database page reports 35,307 items as of April 2025 and notes that historical records can contain discriminatory language (International Research Center for Japanese Studies, 2025); that count is not substituted for the local export count.

The source database's redistribution terms were not located in the supplied materials. This article therefore reports aggregate results without redistributing raw summaries. Current GADM boundaries are applied to potentially historical records only as contemporary support proxies. The analysis neither reconstructs historical boundaries nor dates individual geometries.

4. Representation and Methods

Figure 1 shows the distinction on which the representation is based. The Folklore Geospatial Support Model is a representation schema, not a predictive model. For record i, it is

$$S_i = (A_i, T_i, P_i, G_i, H_i, I_i, B_i). \tag{1}$$

where A_i is administrative support, T_i candidate-toponym evidence, P_i non-toponymic place-description evidence, G_i coordinate-derived geographic context, H_i human-condition evidence, I_i derived interface evidence, and B_i the support basis and provenance. An anchor a_i is stored for display but is not part of the evidential support claim. Table 1 specifies the source, rule, output resolution, ambiguity, provenance, and missing-data handling of each channel.

4.1 Textual Evidence

A Japanese dependency-based named-entity recogniser identifies candidate place entities in each summary (Matsuda, 2020). Fixed lexical rules independently identify place-function and human-condition surface forms. Place functions are hydrology, mobility, boundary, livelihood, dwelling, death ritual, and taboo/time/weather; textual landscape mentions separately record water, mountain, coast, boundary, and agrarian cues. These are multi-label textual features, not observed terrain.

Human-condition evidence is the union of actor, action, and taboo/time/weather terms. The broad human–environment interface is explicitly

$$\begin{aligned} E_i &= \text{hydrology} \vee \text{livelihood} \vee \text{water} \\ &\quad \vee \text{mountain} \vee \text{coast}, \\ H_i &= \text{actor} \vee \text{action} \\ &\quad \vee \text{taboo/time/weather}, \\ \text{broad interface}_i &= E_i \wedge H_i. \end{aligned} \tag{2}$$

This definition makes every broad-interface record a subset of the human-condition records. Strict interfaces are named conjunctions: water crossing, waterside boundary, water-danger norm, mountain route, mountain boundary, coastal landing, domestic threshold, mortuary boundary, and water livelihood.

No human gold labels are available by design for this revision. We therefore report rule-based coverage, never precision, recall, F_1, or validated extraction. Candidate toponym extraction, place-description extraction, human-condition extraction, interface derivation, and municipality matching are represented separately so that a future annotation study can evaluate them without changing the schema.

4.2 Rule-Constrained Municipality Support

GADM 4.1 level-2 units are filtered to municipality-like types Shi, Machi, Mura, SpecialWard, Capital, and Son (GADM, 2022). The Japanese name field is `NL_NAME_2`; `NAME_2` supplies English context. Names and summaries are whitespace-normalised. A direct summary match uses an official-name substring with a left-boundary constraint inside the record's prefecture. If that stage does not assign a unit, an exact candidate-toponym match is attempted. A record is refined only when the deduplicated candidate set contains one unit; multiple or same-named candidates remain at prefecture support. Direct summary matching has precedence, making the 329 direct matches and 902 entity-based matches mutually exclusive.

This procedure assigns municipality support to 1,231 records across 452 unique units. Accepted record types are 929 Shi, 231 Machi, 53 Mura, and 18 SpecialWard; no accepted records use the other allowed types. Historical names are not resolved. GADM polygons are projected to EPSG:3095, and Shapely's interior representative point is transformed to WGS84 for display. The full polygon remains the evidential support.

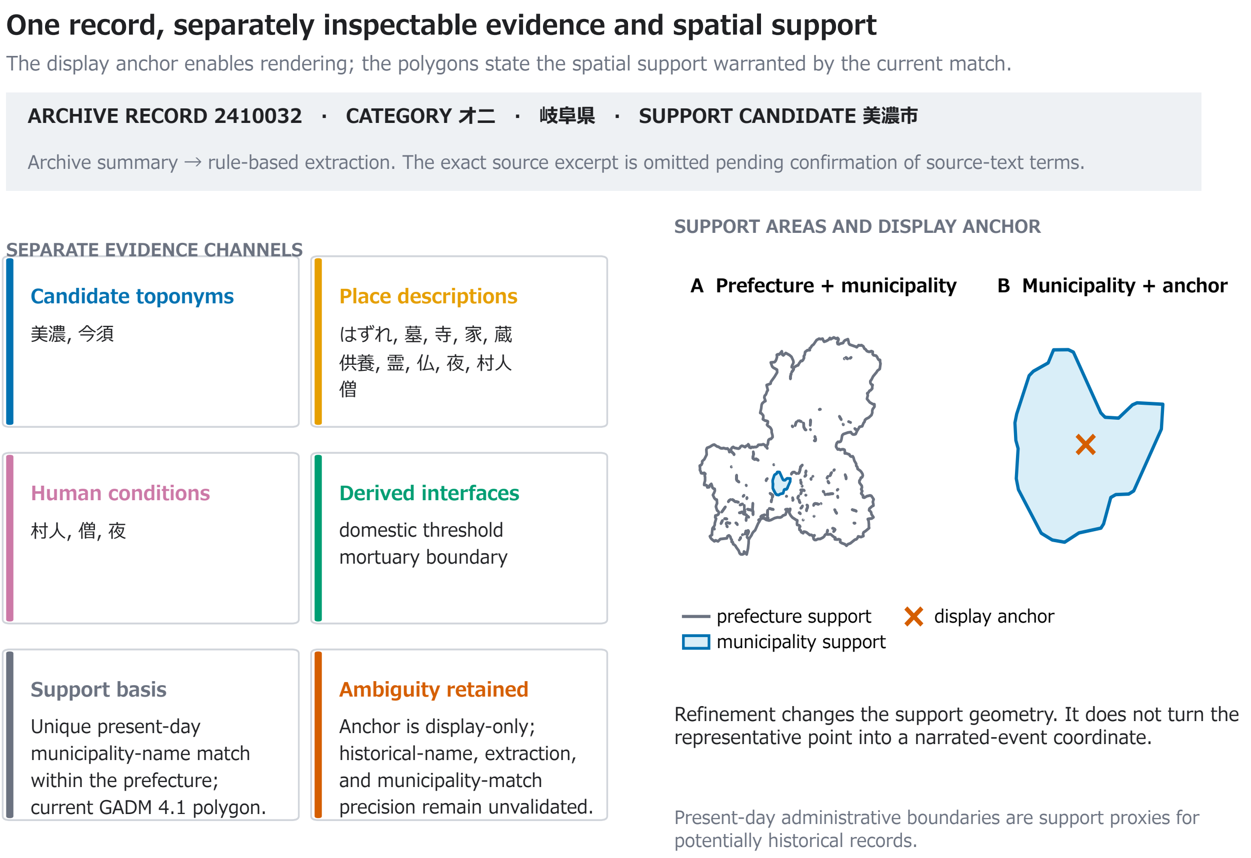

Figure 1. Crown-jewel view of one record. Archive metadata and extracted summary evidence are separated into candidate-toponym, place-description, human-condition, interface, support-basis, and ambiguity channels; the exact source excerpt is omitted pending confirmation of source-text terms. The maps distinguish the prefecture support, the rule-matched municipality support, and the interior display anchor. The anchor is not asserted to be the narrated event location.

4.3 Polygon-Based Geographic Context

Support metrics use GADM 4.1 polygons, Natural Earth 1:10m coastlines and lakes, and National Land Numerical Information W05 river centre-lines (GADM, 2022, Natural Earth, 2026, Ministry of Land, Infrastructure, Transport and Tourism, 2009). Distances and areas are computed in EPSG:3095. "Major water" is the union of the supplied lake boundaries and W05 river lines; river metrics use W05 lines only.

For support polygon P_i and feature geometry F, the minimum distance is

$$d_{\min}(P_i, F) = \min_{x \in P_i, y \in F} \|x - y\|. \tag{3}$$

We additionally draw 128 deterministic area-uniform points x_{ij} per unique support and report the sampled median distance and the share meeting each threshold. Outputs include polygon area; minimum and sampled-median distance to coast, rivers, and major water; proportions within 0.5 km of major water, 2 km of a river, and 10 km of coastline; river and coastline density; and lake-area share. Sensitivity grids use coast thresholds 5, 10, and 20 km; river thresholds 1, 2, and 5 km; and water thresholds 0.25, 0.5, and 1 km. These thresholds are diagnostics rather than natural classes. No exclusive terrain classification or classification precedence is used.

4.4 Archive-Language Associations

RQ3 concerns semantic metadata within the archive, not physical geography. For each selected category–feature cell we report observed prevalence, the median null prevalence, observed-to-null ratio, prevalence difference, an odds ratio with a 95% confidence interval, a two-sided empirical permutation value, and a Benjamini–Hochberg adjusted q-value (Benjamini and Hochberg, 1995). The null shuffles category labels within prefectures 10,000 times, thereby preserving prefectural composition and category size. The empirical tail probability uses $(b + 1)/(B + 1)$. The screen contains 120 category–feature cells. Source-document blocking was not used because 50 joined records have missing or malformed source-group metadata; this remaining dependence limits inference.

Kappa–water and Yurei–death are prespecified construct checks: recovering known archive-language structures shows that the representation can encode them, not that the study discovered them. Snake/Dragon–strict boundary, Kappa–livelihood, and post-hoc mobility, dwelling, and taboo screens are exploratory. Surface-form leakage is reduced by masking category names, but semantic dependence is not claimed to be removed. Dictionary deletion is repeated 100 times at each of 10%, 20%, and 30%.

Channel	Source	Rule and output	Resolution	Ambiguity, provenance, and missing data
A_i administrative support	Prefecture field; GADM 4.1	Prefecture polygon; unique municipality polygon when matched	Prefecture or current municipality-like unit	Current-boundary source and matching method retained; prefecture fallback on no unique match
T_i candidate toponym	Japanese summary	Named-entity candidates; optional exact municipality candidate	Mention and candidate unit	All candidates retained; no calibrated confidence; empty if none detected
P_i place description	Summary and fixed lexical rules	Multi-label hydrology, mobility, boundary, livelihood, dwelling, death-ritual, taboo/time/weather terms	Textual, not geographic	Matched surface forms retained; empty means no rule match
G_i geographic context	Support polygon, coastline, rivers, lakes	Continuous distances, densities, shares, and threshold proportions	Entire support polygon	Data source and scale identified; finite polygon sampling remains approximate
H_i human condition	Summary and dictionaries	Actor, action, or taboo/time/weather terms	Textual, not geographic	Matched terms retained; no inference about actual participants
I_i interface	Other evidence channels	Deterministic conjunctions such as hydrology with mobility	Record-level semantic metadata	Derivation path retained; not a physical site observation
B_i support basis	Administrative and textual evidence	Source, matching basis, resolution, and candidate count	Record	Missing values remain explicit; calibrated confidence is not reported

Table 1. Evidence-channel schema. Missing channels remain empty and do not erase evidence in other channels. "Rule constrained" describes matching restrictions, not measured precision.

Evidence channel	Records	Share
Candidate toponym	9,361	28.0%
Non-toponymic place description	25,023	75.0%
Human condition	16,066	48.1%
Broad human–environment interface	9,518	28.5%
Strict derived interface	9,047	27.1%

Table 2. Coverage of separately represented evidence channels. Counts are rule-based coverage, not extraction accuracy; channels are non-exclusive. The denominator is 33,378 records.

5. Results

5.1 Evidence-Channel Coverage

Figure 2 and Table 2 report non-exclusive, rule-based coverage. Candidate toponyms occur in 9,361 records (28.0%), while non-toponymic place descriptions occur in 25,023 (75.0%). Human-condition evidence occurs in 16,066 (48.1%). The broad interface contains 9,518 records (28.5%) and is a verified subset of the human-condition set.

5.2 Resolution Sensitivity of Polygon Support

Figure 3 and Table 3 compare the 452 uniquely matched municipality polygons with their prefecture polygons. Municipality support reduces polygon area by a median 97.5% (interquartile range 94.6–98.9%). The median absolute change in sampled median coastline distance is 9.2 km (4.2–17.0 km); for major water it is 0.1 km (0.0–0.1 km after resolution-aware rounding). These are changes under administrative-support refinement, not corrected coordinates or errors against event locations.

5.3 Archive-Language Associations

Figure 4 and Table 4 replace the saturated standardised-residual heatmap with interpretable effect estimates. Kappa–water has

Metric	Median	IQR
Support-area reduction (%)	97.5	94.6–98.9
Coast-distance change (km)	9.2	4.2–17.0
Major-water-distance change (km)	0.1	0.0–0.1

Table 3. Resolution sensitivity for 452 uniquely matched municipality support polygons relative to their prefecture support polygons. Values are medians and interquartile ranges. Distances summarise 128 deterministic area-uniform sample points per support and are not errors against event-location ground truth.

78.0% observed versus 23.1% median null prevalence (odds ratio 15.72, 95% CI 13.82–17.89, $q = 0.0002$); Yurei–death has 75.1% versus 16.1% (odds ratio 23.01, 20.47–25.87, $q = 0.0002$). These strong values are construct checks for archive wording.

Among the exploratory cells, Snake/Dragon–strict boundary has an observed-to-null ratio of 1.34 and odds ratio 1.49 (1.34–1.66, $q = 0.0002$); Kappa–livelihood has a ratio of 1.47 and odds ratio 1.69 (1.50–1.91). The Oni, Inugami, and Hitodama rows are post-hoc screens, not independent confirmation. These exploratory cells are treated as prompts for close reading rather than as independently confirmed findings.

5.4 Kappa Water-Distance Diagnostic

The one-sided alternative is that Kappa-labelled municipality-supported records have a smaller median representative-point distance to mapped major water than expected after within-prefecture category shuffling. Among 1,231 municipality-supported records, 48 are Kappa. The observed median is 0.371 km, while the median of the 10,000-permutation null distribu-

Category	Feature	Obs. %	Null %	O/E	OR	95% CI	q
Kappa	Water setting	78.0	23.1	3.38	15.72	[13.82, 17.89]	0.0002
Yurei	Death ritual	75.1	16.1	4.65	23.01	[20.47, 25.87]	0.0002
Snake/Dragon	Strict interface	18.4	13.7	1.34	1.49	[1.34, 1.66]	0.0002
Kappa	Livelihood	28.2	19.2	1.47	1.69	[1.50, 1.91]	0.0002
Oni	Mobility	24.2	16.7	1.44	1.60	[1.36, 1.88]	0.0002
Inugami	Dwelling	36.6	24.3	1.50	1.77	[1.49, 2.11]	0.0002
Hitodama	Taboo, time, or weather	32.1	21.3	1.51	1.74	[1.33, 2.28]	0.0002

Table 4. Prefecture-stratified archive-language associations. Null is the median prevalence under 10,000 within-prefecture label permutations. The q value is Benjamini–Hochberg adjusted across 120 screened cells. Kappa–water and Yurei–death are construct checks; other rows are exploratory.

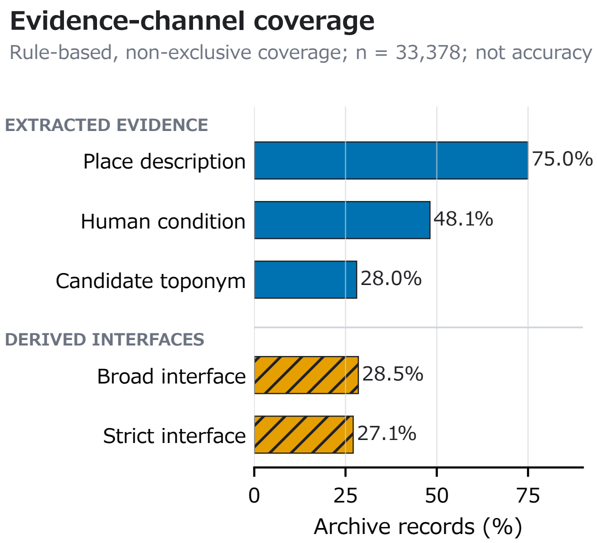

Figure 2. Rule-based coverage of non-exclusive evidence channels among 33,378 records. Bars report coverage, not extraction accuracy. The broad human–environment interface is defined as environmental evidence conjoined with actor, action, or taboo/time/weather evidence and is therefore a subset of the human-condition channel.

tion is 0.283 km with a 95% null interval of 0.218–0.377 km. The one-sided p-value for a smaller median is 0.9323 and the two-sided value is 0.0892; all 10,000 permutations are valid. Because the observed median exceeds the null median, the physical-proximity hypothesis is not supported. Repeated municipalities, rule-matched units, and use of representative points make this a limited negative diagnostic.

5.5 Robustness Diagnostics

Table 5 shows that dictionary deletion gradually changes the full 120-cell ordering: median Spearman correlation is 0.975 at 10% deletion, 0.940 at 20%, and 0.895 at 30%. Category-name masking removes 22,974 occurrences in 13,140 records and reduces surface-form leakage; it changes the Kappa–water observed-to-null ratio from 3.38 to 2.73 and Yurei–death from 4.65 to 3.57. It does not establish semantic independence. For the Snake/Dragon strict-boundary cell, the positive direction is retained in 100%, 99%, and 98% of deletion replicates at the three rates; Kappa–livelihood retains it in 87%, 85%, and 70%.

Table 6 demonstrates definition sensitivity rather than choosing a uniquely correct boundary construct.

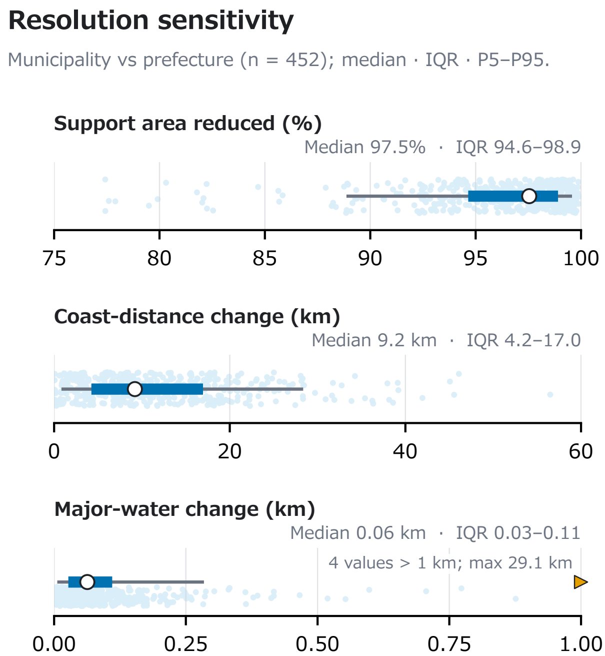

Figure 3. Resolution sensitivity for 452 municipality support polygons relative to their prefecture supports. Pale points show support units; white points are medians, thick intervals are interquartile ranges, and thin intervals span the 5th–95th percentiles. The major-water panel focuses on 0–1 km and explicitly marks values beyond that range. Metrics use whole polygons and 128 deterministic area-uniform samples per support; no representative point is treated as ground truth.

6. Discussion

RQ1. The archive can be represented without collapsing all spatial evidence into one coordinate. Administrative polygons, candidate toponyms, non-toponymic place descriptions, geographic context, human-condition terms, derived interfaces, and provenance remain inspectable. The present analysis establishes representational coverage, not extraction accuracy.

RQ2. Municipality support produces much smaller evidential areas and materially different coastline summaries for the matched subset. The small median change in major-water distance reflects the density of the mapped river network and the sampled metric; it does not validate either administrative level as an event location. Polygon distributions are more faithful to the evidence than interpreting one display anchor, but present-day boundaries still remain proxies.

RQ3. Prefecture-stratified permutations recover known

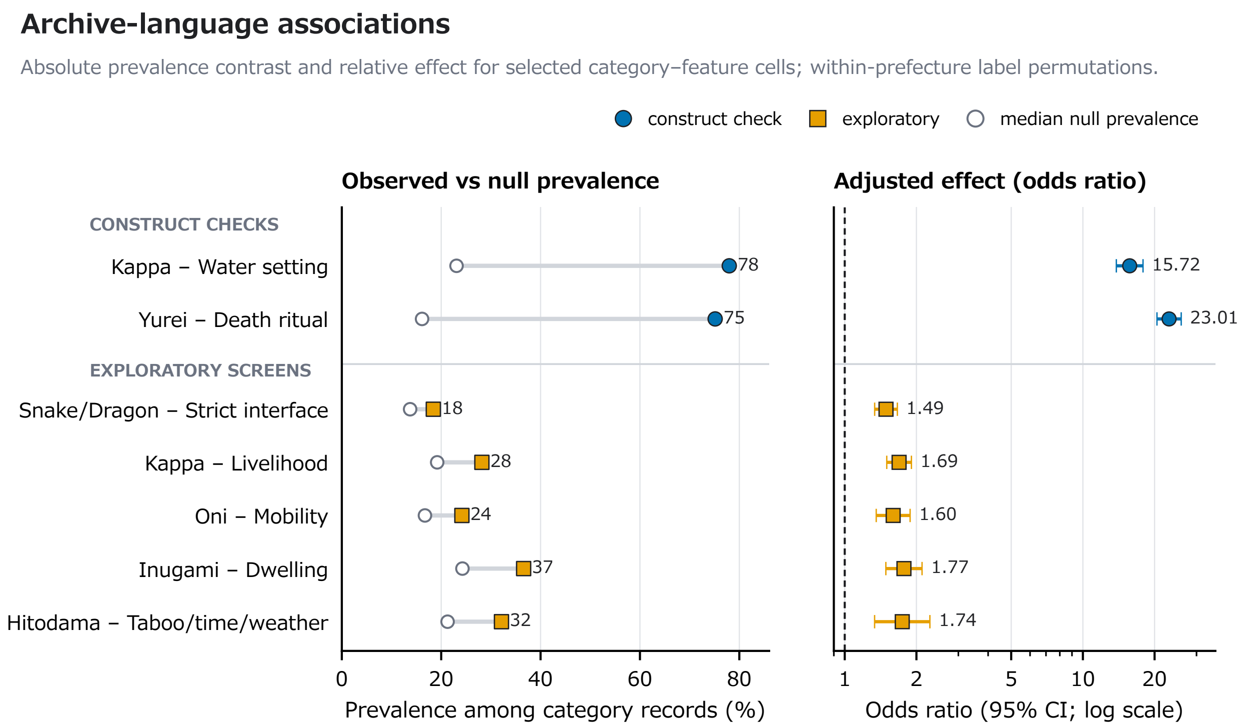

Figure 4. Prefecture-stratified archive-language associations. The left panel compares observed prevalence with median prevalence under within-prefecture label permutations; the right panel shows odds ratios and 95% confidence intervals on a log scale, with the dashed line denoting no association. Kappa–water and Yurei–death are construct checks. Other rows are exploratory and describe archive-language co-occurrence, not physical ecological relationships.

Deletion	Median ρ	IQR	Min–max
10%	0.975	0.942–0.987	0.878–0.999
20%	0.940	0.916–0.963	0.846–0.994
30%	0.895	0.852–0.925	0.665–0.979

Table 5. Dictionary-deletion robustness over 100 deterministic replicates per rate. Spearman correlation compares the full category–feature effect-ratio ordering.

Boundary-interface definition	O/E ratio
Broad environmental × boundary	1.68
Strict derived boundary interfaces	1.34
Hydrology/mountain boundary interfaces	1.83

Table 6. Interface-definition ablation for the exploratory Snake/Dragon–boundary association. Construct-check rows are not repeated because these definitions do not alter their features.

Kappa–water and Yurei–death archive-language associations and identify exploratory boundary, livelihood, mobility, dwelling, and taboo cells. The construct checks show that the schema captures obvious language structures; they are not new folklore findings. The exploratory cells require close reading and independent extraction validation before interpretation.

The distinction between textual and geographic evidence is decisive. A water term in a summary is textual landscape evidence; distance from a support polygon to a mapped river is geographic proximity. Conjoining them can produce a derived semantic interface, but it does not observe a local ecology. Likewise, dense archive coverage may reflect source collection, indexing, or repetition rather than regional folklore prevalence. Maps can over-objectify these choices, which is why the interface must show resolution, support polygon, display anchor, matching method, and source trace together.

Transfer to another archive requires a documented record unit, stable source identifiers, local-language place and function dictionaries, an appropriate historical or contemporary gazetteer, explicit redistribution terms, and a new human evaluation if extraction performance is to be claimed.

7. Limitations

The rule dictionaries can miss obsolete names, shrine names, spelling variants, and implicit expressions, and can mistake common nouns or beings for places. Current administrative names cannot resolve many historical toponyms.

Records sharing a prefecture, municipality, source document, or repeated narrative are dependent. Prefecture-stratified permutations address only the first dependence; source-group blocking was not used because part of the local source metadata is malformed. The polygon sample is deterministic but finite, and W05 river vintages differ by region. Natural Earth 1:10m is appropriate for national context, not detailed local hydrology. No digital elevation model, slope, or terrain-position index is used.

The municipality subset is rule-matched rather than manually verified, and the representative-point distance diagnostic repeats shared municipal anchors. The analysis therefore does not establish exact event locations, coordinate accuracy, physical ecological relationships, diffusion from Kyoto, or population-level prevalence.

8. Ethical Considerations

This computational study involved no human participants. It uses historical source descriptions that may contain discriminatory language, as the database provider itself cautions (International Research Center for Japanese Studies, 2025). Public displays should avoid unnecessary exposure of personal names, sacred sites, or small communities and should not turn uncertain narratives into precise-point claims or regional stereotypes.

GADM 4.1 permits academic and other non-commercial use but restricts redistribution and commercial use without permission (GADM, 2022). Natural Earth data are public domain (Natural Earth, 2026). The W05 river page identifies non-commercial conditions (Ministry of Land, Infrastructure, Transport and Tourism, 2009). This article reports derived summaries and does not reproduce source geometries or archive excerpts.

Current boundaries are labelled as present-day support proxies when applied to historical records. Archive density is not presented as folklore prevalence.

9. Conclusion

We represented 33,378 folklore records through separate administrative, toponymic, non-toponymic, geographic, human-condition, interface, and provenance channels. For RQ1, this schema preserves what kind of evidence supports each spatial interpretation, although its extractors remain unevaluated. For RQ2, municipality polygons substantially reduce support area and change polygon-derived coastline summaries relative to prefecture support; these changes measure resolution sensitivity, not accuracy. For RQ3, prefecture-stratified analysis recovers expected archive-language constructs and identifies exploratory cells for close reading, but it does not establish physical ecological relationships.

The next necessary step is a preregistered human evaluation of candidate toponyms, place descriptions, human-condition terms, interface derivations, and municipality matches, followed by source-aware models using verified historical place resolution. Until then, the defensible contribution is a failure-aware baseline that distinguishes display anchors from the support areas warranted by the available evidence.

Acknowledgements

This work was supported by The Nippon Foundation HUMAI Program.

References

Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Cura, R., Dumenieu, B., Abadie, N., Costes, B., Perret, J., Gribaudi, M., 2018. Historical Collaborative Geocoding. *ISPRS International Journal of Geo-Information*, 7(7), 262. doi.org/10.3390/ijgi7070262.

Foster, M. D., 2024. *The Book of Yōkai: Mysterious Creatures of Japanese Folklore*. Expanded 2nd edn, University of California Press, Oakland, California.

GADM, 2022. GADM database of global administrative areas, version 4.1. Geospatial dataset; academic and non-commercial use; redistribution requires prior permission. gadm.org/license.html (31 July 2026).

Gregory, I. N., Geddes, A. (eds), 2014. *Toward Spatial Humanities: Historical GIS and Spatial History*. Indiana University Press, Bloomington.

Higashiyama, S., Ouchi, H., Teranishi, H., Otomo, H., Ide, Y., Yamamoto, A., Shindo, H., Matsuda, Y., Wakamiya, S., Inoue, N., Yamada, I., Watanabe, T., 2024. Arukikata travelogue dataset with geographic entity mention, coreference, and link annotation. *Findings of the Association for Computational Linguistics: EACL 2024*, Association for Computational Linguistics, St. Julian's, Malta, 513–532. doi.org/10.18653/v1/2024.findings-eacl.35.

International Research Center for Japanese Studies, 2025. The Database of Folktales of Mysterious Phenomena and Yōkai (spirits, ghosts, and monsters). Online database; database count updated April 2025. www.nichibun.ac.jp/en/db/category/yokai/ (31 July 2026).

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., 2005. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartography and Geographic Information Science*, 32(3), 139–160. doi.org/10.1559/1523040054738936.

Matsuda, H., 2020. GiNZA: Practical Japanese NLP Based on Universal Dependencies. *Journal of Natural Language Processing*, 27(3), 695–701. doi.org/10.5715/jnlp.27.695.

Ministry of Land, Infrastructure, Transport and Tourism, 2009. National Land Numerical Information: River data (W05). Geospatial dataset, regional source years 2006–2009; JGD2000; non-commercial use. nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-W05.html (31 July 2026).

Natural Earth, 2026. Natural Earth vector data at 1:10m scale. Public-domain geospatial dataset. www.naturalearthdata.com/about/terms-of-use/ (31 July 2026).

Panagiotidou, G., Lamqaddam, H., Poblome, J., Brosens, K., Verbert, K., Vande Moere, A., 2023. Communicating Uncertainty in Digital Humanities Visualization Research. *IEEE Transactions on Visualization and Computer Graphics*, 29(1), 635–645. doi.org/10.1109/TVCG.2022.3209436.

Richter, L., Geiß, J., Spitz, A., Gertz, M., 2017. HeidelPlace: An extensible framework for geoparsing. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Association for Computational Linguistics, Copenhagen, Denmark, 85–90. doi.org/10.18653/v1/D17-2015.

Simon, R., Barker, E., Isaksen, L., de Soto Cañamares, P., 2017. Linked Data Annotation Without the Pointy Brackets: Introducing Recogito 2. *Journal of Map & Geography Libraries*, 13(1), 111–132. doi.org/10.1080/15420353.2017.1307303.

Stock, K., Jones, C. B., Russell, S., Radke, M., Das, P., Aflaki, N., 2022. Detecting Geospatial Location Descriptions in Natural Language Text. *International Journal of Geographical Information Science*, 36(3), 547–584. doi.org/10.1080/13658816.2021.1987441.