The Social Morphology Ball Model: A Proof-of-Concept for Spherical Representation of Multivariate Social Data

Ryu Murakami[1,2]

[1]Chiba Institute of Science, Faculty of Risk and Crisis Management. Chiba, JAPAN

[2]Meiji University of Integrative Medicine, School of Health Science and Medical Care, Faculty of Emergency Medical Science, Kyoto, JAPAN

e-mail: ryu-murakami@umin.ac.jp

This proof-of-concept study introduces the Social Morphology Ball, a framework for visualizing multivariate social data as three-dimensional morphological structures. The method normalizes variables using the median and interquartile range (IQR), followed by nonlinear compression with a hyperbolic tangent function to reduce the impact of extreme outliers. Each normalized variable is mapped onto spherical basis functions, transforming numerical relationships into geometric deformations—bulges or concavities—on a sphere. When all variables align with their long-term medians, the model yields a perfect sphere representing a morphologically neutral, structurally balanced state. Implemented in Python, the pipeline automates normalization, shape generation, and visualization, exporting results in STL, PNG, and GIF formats. Applied to Japanese national statistics from 2015–2023, the model captured temporal variations in the balance and distortion of societal indicators through intuitive 3D forms. By interpreting data as morphology rather than numbers, this approach extends social data visualization into a sensory and cognitive domain, providing a conceptual foundation for morphological visualization of social phenomena and potential applications in education, policy, and data-driven art.

Key words： social morphology; data physicalization; geometric representation; robust normalization; spherical basis functions; social systems modeling

1. Background

Multivariate datasets that simultaneously handle heterogeneous variables—such as medical, social, statistical, and public health data—are indispensable for understanding the underlying mechanisms, structures, and phenomena of complex systems.[1-3] Visualizing numerical data provides researchers with valuable insights across a wide range of academic disciplines. However, such information is typically presented in two-dimensional formats—such as tables or graphs—which often makes it difficult to intuitively grasp the interrelationships and overarching structures among variables. In the social sciences, in particular, datasets that span multiple domains—demography, economics, health, environment, welfare, and social institutions—are frequently reduced to single-index evaluations that fail to capture essential variations across dimensions. To enable simultaneous comparison and comprehensive understanding across these domains, more integrative visualization methods are needed to generate new perspectives.The author hypothesized that treating data as morphological entities—that is, as structured forms—could facilitate a more intuitive understanding of multidimensional relationships within social structures. Here, the term "morphological" refers to the structural or geometric forms of multivariate data, rather than to its biological or linguistic usage.This study focuses on a spherical "Ball" model as the core framework, though the principle can be extended to other geometrical forms such as polyhedra, tori, or multilayered spheres. The proposed approach represents an attempt to express interdependencies and cyclicity among data in a tangible, visual form. Beyond serving as a methodological tool for both practical and academic interpretation of social statistics, it may also function as an interactive, art-inspired medium for conveying the "pulsation" of society to the general public, including audiences outside specialized fields.

The artist Mathieu Lehanneur has created sculptural artworks that materialize demographic data into three-dimensional forms, exemplifying how quantitative information can be transformed into tangible, spatial representations.[4] In this proof of concept, we sought to transcend the planar and static limitations of conventional visualization techniques and to represent the overall structure of society as a sculptural, three-dimensional form. Specifically, multiple social indicators were transformed into three-dimensional geometries and rendered as spheres to depict societal equilibrium and distortion in a geometric manner. This approach is grounded in the concept of mathematical morphology, reconstructing numerical datasets into physical and visual forms. The objective of this proof-of-concept study is to present a methodological framework that reconstructs social data—as a collection of numerical values—into geometric forms, thereby capturing equilibrium and variation as geometric features. While previous studies in data visualization have mainly focused on two-dimensional representations (e.g., maps, scatterplots, and network diagrams), few have explored the possibility of morphologically reconstructing social data into tangible, three-dimensional forms. The present study introduces the concept of the "*social morphology ball*", a framework that treats multivariate social indicators as a unified geometric entity, thereby extending the idea of data visualization into the realm of morphological cognition. This conceptual shift—from observing data to feeling societal structure—constitutes the novel contribution of this work.

2. Methods

2.1 Data used for implementation testing

To ensure the reproducibility of the proposed method, both the datasets and the Python scripts used for analysis are made publicly available as supplementary materials. The statistical data were prepared in CSV format, in which each row corresponds to a year and each column (after the first) represents a social indicator (variable). The first column stores the year. The data were entered using Microsoft Excel (Microsoft Corporation). For the purpose of this preliminary implementation test, all datasets were obtained from the statistical database managed by the Ministry of Internal Affairs and Communications of Japan.[5] The dataset consisted of annual data from 2015 to 2023, which were manually retrieved by the author from the official database and compiled into a CSV file. The CSV file used in this analysis comprised the following variables:

- Year, Total population (persons), Number of births (persons), Number of deaths (persons), Japanese population (persons), Number of general hospitals (facilities), and Number of new car registrations (units).

2.2 Code generation

At the initial stage of this study, ChatGPT-5 (OpenAI) was utilized to generate a Python script for the purpose of conceptual validation. The generated code was subsequently reviewed, refined, and executed by the author within the JupyterLab environment (Version 4.4.3; Project Jupyter Contributors) to verify its reproducibility and to confirm the successful generation of graphical outputs.

2.3 Data input and robust normalization

Let $t = 1,\ldots,T$ index years (or regions) and $j = 1,\ldots,p$ index variables. *Denote* $x_{t,j}$ as the observed value of variable j at time t. For each variable j, the median and interquartile range (IQR) are computed over the entire observation window. Missing values are handled as stated in the Data Processing subsection. A small constant $\varepsilon > 0$ prevents division by zero when IQR = 0.

Equation (1) defines the robust normalization; equation (2) applies bounded nonlinear compression.

$$z_{t,j} = \left(x_{t,j} - \mathrm{median}\left(x_{*,j}\right)\right) / \left(\mathrm{IQR}\left(x_{*,j}\right) + \varepsilon\right) \tag{1}$$

$$\tilde{z}_{t,j} = \tanh\left(\alpha z_{t,j}\right),\ \alpha = 0.5 \tag{2}$$

In practice, a uniform CSV in which every entry is identical (e.g., 0.5 across all years and variables) will produce perfect spheres for every year and a blank-looking heatmap. This happens because the preprocessing centers each variable at its long-term median and rescales by its IQR; when all values are identical, the centered values become zero and no deformation remains.

2.4 Basis, coefficients, shape field, and change detection

2.4.1 Basis set

This proof-of-concept uses eight spherical trigonometric basis functions: $\{ 1, \sin\theta, \cos\theta, \sin\phi, \cos\phi, \sin 2\theta, \cos 2\theta, \sin\theta\cos\phi \}$. The constant basis is constrained to zero to avoid redundant global scaling.

2.4.2 Coefficients and variable-to-basis mapping

Equations (3)–(4) define the coefficients used to deform the radius field; (4′) is the general form with a fixed weight matrix. If $p < 8$, remaining coefficients are padded with zero.

$$a_0(t) = 0 \tag{3}$$

$$a_k(t) = s\,\tilde{z}_{t,k} \ (k = 1, \dots, 7) \tag{4}$$

$$a_k(t) = \sum_{j=1}^{p} w_{k,j}\ \tilde{z}_{t,j} \tag{4′}$$

Here $s > 0$ denotes the global deformation scale (`CLI option --scale, default 0.30`), while the reference radius is controlled by `--radius.`

In this proof-of-concept the first eight columns of the input CSV map one-to-one to $\Phi_1 \dots \Phi_7$ (the constant term Φ_0 is constrained to zero). If $\mathbf{p < 8}$, the remaining coefficients are zero-padded. The exact column-to-basis correspondence table is provided in the Supplement.

2.4.3 Radius (shape) field and clipping

Equations (5)–(6) define the radius field on the sphere and the clipping operator.

$$R(\theta, \phi, t) = R_0\, \mathrm{clip}\left(1 + \sum_{k=0}^{7} a_k(t)\ \Phi_k(\theta, \phi),\ 0.55,\ 1.60\right) \tag{5}$$

$$\mathrm{clip}(u, L, U) = \max\left(L, \min(u, U)\right) \tag{6}$$

2.4.4 Year-to-year change magnitude and CUSUM

Equations (7)–(9) define the year-to-year magnitude, its mean, and the CUSUM used for change-point inspection.

$$\Delta_t = \sqrt{\sum_{j=1}^{p} \left(\tilde{z}_{t,j} - \tilde{z}_{t-1,j}\right)^2} \tag{7}$$

$$\bar{\Delta} = \frac{1}{T-1} \sum_{t=2}^{T} \Delta_t \tag{8}$$

$$\mathrm{CUSUM}(u) = \sum_{t=2}^{u} \left(\Delta_t - \bar{\Delta} \right)\ (u = 2, \dots, T) \tag{9}$$

2.5 Implementation environment

The proposed visualization pipeline was implemented in Python 3.13.5 under the following computational environment:

- Main libraries: NumPy, Pandas, Matplotlib, *mpl_toolkits.mplot3d*, and Pillow (for GIF animation);
- Execution environment: JupyterLab (Version 4.4.3; Project Jupyter Contributors);
- Script: *Soc_ball_pipeline.py*;
- Output formats: PNG (heatmaps and coefficient plots), STL (3D-printable polygon meshes), and GIF (morphing animations of annual transitions).

The script accepts any CSV file containing annual or regional indicators as input and automatically performs the following steps:

- robust normalization (median-centered and IQR-scaled);
- nonlinear data compression using the hyperbolic-tangent function;
- spherical-shape generation via basis-function expansion; and
- visualization and export of results.

The pipeline also generates auxiliary outputs—such as turning-point analyses (Δ-magnitude and CUSUM-drift plots)—to quantify year-to-year variations in the geometric representation of social indicators.

- 3D Models (ball_YYYY.stl);
- Morphing Animation (balls_morph_gradient.gif);
- Heatmap (neutral_heatmap.png);
- Preview Grid (preview_grid.png);
- Coefficient Plot (basis_coefficients.png); and
- Change-Point Analysis (turningpoint_*.png) turningpoint_delta.png, turningpoint_cusum.png.

3. Results

First, the visualization results obtained when all social indicators were uniformly set to 0.5 are presented (Supplement; Test_05.csv). Under this uniform condition, the model generated a perfectly spherical structure, and the heatmap showed no variation across variables or years. This configuration represents a morphologically neutral baseline, indicating a completely balanced state without directional bias or deformation (Figures 1 and 2). Under a uniform CSV (all entries = 0.5), preprocessing centers every variable at its median (IQR = 0), so no deformation occurs; spheres remain perfectly neutral.

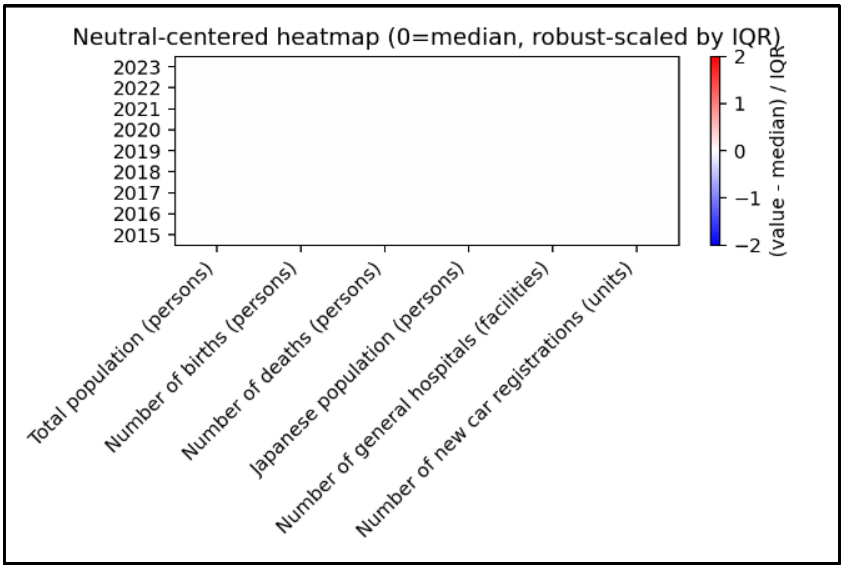

Figure 1. Heatmap generated by all items and years were set to 0.5.

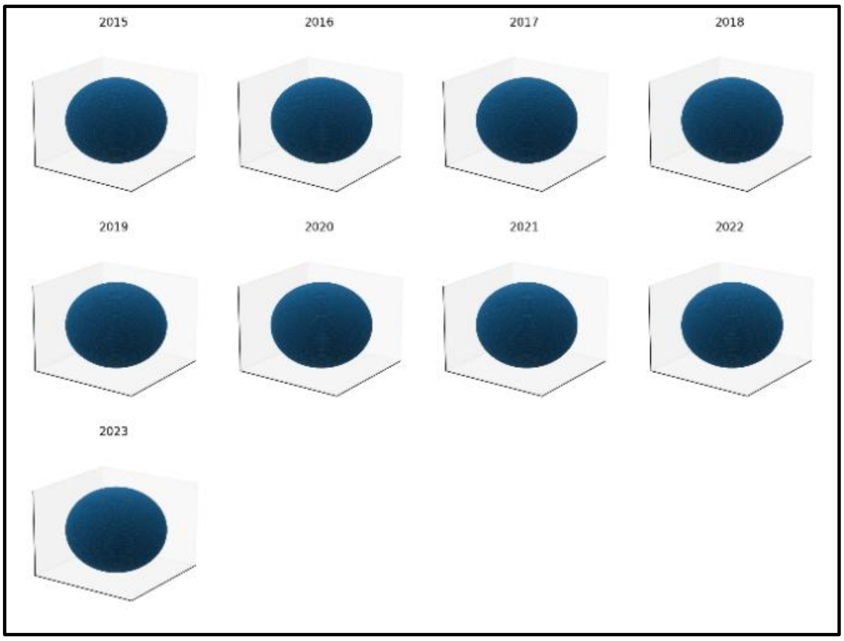

Figure 2. Social Morphology Ball generated by all items and years were set to 0.5.

Only two specific variables (Supplement; Soc_OppositeTrend.csv) —Total Population and Number of New Car Registrations—were varied in opposite phases while all other variables were fixed at their median values, the resulting geometry exhibited a pronounced polar elongation. As one variable increased and the other decreased relative to their respective medians, the sphere deformed symmetrically along the vertical axis, forming a distinct bipolar distortion (Figures 3 and 4). Note that the geometric effect of any pair of indicators depends on their assignment to bases; e.g., "polar elongation" arises when indicators are mapped to polar-sensitive bases such as cos θ.

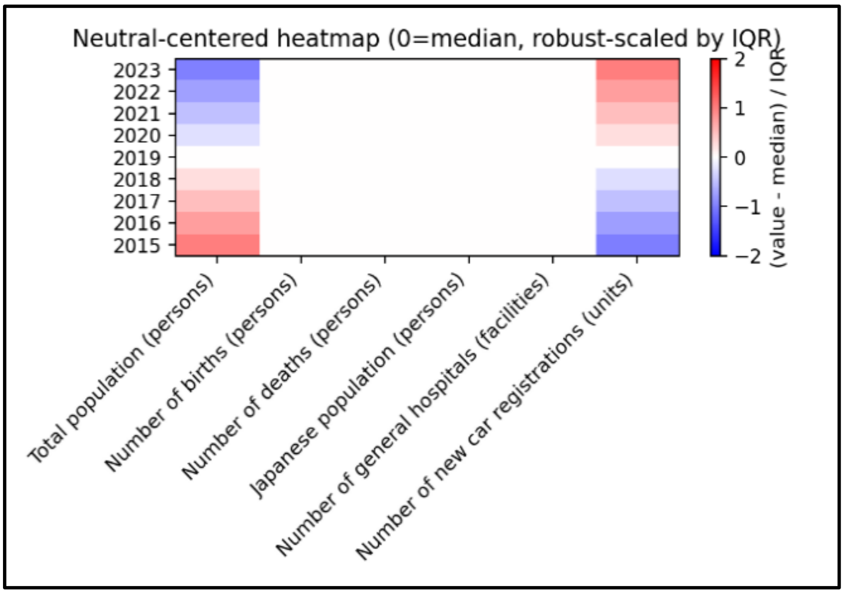

Figure 3. Heatmap generated by adjusting two items.

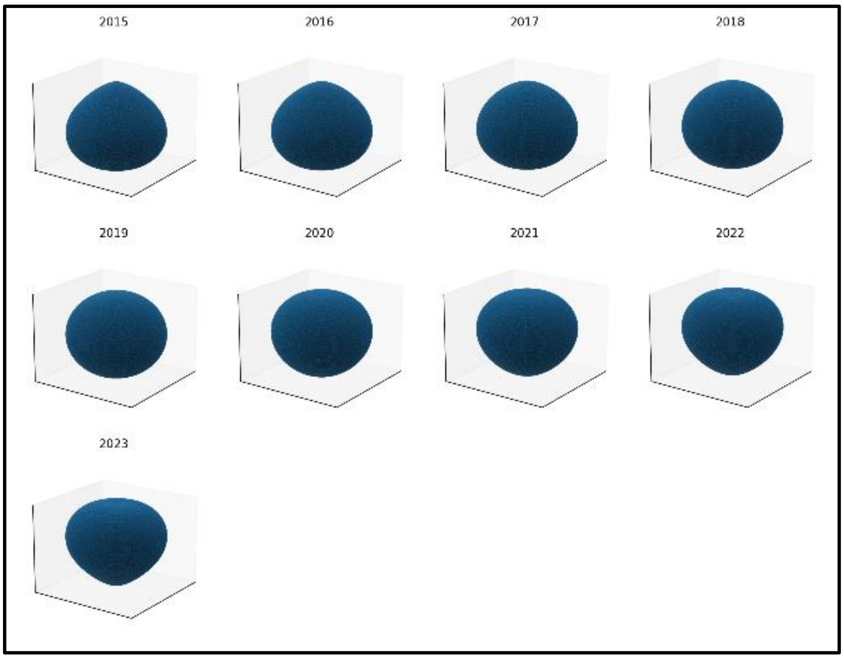

Figure 4. Social Morphology Ball generated by adjusting two items.

Next, a CSV file containing actual social statistical data (Supplementary; Soc_Indicators_Jpn.csv) was used to generate the spherical models. The resulting visualizations revealed that fluctuations in the input variables produced corresponding geometric deformations in the spheres, indicating that the model successfully translated variations in the multidimensional data into three-dimensional morphological changes (Figures 5 and 6).

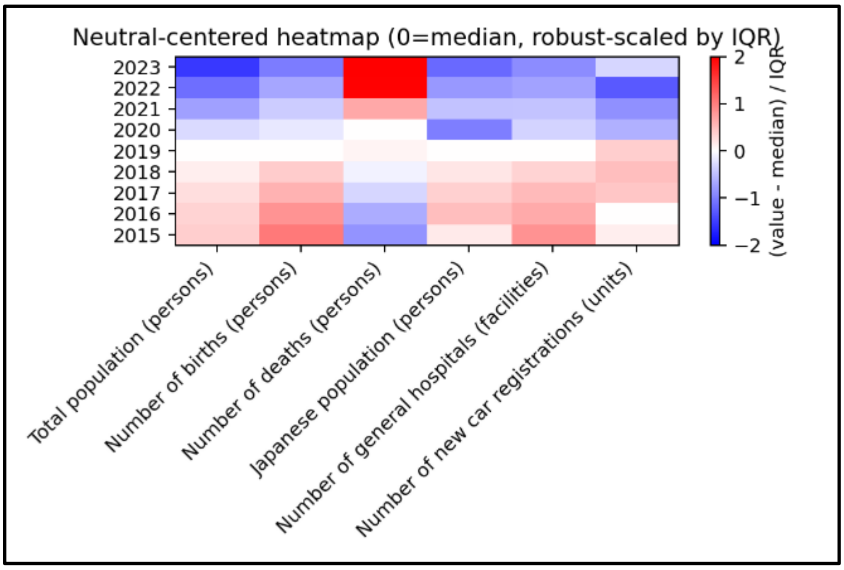

Figure 5. Heatmap generated from Japanese social data.

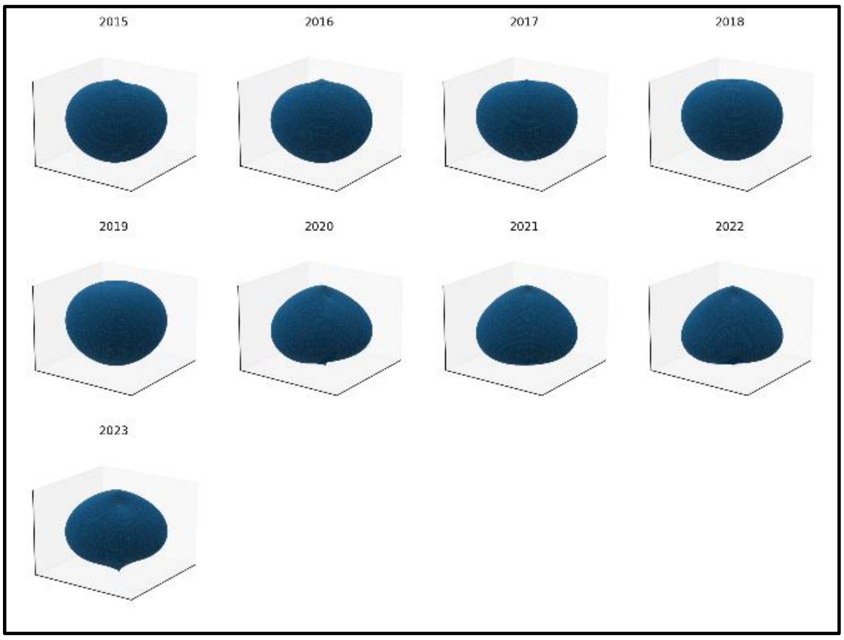

Figure 6. Social Morphology Ball generated from Japanese social data.

4. Discussion and Outlook

4.1 Morphological extensions and potential applications

In the visualization generated from actual social statistical data, the sphere corresponding to 2019 appeared morphologically smoother than those of other years. This outcome reflects that the input variables for 2019 were numerically closer to their respective medians, resulting in smaller normalized deviations and, consequently, a shape nearer to the morphologically neutral baseline. In this model, each social indicator is mapped to a specific spherical basis function, and the overall form of the sphere emerges as the superposition of these basis components weighted by the normalized values of each variable. Therefore, variations in the numerical values of each indicator directly translate into geometric deformations along the directions represented by their respective basis functions. Furthermore, because the assignment of variables to basis functions depends on the order of the column in the input CSV file, rearranging the sequence of variables alters the spatial correspondence between social indicators and the surface regions of the sphere. This property indicates that the "morphological correspondence" between data dimensions and spatial directions can be flexibly defined, providing an experimental framework for exploring structural relationships among multivariate indicators through shape transformations. In this framework, the spherical model serves as the fundamental form for morphologically reconstructing multivariate social data. By extending the model to other geometric structures, the proposed approach could help represent the multilayered and nonlinear characteristics inherent in social systems. In contrast to spherical visualization techniques commonly used in fields such as neuroscience or geospatial analysis, this framework does not merely project existing variables onto a fixed sphere. Instead, it reconstructs the sphere itself as a dynamic morphological representation, in which surface deformations reflect the relative balance and distortions among multiple social dimensions. The spherical basis expansion that forms the foundation of this method can, in principle, be extended to other geometric structures with relative ease. Several possible extensions are envisioned:

• Polyhedral model: By treating variable groups as faces or vertices, this model can emphasize the tensions and network structures among the elements constituting society.

• Torus model: Suitable for representing cyclic or recurrent social phenomena such as economic fluctuations or seasonal health indicators.

• Hyperbolic or saddle-surface model: Applicable to illustrating social polarization or bifurcated structural patterns within a population.

4.2 Applications and Perspectives

This sensory interpretation aligns with emerging discussions in data physicalization and embodied cognition, both of which emphasize the importance of tactile and spatial perception in understanding complex information systems. Within this context, the proposed model transforms social indicators into physical morphology, enabling observers to intuitively perceive social equilibrium and distortion as tangible forms. In this proof of concept, we introduce a novel approach to interpreting social phenomena from a sensory perspective—"feeling" the texture of the ball as a metaphor for perceiving social fluctuations. This framework extends beyond conventional data analysis, offering potential applications in education, art, and policy development.

• Education: In educational settings, 3D-printed models can serve as tactile teaching materials that allow students to grasp the structure of society in a tangible manner.

• Policy-making: In policy contexts, the model can intuitively visualize imbalances among social indicators, contributing to clearer identification of societal challenges.

• Art and media: In artistic and media domains, data can be presented as sculptural objects, enabling new forms of expression that convey the "shape" of social phenomena.

Furthermore, by focusing on the structural characteristics of data—such as morphological distortions—rather than on numerical differences alone, this approach provides a novel analytical perspective distinct from conventional statistical or graph-based methods. By differentially analyzing temporal changes, it becomes possible to trace dynamic structural deformations and identify turning points in social transformation as morphological transitions. Given that social phenomena are inherently complex and shaped by multiple interacting factors, integrating and visualizing these multidimensional relationships in three-dimensional form offers a meaningful framework for capturing the structural dynamics of social systems.

4.3 Limitations of the Present Method

This method presents certain limitations. First, there is an inherent ambiguity in determining which specific variables and basis function combinations contribute to localized deformations on the spherical surface. Therefore, rigorous scientific interpretation should be complemented by conventional statistical analyses for validation. Second, because variables such as population or number of medical institutions are mapped onto spatial basis functions for visualization, the resulting shapes do not derive from strict mathematical foundations. As a result, the analysis may carry elements of interpretive arbitrariness, and caution is required when deriving theoretical conclusions. Finally, this study primarily aims to propose a conceptual framework rather than a finalized analytical model. The Python script and mathematical expressions presented herein were developed for proof-of-concept purposes and may contain incomplete implementations or potential errors. These components will be refined and validated in future work as the theoretical model is further developed.

Conflict of Interest
The author declares that there are no known conflicts of interest or competing financial interests related to this study.

References
1. Zakkar M, Sedig K. Interactive visualization of public health indicators to support policymaking: an exploratory study. Online Journal of Public Health Informatics. 2017;9(2):e190.
2. Shaban-Nejad A, Lavigne M, Okhmatovskaia A, Buckeridge DL. PopHR: a knowledge-based platform to support integration, analysis, and visualization of population health data. Annals of the New York Academy of Sciences. 2017;1387(1):44–53.
3. Leong P, Carlile S. Methods for spherical data analysis and visualization. Journal of Neuroscience Methods. 1998;80(2):191–200.
4. Lehanneur M. State of the World, Sculptures. Mathieu Lehanneur Studio; 2021. Available from: https://www.mathieulehanneur.fr/project/state-of-the-world-sculptures-297 . Accessed October 16, 2025.
5. Ministry of Internal Affairs and Communications (Japan). Statistics Dashboard. Available from: https://dashboard.e-stat.go.jp/ . Accessed October 14, 2025.