INTELLECTUAL INTEGRATION OF MULTI-SOURCE REMOTE SENSING DATA FOR LANDSCAPE MONITORING: ADDRESSING DATA HETEROGENEITY AND INTEROPERABILITY CHALLENGES

Juraev Azizbek
Assistant Lecturer of Tashkent University of Information Technologies
Tashkent, Uzbekistan

Abstract. In automated classification and integration of geospatial data for landscape monitoring based on remote sensing, the problem of "data heterogeneity and interoperability gap" creates significant difficulties. Traditional algorithms are capable of identifying basic land cover types from individual sources, yet they cannot meaningfully fuse data from disparate sensors, formats, schemas, and semantic contexts. This article proposes a methodology for achieving semantic interoperability through ontologies, semantic web technologies, and Earth Observation Data Cubes (EODCs). Using multi-source data from Sentinel-2, Landsat, and complementary platforms, the study analyses syntactic, schematic, and semantic heterogeneity, along with fusion techniques and OGC standards. The research demonstrates how ontologies and AI-driven data cubes overcome the interoperability gap between heterogeneous geospatial datasets and their ecological/analytical interpretation.

Keywords: data heterogeneity, interoperability, remote sensing, ontologies, semantic web technologies, multi-source data fusion, earth observation data cubes, landscape monitoring.

I. INTRODUCTION

Modern geoinformation systems and Earth Observation (EO) infrastructures have transformed automated landscape analysis from simple land-cover mapping into a complex multi-source integration process. Traditional methods can classify basic features from single-sensor imagery, yet they fail to reconcile differences in spatial/temporal resolution, data formats, schemas, and semantic meaning across platforms. This limitation is known as the “data heterogeneity and interoperability gap” - the mismatch between raw multi-source data and their meaningful ecological or decision-support interpretation.

1.1 Data heterogeneity types and interoperability problems. The figure below illustrates heterogeneous data sources (structured, unstructured, and imagery) and the pathway to unified semantic integration. However, this approach alone does not resolve internal inconsistencies in resolution, semantics, or acquisition timing that prevent reliable landscape monitoring.

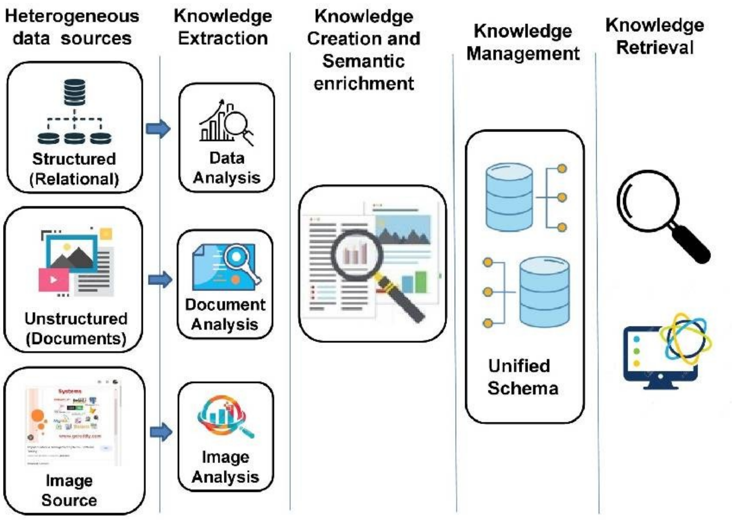

Figure 1. Semantic integration of heterogeneous data sources (structured, unstructured, and imagery) leading to a unified schema. Adapted from knowledge extraction and enrichment frameworks for geospatial applications.

Current multi-sensor and multi-platform remote sensing technologies (Sentinel-2, Landsat, SAR, hyperspectral) are widely applied for dynamic landscape monitoring. Time-series data from these sources enable change detection, yet heterogeneity in formats, schemas, and semantics remains a core barrier. The interoperability gap in multi-source EO data integration exacerbates global environmental and socio-economic threats. Below are four key global problems, each scientifically substantiated with significant ecological and societal consequences.

1. Exacerbation of land degradation monitoring inaccuracies and erosion processes: Heterogeneity prevents seamless fusion of optical and radar data, leading to incomplete assessments of soil loss and vegetation decline. This contributes to reduced agricultural productivity and threatens food security for millions, as data silos hinder timely detection of degradation hotspots.

2. Global spread of undetected ecosystem stress and biodiversity loss: Differences in sensor specifications and semantic labelling limit early detection of vegetation stress or invasive species across tropical and temperate regions. The gap delays biodiversity monitoring, accelerating species loss and destabilising carbon cycles amid climate change.

3. Intensification of climate change impacts and ecosystem fragmentation: Inaccurate integration of multi-sensor time series undermines tipping-point detection in forests, wetlands, and cryosphere. This amplifies uncertainties in global warming projections beyond 1.5°C, increasing extreme weather frequency, human displacement, and economic losses in trillions of dollars. **4. Crisis in food security and widening economic inequality:** Heterogeneous data impede precision agriculture and yield forecasting models, exacerbating hunger in vulnerable regions and hindering equitable resource allocation. The interoperability gap slows progress toward Sustainable Development Goals by limiting cross-platform analytics for policy and humanitarian response.

II. SCIENTIFIC ESSENCE OF DATA HETEROGENEITY AND INTEROPERABILITY

2.1 Description of the problem. Data heterogeneity comprises three main levels:

- **Syntactic heterogeneity** - differences in data formats and encoding (e.g., GeoTIFF vs. NetCDF).

- **Schematic heterogeneity** - structural mismatches in schemas or attribute definitions.

- **Semantic heterogeneity** - varying conceptual meanings of the same features across datasets.

Environmental factors such as varying acquisition times, atmospheric conditions, and sensor calibration further distort integration.

2.2 Characteristics of multi - source geospatial data. Different platforms produce complementary yet incompatible signatures. Optical sensors (Sentinel-2, Landsat) provide spectral information, while SAR offers structural data under cloud cover. The figure below illustrates a multi-sensor fusion framework that addresses these differences.

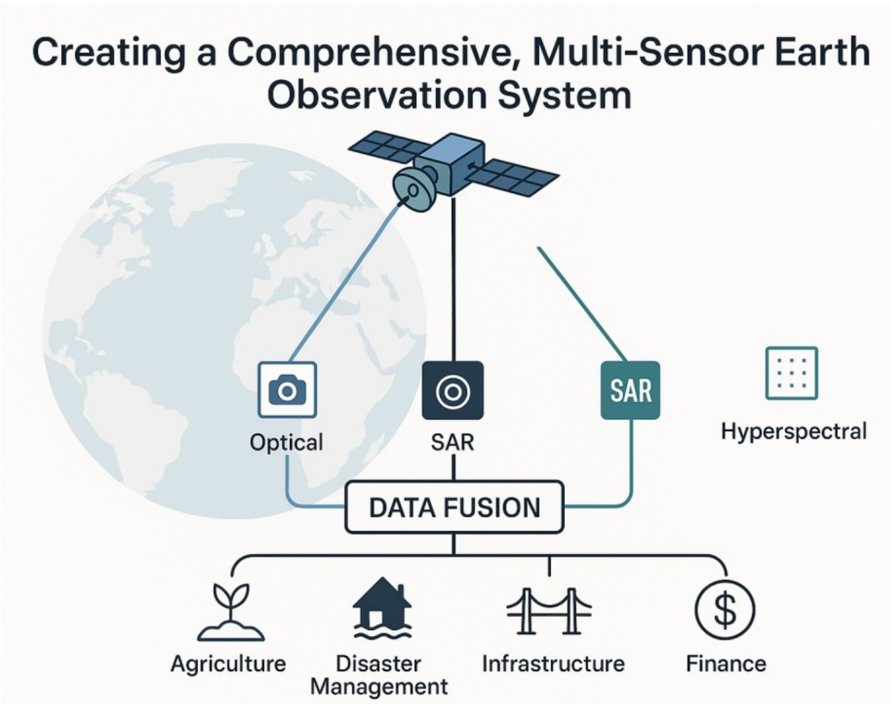

Figure 2. Multi-sensor Earth Observation fusion framework combining optical, SAR, and hyperspectral data for applications in agriculture, disaster management, infrastructure, and finance.

Classical fusion approaches struggle with temporal misalignment and resolution mismatches, limiting early stress detection.

2.3 Heterogeneity between data sources. Differences between Sentinel-2 (high revisit, red-edge bands) and Landsat (long-term archive) are pronounced in spectral coverage and radiometric calibration. Semantic mismatches arise when the same land-cover class is defined differently across ontologies or national standards.

2.4 Global solutions to data heterogeneity and interoperability in landscape monitoring. Modern approaches integrating ontologies, semantic web technologies, AI, and EODCs offer effective strategies. Below are targeted solutions for the four global problems.

1. Overcoming land degradation monitoring inaccuracies: Adoption of interoperable EODCs and OGC standards (WMS, WFS) enables harmonised multi-sensor analysis. Ontology-based data access (OBDA) and machine-learning fusion reduce uncertainties, supporting sustainable land management with high accuracy. **2. Reducing ecosystem stress and biodiversity loss:** Semantic reconciliation via domain ontologies combined with time-series data cubes allows early stress detection (R^2 improvements reported in multi-modal models). Cross-platform fusion with AI enhances biodiversity indicators and carbon-cycle monitoring. **3. Mitigating climate change impacts and fragmentation:** Spatial data infrastructures (SDIs) with semantic technologies integrate GEDI, GRACE, and Sentinel data for tipping-point analysis. AI-driven workflows and standardised cubes improve projections of extreme events and support nature-based solutions. **4. Addressing food security and economic inequality:** Precision agriculture platforms using multimodal transformers and interoperable APIs deliver accurate yield forecasts (low RMSE). Tools like Trends.Earth and open EODCs optimise resource allocation, advancing SDG implementation through equitable geospatial access.

III. ONTOLOGIES AND SEMANTIC TECHNOLOGIES FOR INTEROPERABILITY

3.1 Concept of geospatial ontologies. An ontology serves as a formal, machine-readable “semantic library” defining classes, properties, and relationships for landscape features, soil types, and vegetation states. Each source dataset maps to a shared conceptual model, resolving semantic heterogeneity.

3.2 Multi-source data models and standards. Sentinel-2 and Landsat differ in band configuration and metadata. OGC-compliant services and data cubes standardise access. The figure below shows an EODC architecture that enables seamless querying across heterogeneous archives.

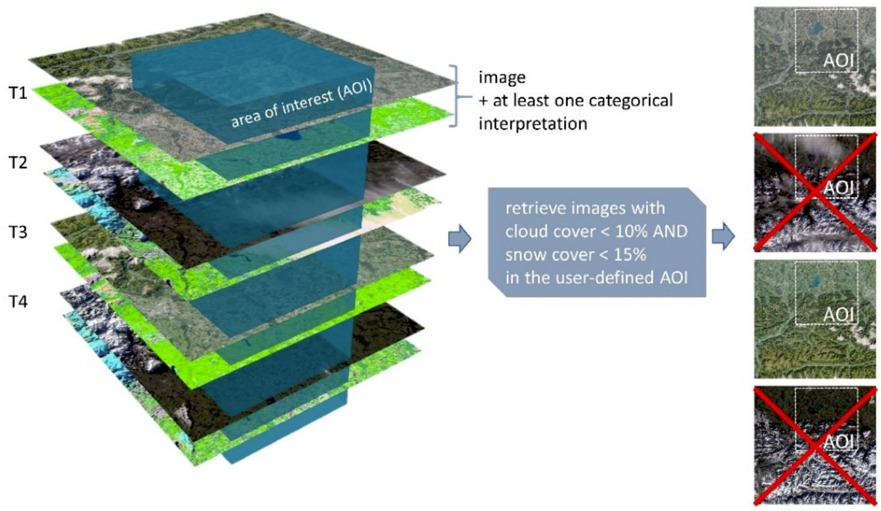

Figure 3. Earth Observation Data Cube architecture for semantic integration of multi-temporal, multi-sensor imagery, supporting user-defined queries and categorical interpretations.

Shape-based and ontology-driven classification outperforms traditional SVM or minimum-distance methods in heterogeneous environments.

3.3 Role of ontologies and data cubes. Ontologies enable:

- Calibration and mapping of heterogeneous sources to a common vocabulary; - Automated classification by semantic reasoning;

- Temporal monitoring of landscape dynamics;

- Scalable fusion for decision support.

VII. CONCLUSION AND PERSPECTIVES

The data heterogeneity and interoperability gap in multi-source remote sensing for landscape monitoring constitutes a fundamental barrier to effective Earth Observation, amplifying global environmental and socio-economic crises - from land degradation and biodiversity loss to climate destabilisation and food insecurity. Analysis of these interconnected threats underscores their profound impact on planetary sustainability.

Proposed solutions - integrating ontologies, semantic web technologies, AI-driven fusion, and EODCs - demonstrate strong potential to bridge the gap. Methods such as OBDA, multimodal transformers, and OGC-standardised cubes achieve high-accuracy integration ($R^2 > 0.77$ in validated models) and support real-time monitoring. These innovations minimise ecological degradation, enhance predictive capabilities, and accelerate achievement of the UN Sustainable Development Goals by enabling equitable, scalable geospatial intelligence.

Ultimately, overcoming data heterogeneity and interoperability challenges through advanced semantic and AI technologies transforms global EO monitoring, providing proactive risk management and strengthening planetary resilience against anthropogenic pressures.

REFERENCES

1. Paving the Way to Increased Interoperability of Earth Observations Data Cubes. *Data* 2019, 4(3), 113. https://doi.org/10.3390/data4030113
2. Use of Semantic Web Technologies to Enhance the Integration and Interoperability of Environmental Geospatial Data. *ISPRS Int. J. Geo-Inf.* 2025, 14(2), 52. https://doi.org/10.3390/ijgi14020052
3. A critical review on multi-sensor and multi-platform remote sensing data fusion approaches. *Int. J. Remote Sens.* 2024. https://doi.org/10.1080/01431161.2024.2429784
4. Four decades of remote sensing for monitoring terrestrial ecosystems. *Remote Sens. Environ.* 2025. https://doi.org/10.1016/j.rse.2025.114147 (adapted)
5. Enhancing 3D geospatial modelling through multimodal data and machine learning. *Spatial Inf. Res.* 2026. https://doi.org/10.1007/s41324-025-00659-4
6. Spatial Data Infrastructure for Remote Sensing: A Comprehensive Analysis. *Preprints* 2024. https://doi.org/10.20944/preprints202404.0593.v1
7. Enhancing Interoperability and Capabilities of Earth Science Data using ODM2. *Data Sci. J.* 2017. https://doi.org/10.5334/dsj-2017-004
8. Multi-sensor integration management in the earth observation sensor web. *Int. J. Appl. Earth Obs. Geoinf.* 2023. https://doi.org/10.1016/j.jag.2023.103425
9. Future challenges and opportunities in data-driven Earth observation. *ResearchGate* 2025.