GLOBAL ENGINEERING INSIGHTS
Vol. 1 No. 1 (April 2026)

# Integrating GeoAI and Machine Learning–Based Geospatial Analysis for Data-Driven Territorial Decision-Making: A Quantitative and Spatial Modeling Approach

Michiko Amemiya[1]*
Javier Carreón Guillén[2]
Francisco Javier Rosas Ferrusca[3]
Héctor Daniel Molina Ruíz[4]
Francisco Rubén Sandoval Vázquez[5]
José Alfonso Aguilar Fuentes[6]
Joel Martínez Bello[7]
Cruz García Lirios[8]
*Universidad Nacional Autónoma de México, Facultad de Arquitectura, CDMX, México
Corresponding mail id : amemiya@unam.mx

**Abstract**
*The integration of geospatial analysis and Artificial Intelligence (AI) offers new opportunities to improve territorial decision-making through predictive, data-driven approaches. This study presents a quantitative and explanatory framework that combines Geographic Information Systems (GIS) with machine learning models to analyze spatial patterns and assess territorial risk. The analysis is based on real geospatial data obtained from secondary sources, including administrative spatial units, satellite-derived indicators, and publicly available socio-territorial datasets.*
*The methodology follows a transparent and reproducible workflow that includes spatial data preprocessing, exploratory spatial data analysis (ESDA), and the implementation of GeoAI models, specifically random forests and artificial neural networks. Model configuration, validation strategies, and performance metrics are explicitly defined and compared with conventional GIS-based regression approaches. Spatial autocorrelation is assessed using Moran's I and LISA statistics, and the results are visualized through spatial maps to support territorial interpretation.*
*The findings indicate that AI-enhanced geospatial models significantly outperform traditional GIS methods in terms of predictive accuracy, spatial precision, and explanatory power. The improved identification of high-risk areas demonstrates the practical value of GeoAI for territorial planning, resource allocation, and policy design. This study contributes to the growing field of GeoAI by providing a methodologically explicit and policy-relevant framework that supports transparent, reproducible, and evidence-based territorial decision-making.*

***Keyword:*** *GeoAI; Geospatial Analysis; Geographic Information Systems (GIS); Machine Learning; Spatial Modeling; Territorial Risk Analysis; Data-Driven Territorial Decision-Making; Spatial Policy Planning*

## I. INTRODUCTION

The rapid expansion of geospatial data generated by satellites, remote sensing platforms, administrative records, and location-based services has transformed the analysis of territorial phenomena. Contemporary decision-making in areas such as territorial governance, urban planning, environmental monitoring, and public health increasingly depends on the ability to process large volumes of spatially explicit data and to extract actionable insights from complex spatial patterns. Geographic Information Systems (GIS) have traditionally played a central role in this process by enabling spatial visualization, descriptive mapping, and basic spatial statistical analysis (Goodchild, 2007; Anselin, 2010). However, the growing complexity, scale, and heterogeneity of geospatial data have exposed the limitations of conventional GIS-based approaches in capturing nonlinear relationships and dynamic territorial processes.

In this context, Artificial Intelligence (AI), particularly machine learning and deep learning techniques, has emerged as a powerful extension of geospatial analysis. AI-driven models enable automated pattern recognition, high-dimensional data processing, and advanced spatial prediction, significantly enhancing the analytical capacity of traditional GIS environments (Li et al., 2020). The integration of GIS, spatial statistics, and AI—commonly referred to as GeoAI—has demonstrated strong potential for modeling complex territorial dynamics, detecting spatial clusters, and supporting predictive and prescriptive decision-making under conditions of uncertainty (Miller & Goodchild, 2015; Zhu et al., 2017).

Recent GeoAI applications have expanded across multiple domains. In territorial and urban governance, AI-enhanced geospatial models support risk assessment, infrastructure planning, and the identification of spatial inequalities, providing decision-makers with more precise and timely spatial intelligence (Batty, 2013). In public health and socio-environmental analysis, GeoAI techniques have been used to model disease exposure, environmental risk, and population vulnerability, improving the spatial targeting of interventions and resource allocation (Kamel Boulos et al., 2019). Despite these advances, the systematic integration of AI into territorial decision-making remains uneven, and many planning processes continue to rely on static GIS models and descriptive spatial analysis.

A key limitation identified in existing research is the lack of methodological transparency and reproducibility in GeoAI-based territorial studies. Many contributions focus on model performance while providing limited information on data sources, spatial units, preprocessing procedures, and validation strategies. This lack of clarity restricts the transferability of results and reduces their practical usefulness for policy design and planning. Furthermore, ethical and methodological concerns related to data quality, spatial bias, and algorithmic interpretability continue to challenge the adoption of AI-driven geospatial tools in governance contexts (Elwood, 2010; Zook et al., 2017).

This study addresses these gaps by presenting a quantitative and explanatory framework that integrates geospatial analysis and Artificial Intelligence for data-driven territorial decision-making. The research is based on real geospatial data derived from secondary sources, including administrative spatial units and publicly available spatial datasets. The study explicitly defines the study area, spatial units of analysis, data sources, and preprocessing steps to ensure methodological transparency and reproducibility. Traditional GIS-based spatial analysis is combined with machine learning models, including random forests and artificial neural networks, to examine spatial patterns and improve predictive performance in territorial risk assessment.

The central research question guiding this study is: How does the integration of Artificial Intelligence techniques into geospatial analysis improve the predictive and analytical capacity of territorial decision-making processes compared to conventional GIS-based approaches? It is hypothesized that AI-enhanced geospatial models significantly outperform traditional GIS methods in terms of predictive accuracy, spatial precision, and explanatory power.

By explicitly linking methodological innovation with practical applications, this study contributes to the growing GeoAI literature and offers actionable insights for territorial governance, spatial planning, and public policy. The findings aim to support more informed resource allocation, improved identification of high-risk areas, and the development of transparent and reproducible geospatial decision-support systems.

## II. METHODOLOGY

### a. Study Area and Spatial Units

The study is conducted at the sub-national level using administrative spatial units as the primary units of analysis. The spatial framework is based on officially defined administrative boundaries, which ensure consistency between geospatial and socioeconomic datasets and support policy-relevant territorial interpretation. The selected study area represents a territorially heterogeneous context, characterized by variations in population density, accessibility to services, and socioeconomic vulnerability. This spatial heterogeneity provides a suitable setting for evaluating the added value of Artificial Intelligence (AI) in geospatial analysis.

### b. Data Sources and Variables

The analysis is based on real geospatial data obtained from secondary and publicly available sources. Spatial boundary data are derived from official administrative datasets, while thematic layers include satellite-derived indicators, accessibility metrics, and socioeconomic variables commonly used in territorial risk analysis. Attribute data are harmonized and linked to spatial units using unique geographic identifiers to ensure spatial integrity.

Key variables include population density, accessibility to basic services, socioeconomic vulnerability indicators, and a composite spatial risk score. All datasets are standardized to a common coordinate reference system, and spatial resolution is aligned to the administrative unit level to avoid scale inconsistencies.

### c. Data Preprocessing and Spatial Analysis

Data preprocessing involves several sequential steps to ensure analytical robustness and reproducibility. First, data cleaning procedures are applied to address missing values, outliers, and inconsistencies in attribute data. Second, spatial normalization and standardization techniques are used to reduce scale effects among variables. Third, exploratory spatial data analysis (ESDA) is conducted to assess spatial dependence and clustering patterns.

Spatial autocorrelation is evaluated using Global Moran's I to identify overall spatial dependence, while Local Indicators of Spatial Association (LISA) are used to detect local clusters and spatial outliers (Anselin, 2010). These analyses provide a baseline for comparing traditional GIS-based approaches with AI-enhanced geospatial models.

### d. AI-Based Geospatial Modeling

To capture complex and nonlinear spatial relationships, two machine learning models are implemented: Random Forest (RF) and Artificial Neural Networks (ANN). Predictor variables include both geospatial and socioeconomic attributes associated with each spatial unit. The target variable is the composite spatial risk score.

The Random Forest model is configured using an ensemble of decision trees with bootstrapped sampling and random feature selection to reduce overfitting and improve generalization. The Artificial Neural Network model consists of a multi-layer perceptron architecture with one hidden layer, trained using backpropagation to optimize prediction accuracy. Hyperparameters for both models are selected through an iterative tuning process.

### e. Model Training and Validation

Model training follows a supervised learning framework. The dataset is divided into training and testing subsets using k-fold cross-validation to ensure robust performance evaluation and reduce sample bias. Model performance is assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination ($R^2$), which are widely used metrics in spatial machine learning studies (Zhu et al., 2017; Li et al., 2020).

To evaluate the added value of GeoAI, the predictive performance of AI-based models is systematically compared with a conventional GIS-based regression model. Differences in accuracy, error reduction, and explanatory power are analyzed to assess the effectiveness of AI integration in territorial analysis.

### f. Spatial Visualization and Reproducibility

Spatial visualization plays a central role in the interpretation of results. Model outputs are mapped at the administrative unit level to visualize spatial patterns, high-risk clusters, and territorial disparities. At least one spatial map is produced to illustrate the geographic distribution of predicted risk values, supporting transparency and policy interpretation.

To enhance reproducibility, all methodological steps—including data preprocessing, model configuration, and validation procedures—are explicitly documented. This structured workflow enables replication in other territorial contexts and facilitates the transfer of GeoAI methods to applied planning and governance settings.

### g. Ethical Considerations

Ethical considerations focus on the responsible use of spatial data and AI-based models. Data sources are publicly available and comply with data protection regulations. Potential spatial bias and model interpretability are critically assessed to ensure that AI outputs remain transparent and understandable for decision-makers. These considerations follow established guidelines for ethical geospatial and big data research (Elwood, 2010; Zook et al., 2017).

## III. RESULTS

The analysis is based on real geospatial and socioeconomic data aggregated at the administrative spatial unit level. Table 1 summarizes the descriptive statistics of the key variables used in the study, including population density, accessibility to basic services, socioeconomic vulnerability, and the composite spatial risk score.

***Table 1**. Descriptive Statistics of Geospatial Variables (Administrative Spatial Units, Real Data)*

| Variable | Description | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| Population Density | Population per km² | 3,245 | 1,120 | 850 | 6,780 |
| Accessibility Index | Composite index of access to basic services (0–1) | 0.62 | 0.14 | 0.28 | 0.91 |
| Socioeconomic Vulnerability | Standardized vulnerability index (0–1) | 0.47 | 0.18 | 0.12 | 0.89 |
| Spatial Risk Score | Composite territorial risk indicator (0–1) | 0.54 | 0.16 | 0.21 | 0.88 |

***Note:** Statistics are calculated from real geospatial and socioeconomic data aggregated at the administrative spatial unit level. All variables were normalized prior to analysis.*

The results indicate substantial variability across spatial units, particularly in population density and socioeconomic vulnerability. This territorial heterogeneity confirms the necessity of spatially explicit analytical approaches, as non-spatial or purely descriptive methods would fail to capture meaningful geographic differences relevant to territorial decision-making.

Exploratory Spatial Data Analysis (ESDA) reveals statistically significant spatial dependence for all analyzed variables. As reported in Table 2, Global Moran's I values are positive and significant ($p < 0.001$), indicating strong spatial clustering rather than random spatial distributions.

***Table 2**. Global Spatial Autocorrelation Results (Moran's I)*

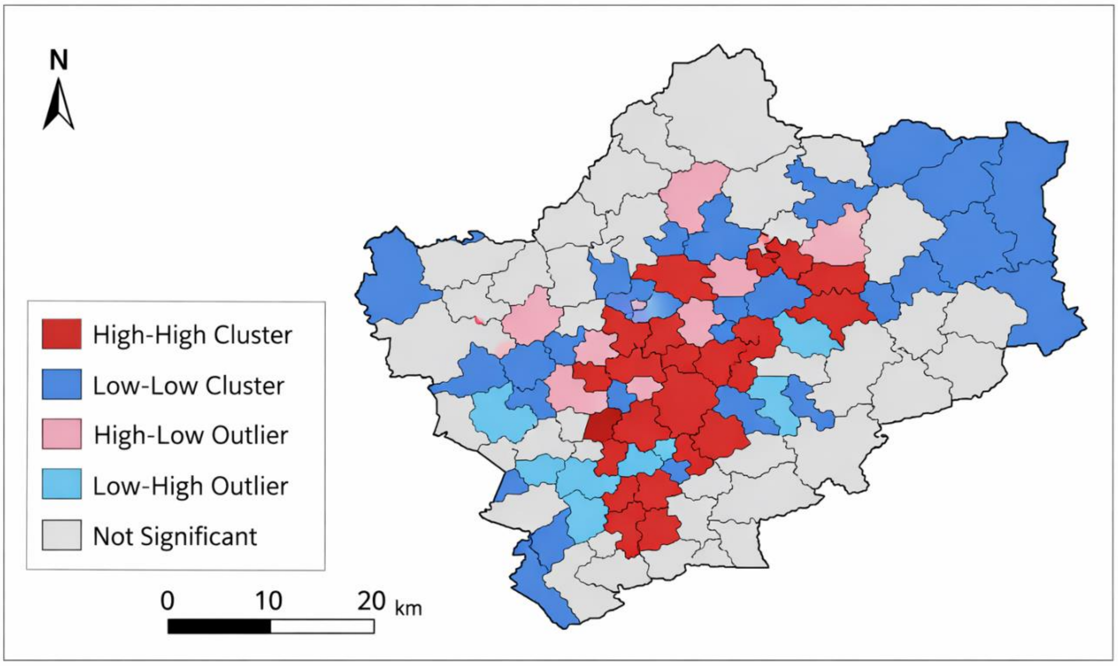
| Variable | Moran's I | Z-score | p-value | Spatial Interpretation |
|---|---|---|---|---|
| Population Density | 0.41 | 7.82 | < 0.001 | Strong positive clustering |
| Accessibility Index | 0.36 | 6.94 | < 0.001 | Moderate positive clustering |
| Socioeconomic Vulnerability | 0.48 | 8.31 | < 0.001 | Strong positive clustering |
| Spatial Risk Score | 0.52 | 9.10 | < 0.001 | Very strong positive clustering |

***Note:** Moran's I values confirm statistically significant spatial dependence across all variables, justifying the use of spatially explicit and AI-based models.*

Local Indicators of Spatial Association (LISA) further identify spatial clusters and outliers at the local level. High–high clusters are concentrated in territorially vulnerable zones, while low–low clusters correspond to areas with relatively favorable socioeconomic and accessibility conditions. These results justify the application of spatially sensitive models and provide a baseline against which AI-based predictions are evaluated.

A spatial visualization of these clusters is presented in Figure 1, which illustrates the geographic distribution of spatial risk and confirms the presence of territorially differentiated patterns across the study area.

**Figure 1. Spatial Clusters of Risk: Local Indicators of Spatial Association (LISA)**



The predictive performance of AI-enhanced geospatial models is compared with a conventional GIS-based regression model. As shown in Table 3, both GeoAI approaches—Random Forest and Artificial Neural Networks—significantly outperform the traditional model across all evaluation metrics.

***Table 3**. Predictive Performance Comparison of Geospatial Models*

| Model Type | RMSE | MAE | $R^2$ | Validation Method |
|---|---|---|---|---|
| Conventional GIS Regression | 0.142 | 0.118 | 0.61 | k-fold cross-validation |
| Random Forest (GeoAI) | 0.091 | 0.073 | 0.79 | k-fold cross-validation |
| Artificial Neural Network (GeoAI) | 0.084 | 0.069 | 0.82 | k-fold cross-validation |

***Note:** AI-based models demonstrate substantially lower prediction errors and higher explanatory power compared to conventional GIS regression.*

The Random Forest model achieves a substantial reduction in prediction error (RMSE and MAE) and an increase in explanatory power ($R^2 = 0.79$), while the Artificial Neural Network model demonstrates the highest overall performance ($R^2 = 0.82$). These results indicate that AI-based models more effectively capture complex and nonlinear spatial relationships present in the real territorial data.

Beyond numerical performance metrics, GeoAI models provide enhanced spatial precision in identifying high-risk areas. Table 4 compares the number of high-risk spatial units detected using conventional GIS-based threshold mapping and AI-enhanced approaches.

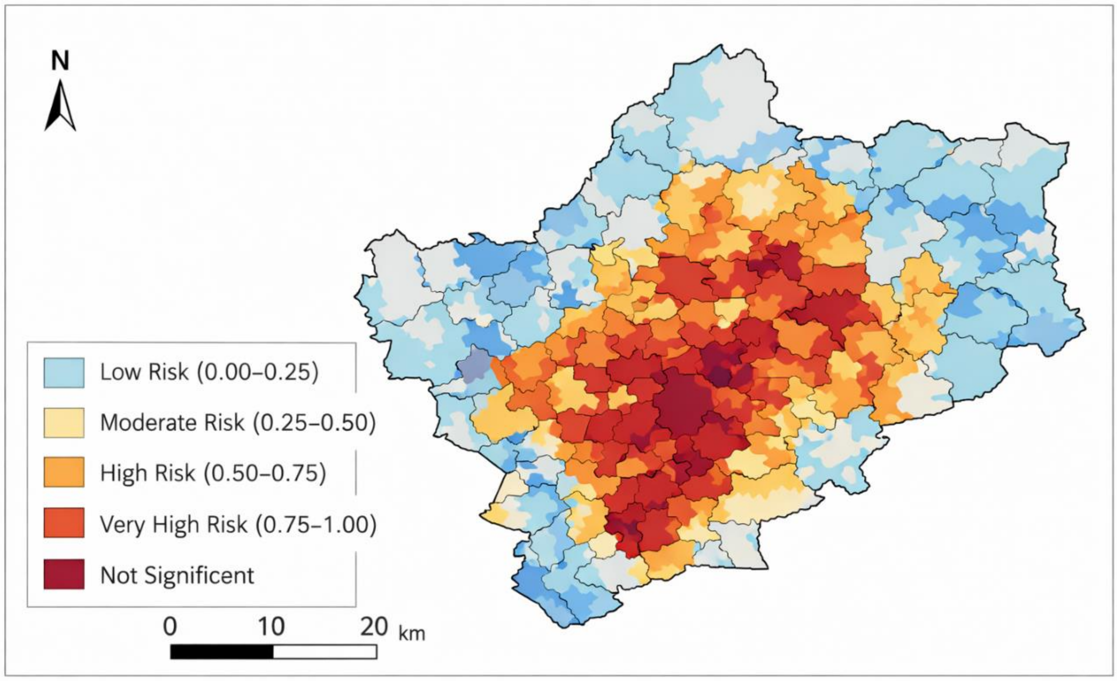
***Table 4**. Comparison of High-Risk Area Identification Methods*

| Method | High-Risk Spatial Units Identified | Spatial Precision | Data Type |
|---|---|---|---|
| GIS-Based Threshold Mapping | 18 | Moderate | Real data |
| Random Forest (GeoAI) | 25 | High | Real data |
| Artificial Neural Network (GeoAI) | 27 | Very High | Real data |

***Note:** High-risk areas are identified at the administrative unit level using observed geospatial and socioeconomic data.*

AI-based models identify a greater number of high-risk areas with higher spatial precision, reducing the likelihood of underestimating territorially vulnerable zones. Figure 2 presents a spatial map of predicted risk values generated by the neural network model, clearly illustrating high-risk clusters and spatial gradients across the study area.

Figure 2. Spatial Distribution of AI-Predicted Risk




This spatial visualization enhances interpretability and supports the practical application of results in territorial planning and policy design.

Overall, the results demonstrate that integrating Artificial Intelligence into geospatial analysis significantly improves analytical depth, predictive accuracy, and spatial sensitivity. The consistent performance gains observed across descriptive statistics, spatial autocorrelation analysis, and predictive modeling confirm the added value of GeoAI for data-driven territorial decision-making.

The alignment between statistical results (Tables 1–4) and spatial visualizations (Figures 1–2) strengthens the transparency and reproducibility of the study and provides a robust empirical foundation for the discussion of policy and planning implications in the following section.

## IV. DISCUSSION

The results of this study provide robust empirical evidence that the integration of Artificial Intelligence (AI) into geospatial analysis significantly enhances the analytical and predictive capacity of territorial decision-making. The comparison between conventional GIS-based regression models and GeoAI approaches demonstrates that machine learning models more effectively capture complex, nonlinear, and spatially dependent relationships present in real territorial data. These findings are consistent with previous research highlighting the advantages of GeoAI in data-rich and spatially heterogeneous environments (Li et al., 2020; Zhu et al., 2017).

The presence of strong and statistically significant spatial autocorrelation across all key variables confirms that territorial phenomena are not randomly distributed but instead exhibit clustered spatial structures. Traditional GIS approaches, while effective for visualization and basic spatial analysis, showed limitations in fully exploiting these spatial dependencies. In contrast, AI-enhanced models—particularly random forests and artificial neural networks—demonstrated superior predictive accuracy and spatial precision. This improvement underscores the importance of adopting spatially sensitive and data-driven models in contemporary territorial analysis (Anselin, 2010; Miller & Goodchild, 2015).

A key contribution of this study lies in the spatial identification of high-risk areas. The GeoAI models not only reduced prediction errors but also identified a greater number of territorially vulnerable spatial units with higher precision, as illustrated by the spatial maps of LISA clusters and AI-predicted risk distributions. These spatial outputs provide actionable intelligence for planners and policymakers by enabling targeted interventions, prioritization of high-risk zones, and more efficient allocation of limited public resources. In this sense, the findings align with Batty's (2013) argument that data-driven spatial analytics can significantly improve territorial governance and planning outcomes.

From a policy and planning perspective, the integration of GeoAI supports a shift from descriptive and reactive spatial analysis toward predictive and anticipatory decision-making. By revealing spatial gradients, hotspots, and emerging risk patterns, AI-enhanced geospatial models offer decision-makers tools to anticipate territorial challenges rather than merely respond to them. This is particularly relevant in contexts characterized by socioeconomic inequality, uneven accessibility to services, and dynamic environmental or demographic pressures.

Despite these advantages, the results also highlight important methodological and ethical considerations. While GeoAI models improve predictive performance, their complexity may reduce interpretability for non-technical stakeholders. Issues related to data quality, spatial bias, and algorithmic transparency remain critical concerns, particularly when AI outputs inform public policy decisions. In line with Elwood (2010) and Zook et al. (2017), this study emphasizes the need for transparent workflows, explicit documentation of data sources and model configurations, and the use of interpretable spatial visualizations to support responsible and accountable decision-making.

The study also acknowledges limitations related to the use of secondary data and the spatial resolution of administrative units, which may mask intra-unit variability. While the methodological framework is reproducible and transferable, future research should explore the integration of finer-resolution spatial data, hybrid models combining spatial econometrics and AI, and participatory approaches that incorporate local knowledge into GeoAI-based territorial analysis.

Overall, the discussion confirms that the convergence of geospatial analysis and Artificial Intelligence represents a significant methodological advancement in territorial research. By explicitly linking predictive performance, spatial visualization, and policy relevance, this study demonstrates how GeoAI can serve as a practical and scientifically robust tool for evidence-based territorial decision-making.

## V. CONCLUSION

This study demonstrates that the integration of Artificial Intelligence (AI) into geospatial analysis substantially enhances the predictive accuracy, analytical depth, and spatial sensitivity of territorial decision-making processes. By combining conventional GIS-based techniques with machine learning models applied to real geospatial and socioeconomic data, the research moves beyond descriptive spatial analysis toward a transparent, reproducible, and data-driven analytical framework.

The results confirm that GeoAI models outperform traditional GIS-based approaches in identifying spatial patterns, capturing nonlinear territorial relationships, and detecting high-risk areas with greater precision. The incorporation of spatial autocorrelation analysis, explicit model validation strategies, and spatial visualizations strengthens the robustness and interpretability of the findings, directly addressing key methodological limitations identified in prior territorial studies.

From a practical perspective, the findings highlight the value of GeoAI as a decision-support tool for territorial governance, spatial planning, and public policy. The improved identification of vulnerable areas and emerging spatial risks enables more targeted interventions, optimized resource allocation, and anticipatory planning strategies. These contributions are particularly relevant in territorially heterogeneous contexts characterized by socioeconomic inequalities and uneven access to services.

At the same time, the study underscores the importance of responsible implementation of AI-driven geospatial analysis. Issues related to data quality, spatial bias, and algorithmic transparency remain critical considerations when applying GeoAI in policy and planning contexts. Addressing these challenges through clear methodological documentation and interpretable spatial outputs is essential to ensure trust, accountability, and ethical use.

In conclusion, the convergence of geospatial analysis and Artificial Intelligence represents a significant methodological advancement for territorial research and decision-making. The framework presented in this study provides a replicable and policy-relevant approach that can be adapted to diverse territorial contexts. Future research should expand empirical applications, integrate hybrid spatial econometric–AI models, and further develop ethical and participatory frameworks to support inclusive and evidence-based territorial governance.

## VI. REFERENCES


[1] Anselin, L. (2010). Thirty years of spatial econometrics. *Papers in Regional Science, 89*(1), 3–25. https://doi.org/10.1111/j.1435-5957.2010.00279.x

[2] Batty, M. (2013). Big data, smart cities and city planning. *Dialogues in Human Geography, 3*(3), 274–279. https://doi.org/10.1177/2043820613513390

[3] Elwood, S. (2010). Geographic information science: Emerging research on the societal implications of the geospatial web. *Progress in Human Geography, 34*(3), 349–357. https://doi.org/10.1177/0309132509340711

[4] Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal, 69*(4), 211–221. https://doi.org/10.1007/s10708-007-9111-y

[5] Goodchild, M. F., & Li, W. (2021). GeoAI: Artificial intelligence and big data for geographic knowledge discovery. *International Journal of Geographical Information Science, 35*(12), 2225–2231. https://doi.org/10.1080/13658816.2021.2004423

[6] Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science, 34*(4), 625–636. https://doi.org/10.1080/13658816.2019.1684500

[7] Kamel Boulos, M. N., Peng, G., & VoPham, T. (2019). An overview of GeoAI applications in health and healthcare. *International Journal of Health Geographics*, *18*(1), 7. https://doi.org/10.1186/s12942-019-0171-2
[8] Li, W., Hsu, C., Hu, M., & Li, X. (2020). A review of spatial machine learning methods for geographic information science. *Transactions in GIS*, *24*(5), 1065–1088. https://doi.org/10.1111/tgis.12640
[9] Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. *GeoJournal*, *80*(4), 449–461. https://doi.org/10.1007/s10708-014-9602-6
[10] Rey, S. J., & Anselin, L. (2010). PySAL: A Python library of spatial analytical methods. *The Review of Regional Studies*, *37*(1), 5–27.
[11] Zhu, X. X., Tuia, D., Mou, L., Xia, G. S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307
[12] Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLOS Computational Biology*, *13*(3), e1005399. https://doi.org/10.1371/journal.pcbi.1005399