remote sensing

MDPI

Systematic Review

GeoPROV: A Semantic Provenance Profile for Coastal EO-Derived Products Informed by a Systematic Review of Spatial Data Supply Chains

Muhammad Azeem Sadiq *, Philip Kibet Langat and Arjun Neupane

School of Engineering and Technology, Central Queensland University, Rockhampton, QLD 4701, Australia; p.langat@cqu.edu.au (P.K.L.); a.neupane@cqu.edu.au (A.N.)
* Correspondence: m.sadiq@cqumail.com

Highlights

What are the main findings?

- A systematic review of 156 publications identifies fragmented provenance support across emerging coastal Earth observation (EO) workflows that combine Low-Earth-Orbit and very-high-resolution satellites, unmanned aerial vehicles (UAVs), LiDAR, hyperspectral/thermal sensors, Internet of Things (IoT) observations, and artificial intelligence (AI)-enabled image analysis.
- Existing provenance models provide limited feature-level traceability for EO-derived coastal products used in shoreline mapping and forecasting, UAV/LiDAR-based 3D reconstruction, topo-bathymetry, coastal erosion and flood monitoring, water-quality mapping, habitat classification, and coastal land-use/land-cover analysis.

What are the implications of the main findings?

- GeoPROV provides a lightweight semantic profile that links sensor acquisition, preprocessing, AI/geoprocessing, validation, uncertainty assessment, and STAC/ISO catalogue publication through PROV-O, GeoSPARQL, SHACL, and ISO 19115/19157 alignment.
- Provenance-aware coastal EO supply chains can improve reproducibility, uncertainty traceability, sensor-to-decision accountability, and administrative defensibility for coastal vulnerability assessment, flood disaster response, shoreline management, ecosystem monitoring, and land–sea planning.

Abstract

Emerging remote sensing technologies increasingly support coastal observation through Earth observation (EO) derivatives such as satellite-derived shorelines, unmanned aerial vehicle (UAV)- and light detection and ranging (LiDAR)-based elevation models, coastal land-use/land-cover products, bathymetric grids, turbidity and chlorophyll-a indicators, habitat maps, flood extents, and artificial intelligence (AI)-generated segmentation outputs. These EO-derived products are central to shoreline monitoring, coastal vulnerability assessment, flood-response planning, marine ecosystem management, and land–sea administration. Yet their evidential value depends on transparent provenance: users must know which sensor, acquisition condition, preprocessing chain, algorithm, model version, validation dataset, and uncertainty estimate produced each coastal product. The review synthesises 156 eligible publications identified through a systematic review process. Of these, 89 studies containing sufficient implementation and methodological detail were selected for detailed analytical coding. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol, the review analyses application domains,

Check for updates

Academic Editors: Anselme Muzirafuti and Andrew Clive Banks

Received: 19 August 2026
Revised: 18 October 2026
Accepted: 21 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

provenance capture modes, representation strategies, semantic web technologies, and interoperability practices. A coastal EO thematic refinement translates the review findings into provenance requirements for satellite, UAV, LiDAR, Internet of Things (IoT), open-water, and AI-enabled coastal observation workflows. The review identifies persistent gaps in feature-level lineage, alignment between the World Wide Web Consortium Provenance Ontology (PROV-O) and International Organization for Standardization (ISO) geospatial metadata standards, spatial semantics, AI workflow transparency, uncertainty traceability, and scalable querying. To address these limitations, the paper introduces GeoPROV, a lightweight semantic profile of PROV-O designed for spatially explicit, feature-level provenance. GeoPROV aligns PROV-O with ISO 19115/19157, GeoSPARQL, Shapes Constraint Language (SHACL), and SpatioTemporal Asset Catalogue (STAC)-compatible catalogue practices. The framework enables EO-derived coastal products to be traced from raw observation through preprocessing, AI/geoprocessing, validation, catalogue publication, and coastal-administrative decision-making. The resulting GeoPROV profile integrates PROV-O, GeoSPARQL, SHACL, ISO 19115/19157 lineage concepts, and STAC metadata within a lightweight semantic framework demonstrated through a coastal EO product workflow and machine-actionable provenance implementation.

Keywords: earth observation; coastal remote sensing; EO derivatives; satellite-derived shoreline; UAV LiDAR; AI-enabled remote sensing; spatial data supply chains; provenance; PROV-O; GeoSPARQL

1. Introduction

Spatial data supply chains (SDSCs) capture the end-to-end flow of geospatial data from acquisition to processing, transformation, and dissemination. They operate across heterogeneous organisational arrangements and technological platforms [1–3]. Unlike conventional physical supply chains, SDSCs involve non-depleting digital assets, complex spatial transformations (e.g., reprojection, resampling, generalisation), multi-scale representations, and iterative reuse in analytical and decision-support workflows [4–8]. These characteristics complicate traceability [9], which is critical for ensuring data quality and trust in decision-making processes [10] and is essential for geospatial intelligence [11].

The evolution of spatial data infrastructures [12–14] (SDIs) into service-oriented [15], cloud-native, and platform-based ecosystems has further increased provenance complexity. Modern SDSCs integrate Earth observation platforms, IoT sensor networks [16], AI-driven analytics, and policy-driven governance frameworks, creating intricate webs of data dependencies that traditional metadata-centric lineage models struggle to capture [17–19]. Provenance in this context is no longer solely a documentation artefact; it has become essential for automated quality assessment, reproducibility, accountability, regulatory compliance, and trust.

Existing research has explored provenance interoperability between ISO lineage standards, W3C PROV, workflow provenance systems, and linked-data environments. However, there remains limited guidance on how these standards can be integrated into a lightweight semantic profile specifically designed for EO-derived coastal products. GeoPROV addresses this implementation-level gap by combining provenance, spatial, demonstrated, and EO metadata standards [20] within a single application-oriented profile.

This review primarily investigates provenance frameworks within Spatial Data Supply Chains (SDSCs) across multiple geospatial application domains. Coastal Earth Observation (EO) is subsequently used as a representative application domain to illustrate provenance requirements and demonstrate the practical applicability of the proposed

GeoPROV framework. The review investigates provenance across SDSCs broadly, while coastal EO serves as an application domain used to derive and demonstrate domain-specific provenance requirements.

The study addresses these challenges through a systematic review of semantic web-based provenance approaches for SDSCs. Beyond synthesising existing work, the review identifies recurring limitations and motivates GeoPROV, a lightweight semantic profile that extends PROV-O with spatial semantics while maintaining standards alignment and practical deployability. The investigation makes three contributions: (i) a comprehensive SDSC-focused synthesis of semantic web provenance approaches; (ii) an evidence-based analysis of constraints limiting operational provenance adoption; and (iii) the introduction of GeoPROV as a minimal, interoperable profile designed to lower barriers to spatial provenance implementation. Collectively, these contributions provide a practical and theoretically grounded pathway for advancing scalable, standards-compliant provenance in geospatial systems.

Coastal observation provides a particularly useful lens through which to examine provenance in SDSCs because decision-makers rarely act on raw satellite scenes, UAV images, LiDAR point clouds, or IoT measurements directly. Instead, they rely on EO derivatives: shoreline vectors, erosion-rate transects, coastal digital elevation models, bathymetric grids, turbidity and chlorophyll-a surfaces, habitat classifications, inundation extents, infrastructure exposure layers, and administrative risk zones. Each derivative is shaped by sensor characteristics, spatial resolution, acquisition timing, tidal stage, atmospheric or water-column correction, cloud and shadow masking, classification thresholds, AI model versions, validation data, and uncertainty estimates [21]. In coastal environments, these dependencies are intensified by the land–sea interface itself: the mapped "shoreline" may represent an instantaneous wet/dry boundary, a tidal datum, a vegetation edge, a geomorphic scarp, or an administrative boundary. Provenance therefore becomes more than technical metadata [22]; it becomes the evidential chain that allows coastal observations to be interpreted, compared, audited, and defended in planning, risk management, environmental regulation, and land–sea governance. The central research question guiding this review is: how can semantic web provenance models support transparent, interoperable, feature-level, and uncertainty-aware traceability for EO-derived coastal observation products generated from satellite, UAV, LiDAR, IoT, and AI-enabled remote sensing workflows?

This framing directly responds to the growing use of emerging remote sensing technologies in coastal observation. Recent coastal workflows increasingly combine Low-Earth-Orbit and very-high-resolution satellite imagery, UAV photogrammetry, UAV and airborne LiDAR, hyperspectral and thermal sensing, IoT-based coastal monitoring, open-source tools such as DSAS, QGIS, and Google Earth Engine, and AI-enabled image analysis for shoreline extraction, land-cover mapping, coastal-risk assessment, and flood-response applications. However, the decision value of these technologies depends not only on spatial resolution or algorithmic accuracy but also on whether their derived products can be traced, demonstrated, queried, and reused across scientific and administrative contexts. GeoPROV therefore addresses an under-documented but critical layer of emerging coastal remote sensing: machine-actionable provenance for EO-derived coastal products.

While the review focuses on provenance management for coastal EO-derived products, the evidence base extends beyond coastal applications to include broader geospatial provenance, spatial data infrastructure, and semantic web studies. These studies are examined because many provenance challenges affecting coastal EO-derived products, particularly interoperability, lineage traceability, workflow documentation, and uncertainty communication, are shared across geospatial data ecosystems. Consequently, findings from the

broader spatial data supply chain literature are synthesised to inform a coastal EO-specific provenance profile.

The novelty of this paper is not the introduction of the GeoPROV ontology itself, which has been reported previously, but the derivation of an EO-specific provenance profile from a systematic synthesis of provenance requirements across SDSCs and coastal EO workflows.

2. Materials and Methods

2.1. Review Protocol and Method

A systematic literature review was undertaken following the PRISMA guidelines to ensure transparency, rigour, and reproducibility. The review focused on peer-reviewed journal articles, conference papers, and book chapters addressing provenance modelling in geospatial or spatial data supply chain contexts using semantic web technologies. Searches were conducted across Scopus, Web of Science, IEEE Xplore, and SpringerLink, covering publications from 2001 to 2025. Search terms combined concepts related to provenance, lineage, semantic web, RDF, PROV, and geospatial data infrastructures. To strengthen the relevance of the synthesis for coastal remote sensing, the included EO-related studies were also interpreted through a secondary thematic lens focused on coastal observation. This refinement did not alter the PRISMA inclusion count; rather, it classified relevant studies according to: (i) EO derivative type, including satellite-derived shorelines, coastal DEMs/DSMs [23], bathymetry, water-quality indicators, flood extents, land-use/land-cover products, and habitat maps; (ii) enabling observation technology, including optical and SAR satellites, UAV photogrammetry, UAV and airborne LiDAR, hyperspectral and thermal sensing, IoT and in situ sensors, cloud-computing platforms, and AI-enabled image analysis; and (iii) decision context, including coastal erosion monitoring, shoreline forecasting, emergency response, marine ecosystem assessment, infrastructure exposure, cadastral or administrative boundary updating, and coastal-zone planning. This thematic refinement enabled the general SDSC provenance findings to be translated into coastal observation requirements without representing the review as a separate coastal remote sensing meta-analysis.

The core search string was adapted to the syntax of each database and followed the general structure:

(“provenance” OR “lineage” OR “traceability” OR “data provenance”)
AND
(“Semantic Web” OR RDF OR OWL OR SPARQL OR “PROV-O” OR “W3C PROV”)
AND
(“geospatial” OR “spatial data” OR “Earth observation” OR “remote sensing” OR “spatial data infrastructure” OR “geoprocessing workflow”)

Following the broader SDSC review, a thematic refinement was undertaken to identify studies relevant to coastal Earth Observation workflows. This refinement was intended to provide an illustrative application context rather than a comprehensive systematic review of coastal EO systems.

The literature search covered Scopus, Web of Science Core Collection, IEEE Xplore, and SpringerLink. Publication coverage was restricted to studies published between 2001 and 2025. The search strategy combined three groups of terms representing provenance, semantic web technologies, and geospatial applications. Synonyms within each group were combined using the Boolean operator OR, and the three groups were linked using AND. The final searches were conducted from 24–30 January 2025 (Scopus), from 26 January–10 February 2025 (Web of Science Core Collection), from 28 January–5 February 2025 (IEEE Xplore), and on 25 February 2025 and revisited in 20 August 2025 (SpingerLink). Searches were limited to peer-reviewed journal articles published in English. The exact search strings,

database editions, applied filters, retrieval counts, and any update searches are provided in Supplementary Table S1.

For Scopus, the strategy was formulated for the title, abstract and keyword fields using the expression: TITLE-ABS-KEY((“provenance” OR “lineage” OR “traceability” OR “data provenance”) AND (“Semantic Web” OR “RDF” OR “OWL” OR “SPARQL” OR “PROV-O” OR “W3C PROV”) AND (“geospatial” OR “spatial data” OR “Earth observation” OR “remote sensing” OR “spatial data infrastructure” OR “geoprocessing workflow”)) AND PUBYEAR > 2000 AND PUBYEAR < 2026.

For the Web of Science Core Collection, the strategy was formulated for the Topic field, encompassing titles, abstracts, author keywords and Keywords Plus, using the expression: TS = ((“provenance” OR “lineage” OR “traceability” OR “data provenance”) AND (“Semantic Web” OR “RDF” OR “OWL” OR “SPARQL” OR “PROV-O” OR “W3C PROV”) AND (“geospatial” OR “spatial data” OR “Earth observation” OR “remote sensing” OR “spatial data infrastructure” OR “geoprocessing workflow”)) AND PY = (2001–2025). The indexes searched were the Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), and Emerging Sources Citation Index (ESCI).

For IEEE Xplore, the strategy was formulated for Command Search using the All Metadata field: (“All Metadata”:“provenance” OR “All Metadata”:“lineage” OR “All Metadata”:“traceability” OR “All Metadata”:“data provenance”) AND (“All Metadata”:“Semantic Web” OR “All Metadata”:“RDF” OR “All Metadata”:“OWL” OR “All Metadata”:“SPARQL” OR “All Metadata”:“PROV-O” OR “All Metadata”:“W3C PROV”) AND (“All Metadata”:“geospatial” OR “All Metadata”:“spatial data” OR “All Metadata”:“Earth observation” OR “All Metadata”:“remote sensing” OR “All Metadata”:“spatial data infrastructure” OR “All Metadata”:“geoprocessing workflow”). The publication-year filter (2001–2025) was applied through the IEEE Xplore advanced search interface. For SpringerLink, the strategy was formulated for the advanced search Keywords field using the expression: (“provenance” OR “lineage” OR “traceability” OR “data provenance”) AND (“Semantic Web” OR “RDF” OR “OWL” OR “SPARQL” OR “PROV-O” OR “W3C PROV”) AND (“geospatial” OR “spatial data” OR “Earth observation” OR “remote sensing” OR “spatial data infrastructure” OR “geoprocessing workflow”). This field searches titles, abstracts and body text, providing broader textual coverage than the metadata fields used in the other databases. The publication date filter (1 January 2001–31 December 2025) was applied during database searching. English-language and publication-type restrictions were applied during database searching and verified during eligibility screening.

Studies were excluded if they were non-English, lacked explicit provenance representations, presented purely conceptual discussions without methodological contributions, or were duplicate or substantially redundant. From an initial set of 734 records, 156 studies met the inclusion criteria (Figure 1).

GeoPROV was derived through a PRISMA-aligned, evidence-driven process, synthesising recurring requirements and identifying gaps in granularity, interoperability, and standards alignment, leading to a minimal, standards-compliant semantic profile demonstrated against representative SDSC use cases.

The review included two analytical stages. First, all eligible studies contributed to the qualitative synthesis examining provenance concepts, standards, and implementation approaches. Second, a subset of studies containing sufficient methodological and implementation detail were selected for quantitative coding and thematic analysis. This distinction ensured consistent comparison across studies while preserving the broader contextual evidence base.

The systematic review yielded 156 eligible publications. A subset of 89 studies were selected for detailed coding because these publications provided sufficient information

regarding provenance architecture, semantic technologies, workflow implementation, metadata models, or interoperability mechanisms.

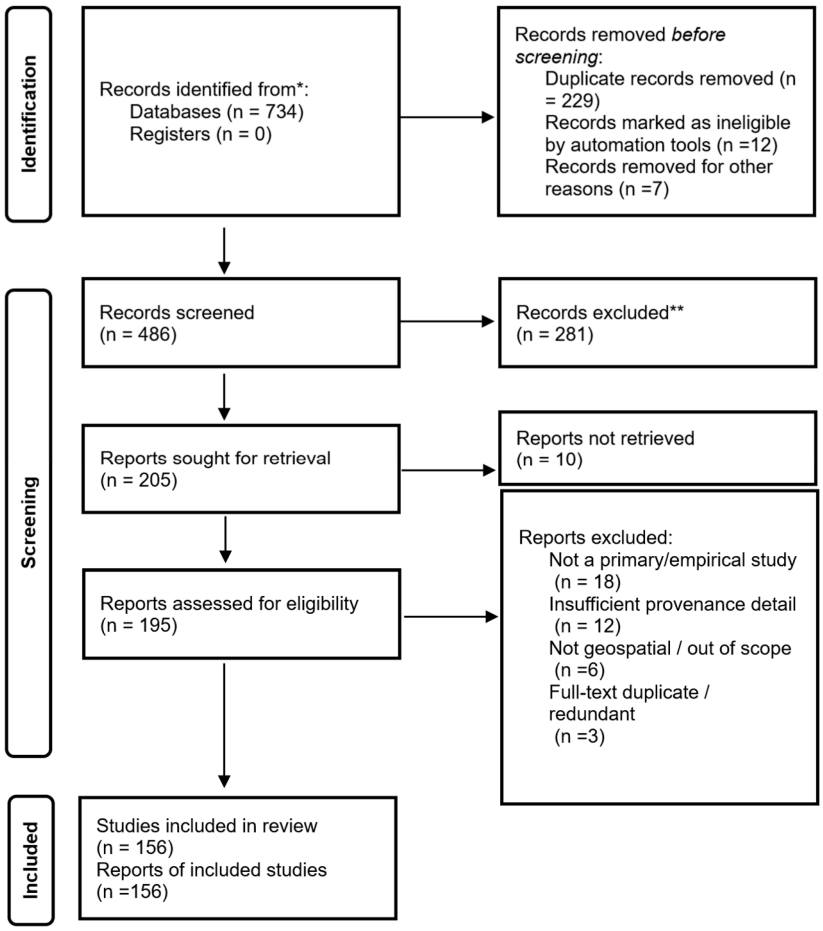


Figure 1. PRISMA flow diagram illustrating the literature search, screening, eligibility assessment, and final selection of studies included in the systematic review. * Records were identified from Scopus, Web of Science Core Collection, IEEE Xplore, and SpringerLink, yielding 734 records before duplicate removal and other exclusions. ** All 281 records excluded during screening were excluded by human reviewers. The 12 automated exclusions occurred before screening and are reported separately.

2.2. Literature Search Strategy

This review followed the PRISMA 2020 reporting guidelines. The review protocol was not published, and the review was not registered. Literature research was conducted in Scopus, Web of Science Core Collection, IEEE Xplore, and SpringerLink between January and March 2025 and revisited from December 2025 to February 2026. Database-specific search strings were developed to account for differences in indexing structures and search syntax. Searches were independently executed by two reviewers, and all retrieved records were imported into EndNote for duplicate removal.

Title and abstract screening was conducted independently by two reviewers. Any discrepancies were resolved through discussion and consensus. Full-text eligibility assessment followed predefined inclusion and exclusion criteria. Data extraction was independently performed using a structured coding framework that captured study objectives, provenance standards, implementation approaches, application domains, validation methods, and reported challenges.

To ensure transparency and reproducibility, database-specific search strategies, screening protocols, exclusion reasons, and coding procedures are provided in Supplementary Materials.

Title/abstract screening and full-text eligibility assessment were independently conducted by two reviewers, with disagreements resolved through discussion and consensus.

A qualitative quality assessment was conducted to evaluate the strength and relevance of the evidence base as described in Table 1. Studies were appraised against five dimensions: (i) methodological transparency [24,25], (ii) provenance implementation maturity [26], (iii) validation evidence [27–29], (iv) interoperability [15] contribution, and (v) application relevance [15]. Rather than assigning numerical scores, studies were classified as having high, moderate, or emerging evidence strength based on the extent of implementation, validation, and practical applicability reported. The assessment indicated that the literature is relatively mature in provenance conceptualisation and interoperability standards but less mature in operational implementation and formal validation [30]. This gap was particularly evident for coastal EO applications, where provenance requirements are often discussed conceptually rather than implemented through machine-actionable frameworks [31–33]. The quality appraisal informed the identification of research gaps and the design requirements of the proposed GeoPROV profile.

Table 1. Coding framework used to classify studies in the systematic review.

Coding Dimension	Categories/Examples	Purpose
Application domain	Earth observation, coastal observation, cadastre, smart cities, hydrology, digital twins, environmental monitoring	Identifies the operational context in which provenance is applied
Provenance capture mode	Embedded capture, service-oriented capture, post hoc reconstruction, hybrid capture	Classifies how provenance information is generated
Representation strategy	PROV-O, ISO 19115/19157 [34] lineage, custom ontology, hybrid model, workflow model	Identifies the conceptual or formal provenance model used
Semantic web technologies	RDF, OWL, SPARQL, GeoSPARQL, SHACL, JSON-LD	Identifies the machine-actionable technologies used
Granularity	Dataset-level, workflow-level, feature-level, attribute-level, parameter-level	Assesses the detail at which provenance is represented
Standards alignment	PROV, ISO 19115/19157, OGC services, STAC, DCAT/GeoDCAT-AP	Assesses interoperability and standards compliance
Coastal EO thematic refinement	Satellite-derived shoreline, UAV/LiDAR DEM, bathymetry, turbidity, habitat map, flood extent, coastal-administrative layer	Translates general SDSC provenance findings into coastal observation requirements

2.3. Coastal EO Thematic Refinement

Because this review targets provenance requirements in remote sensing and coastal observation, a secondary thematic refinement was applied to the EO-related subset of the included studies. This refinement classified studies and the supporting coastal EO literature according to four dimensions: observation technology, EO derivative, coastal process, and decision context. Observation technologies included optical and SAR satellites, very-high-resolution commercial satellite imagery, UAV photogrammetry, UAV and airborne LiDAR, hyperspectral and thermal sensing, IoT and in situ sensors, cloud-native processing platforms, and AI-enabled image analysis. EO derivatives included shoreline vectors, coastal DEMs/DSMs, bathymetric grids, turbidity and chlorophyll-a products, flood extents, land-use/land-cover maps, habitat classifications, and administrative risk zones. Coastal

processes included erosion, accretion, flooding, subsidence, sea-water intrusion, open-water dynamics, wetland change, and shoreline migration. Decision contexts included coastal vulnerability assessment, emergency response, infrastructure exposure, ecosystem monitoring, development control, and land–sea administration.

This refinement did not replace the PRISMA-based provenance review. Instead, it translated the general SDSC provenance evidence into coastal observation requirements, enabling GeoPROV to be evaluated as a standards-aligned provenance profile for EO-derived coastal products. The complete ontology, SHACL shapes, event-contract schema, representative tool-specific fixtures, competency queries, evaluation scripts, and benchmark outputs are provided in the Supplementary Materials (Tables S1–S4).

3. Results

3.1. SDSCs and Provenance Requirements

SDSCs involve multiple interacting actors such as data producers, processors, custodians, and users; each of whom may introduce transformations, uncertainty, or reinterpretation [24]. Unlike linear data pipelines, SDSCs exhibit networked and iterative flows, where outputs are repeatedly repurposed, combined, and re-analysed across organisational and technological boundaries. As a result, provenance in SDSCs [2] must capture multi-actor, multi-stage, and multi-granularity lineage rather than coarse, dataset-level histories. Across the reviewed literature, a consistent set of provenance requirements emerges. These include support for feature- and attribute-level traceability, explicit representation of spatial transformations and associated parameters, alignment with data quality and governance frameworks, and machine-actionable querying and reasoning capabilities. These requirements, and the extent to which they are addressed across application domains, are summarised in Table 2.

Recent work has extended generic provenance frameworks through GeoPROV, a domain-specialised provenance model designed specifically for Spatial Data Supply Chains. GeoPROV integrates geospatial entities, workflow activities, and provenance relationships while supporting interoperability with established standards such as PROV-O, GeoSPARQL, and RDF. GeoPROV was previously introduced and evaluated as a provenance-centric framework and infrastructure for geospatial data management and interoperability [35,36]. While those studies focused on the design, implementation, and validation of the GeoPROV framework and ontology, the present paper has a broader scope. This study provides a systematic literature review of geospatial data provenance research, synthesising existing approaches, technologies, standards, challenges, and research trends across the field. GeoPROV is included as one of the analysed provenance frameworks and is evaluated alongside other relevant approaches identified in the literature. Consequently, the contribution of this paper is not the development of GeoPROV itself, but rather a comprehensive and systematic assessment of the geospatial provenance landscape that contextualises and positions GeoPROV within the broader state of the art.

The quality appraisal indicated that most studies demonstrated moderate to high methodological transparency and strong interoperability contributions. However, comparatively fewer studies reported operational implementations, machine-actionable provenance artefacts, or formal validation procedures. This finding highlights a persistent gap between conceptual provenance frameworks and deployable provenance solutions within geospatial and EO workflows.

Table 2. Core semantic technologies, standards, and provenance capabilities identified in reviewed SDSC studies.

Technology/Standard	Primary Purpose	Provenance Function	Spatial Capability	Typical SDSC Applications	Representative References
PROV-O	W3C provenance ontology	Represents entities, activities, and agents involved in data creation and transformation	Limited (requires extensions)	Workflow lineage, audit trails, reproducibility	[2,3,35–38]
RDF	Resource Description Framework	Structured representation of provenance relationships as triples	Indirect	Knowledge graphs, metadata integration	[21,39,40]
OWL	Web Ontology Language	Formal ontology definition and reasoning	Indirect	Semantic modelling, interoperability	[15,35,41–44]
SPARQL	RDF query language	Querying provenance graphs and metadata repositories	Limited	Provenance search, lineage tracing	[45,46]
GeoSPARQL	Geospatial semantic standard	Enables spatial relationships and geospatial querying	High	Spatial provenance, location-aware lineage	[15,42,47]
SKOS	Knowledge Organisation System	Controlled vocabularies and semantic classification	Limited	Metadata harmonisation	[28,42]
DCAT	Data Catalogue Vocabulary	Dataset discovery and cataloguing	Limited	Data portals and interoperability	[48,49]
Dublin Core	Metadata standard	Resource description and metadata management	Limited	Dataset documentation	[25,50]
ISO 19115	Geographic metadata standard	Dataset lineage and quality documentation	High	Geospatial metadata management	[38,51]
ISO 19157	Geographic data quality standard	Quality reporting and lineage documentation	High	Data quality assessment	[52]
OGC Standards	Spatial interoperability framework	Supports exchange and discovery of geospatial resources	High	SDIs and EO platforms	[49,53]
STAC	SpatioTemporal Asset Catalogue	EO asset discovery and metadata interoperability	High	Earth Observation supply chains	[43]
SensorML	Sensor description framework	Documents sensor provenance and acquisition parameters	High	EO, UAV and IoT workflows	[39]
SOSA/SSN	Semantic sensor ontology	Models sensors, observations and sampling activities	High	Environmental monitoring and EO workflows	[54,55]
Knowledge Graphs	Semantic integration framework	Links provenance, metadata and domain knowledge	High	Multi-source SDSC integration	[36,56,57]
Linked Data	Data publication framework	Persistent linkage between datasets and provenance records	Moderate	Interoperable geospatial ecosystems	[58,59]

3.2. Application Domains

Concepts from traditional supply chain management (SCM) provide a useful starting point for reasoning about spatial data flows but require substantial adaptation to reflect the distinctive properties of geospatial information [60]. Unlike physical goods, spatial data are non-depleting resources whose provenance documents data origin, ownership, processing history, analytical transformations, and usage contexts. Although provenance concepts are also applied in domains such as palaeontology, food supply chains, and forensic science, the focus of this study is semantic and computational provenance for digital geospatial information systems.

Within this context, SDSCs exhibit requirements that extend beyond conventional metadata-centric lineage models. Geospatial processing routinely introduces explicit spatial transformations (e.g., reprojection, resampling, generalisation), temporal evolution through versioning, updates, and aggregation, and semantic enrichment operations such as classification, feature extraction, and spatial reasoning [34,61–64]. Each of these operations directly affects recognised data quality dimensions such as positional accuracy, thematic accuracy, completeness, logical consistency, and lineage, thus requiring provenance to be captured at the point where quality is created, assessed, or modified.

Across application domains such as environmental monitoring, transport planning, mining, smart cities, and emergency management, SDSCs involve multiple actors—data collectors, processors, analysts, custodians, publishers, and end-users—each introducing value, uncertainty, or risk. Governance in these settings is inherently distributed, with differing roles, responsibilities, and authority, creating complex trust relationships that must be made explicit and verifiable [34,38]. As synthesised in Table 3, effective operational governance depends on the ability to associate quality diagnostics (e.g., accuracy metrics, validation results, completeness indicators) with the specific processing steps, agents, and reference data that produced them, and to store these elements in appropriate standards-compliant locations.

Table 3. Data quality elements, practical metrics, and recommended Storage Locations.

Quality Element	Typical Metrics/Diagnostics	Capture Point	Store in...
Positional accuracy [65]	RMSE vs. reference, offset stats, checkpoints	Transform/QA step	ISO 19157 quality section; PROV note on reference used; QA tables
Attribute (thematic) accuracy [66]	Confusion matrix, precision/recall	Transform/QA step	ISO 19157 measures; PROV entity for the validation dataset.
Logical consistency	Topology errors, domain checks	Ingest/Transform	ISO 19157 logical checks; graph constraints
Completeness	Missing features/attributes, coverage	Capture/Ingest	ISO 19157 completeness; cube/tiling coverage metadata
Lineage	Sources, steps, parameters, agents, dates	All steps	ISO 19115-2 lineage; PROV-O graph

However, the literature summarised in Table 3 indicates that while semantic provenance approaches increasingly capture high-level lineage, they often fail to integrate feature-level provenance with data quality elements across the full SDSC lifecycle. This gap highlights a fundamental misalignment between how geospatial data are processed in practice and how provenance is represented. Addressing this misalignment requires a lightweight, interoperable provenance profile that explicitly links spatial processing activities, quality assessments, and standards-based metadata—motivating the design of GeoPROV as an enabling layer between operational SDSCs and existing ISO- and PROV-based frameworks.

3.3. Provenance Concepts and Standards

Provenance in spatial data encompasses both retrospective provenance, documenting what has occurred, and prospective provenance [67], describing intended processes across the data lifecycle [9,68] (documenting what should happen) across the entire data lifecycle. It has evolved from basic audit trails into semantic representations that support machine-actionable reasoning about data quality, trustworthiness, and fitness for purpose, elevating

data from passive inputs to strategic assets [69–71]. In domains such as autonomous vehicles, underground utilities, and geofencing, high-stakes decisions increasingly depend on traceable, machine-readable spatial context [72,73]. Central to this capability is data lineage, which records data origins, transformations, and dissemination pathways [6,74].

Versioning is essential for reproducibility and auditability, as sequential dataset histories enable reliable re-analysis [75]. Large-scale operational systems such as Digital Earth Australia [76] highlight the need for feature-level versioning [77] with persistent identifiers [78] to manage algorithmic updates and complex data interdependencies. Experience from DEA [79] further emphasises the importance of explicitly capturing ingest and processing properties—including coordinate reference systems, resolution, tiling, compression, and metadata conventions—to stabilise downstream analyses and ensure robust provenance chains [80].

The W3C PROV standard dominates provenance representation, providing a flexible entity–activity–agent model adopted across numerous geospatial and scientific domains [8,38,81,82]. However, PROV lacks native geospatial semantics, necessitating integration with spatial standards and domain vocabularies. As data volumes and processing complexity increase, semantic web-based provenance modelling has become essential for SDSCs, from early sensor web experiments to ontology-driven interoperability frameworks [83–86]. Despite this progress, challenges persist in granular capture, dynamic workflows, and semantic integration [3].

A coherent framework linking SDSCs, provenance metadata, and semantic web technologies—emphasising ontology-driven representation and querying—is therefore critical [87–89]. Provenance plays a foundational role in transparent, interoperable workflows and contextual quality assessment [90]. ISO 19157-1:2023 defines core data quality elements such as positional accuracy, thematic accuracy, logical consistency, completeness, and lineage, each of which requires systematic capture at specific processing stages. Table 3 summarises practical metrics, capture points, and storage strategies, illustrating how ISO quality elements can be complemented with PROV-O entities and activities. While prior studies demonstrate semantic integration of PROV-O with ISO lineage [38,91,92], workflow-centric graph models [38,93–95], and scalable big data lineage solutions [51,96–98], gaps remain in proactive capture, intermediate result validation, and domain-specific profiling—motivating the need for standardised, spatially explicit provenance frameworks.

3.4. Evolution of Semantic Provenance Research in Spatial Data

Research on provenance modelling for spatial data supply chains using semantic web technologies has evolved from foundational frameworks to advanced, domain-specific applications. Early studies focused on developing semantic provenance models [87,93], architectures, and standards for integration to support traceability and data quality assessment [13,47], Later work shifted towards geospatial workflows, feature-level granularity, and interoperability challenges, leveraging ontologies, linked data, and semantic execution engines [99]. Recent developments incorporate big data integration, automated reasoning, and practical implementations across specific domains, underscoring the role of provenance in trust [100], interoperability [82], and data governance [25,98].

The reviewed literature consistently highlights the importance of semantic web technologies for enhancing provenance modelling in spatial data supply chains, particularly through extensions to standards such as W3C PROV [32,81,82,90,101–103] and ISO geospatial metadata [34,62,104,105]. There is broad consensus on the need to capture provenance at multiple levels of granularity, including feature and attribute levels, to ensure data quality and trustworthiness. However, differences emerge regarding the degree of integration with automated geospatial workflows and the sophistication of querying and

reasoning capabilities. These variations reflect differences in domain focus, technical scope, and implementation maturity. Interoperability challenges and inconsistencies in standard compliance also persist, often due to heterogeneous application contexts and evolving semantic frameworks.

3.5. Semantic Web Technologies

The semantic web technology stack provides the foundational technologies for representing, storing, and querying provenance information in machine-readable formats [81,106–109]. RDF serves as the basic data model for expressing provenance relationships, while OWL enables the definition of domain-specific vocabularies and reasoning rules [96,110–112]. SPARQL Protocol and RDF Query Language (SPARQL) provides the standard query interface for semantic provenance repositories [62,104].

The application of semantic web technologies to spatial data provenance presents both opportunities and challenges [113]. On one hand, the flexibility of RDF enables the representation of complex spatial relationships and transformations that are difficult to capture in traditional relational models. On the other hand, the verbosity of RDF can lead to performance challenges when dealing with large-scale spatial datasets, requiring careful optimisation of storage and query strategies. Table 4 shows that most provenance studies rely on the classic semantic web stack: RDF (88.8%) for representation, SPARQL (80.9%) for querying, and OWL (71.9%) for ontologies and reasoning. Specialised technologies appear less frequently but serve specific roles: GeoSPARQL (33.7%) for spatial queries, SHACL (20.2%) for validation, and JSON-LD (28.1%) for web-friendly RDF serialisation. Overall, RDF, SPARQL, and OWL form the core, while GeoSPARQL, SHACL, and JSON-LD address spatial, validation, and integration needs.

Table 4. Semantic technology adoption patterns.

Technology	Studies (n)	Percentage	Primary Use Cases
RDF	79	88.8%	Basic provenance representation
SPARQL	72	80.9%	Provenance querying and discovery
OWL	64	71.9%	Ontology definition and reasoning
GeoSPARQL	30	33.7%	Spatial queries over provenance
SHACL	18	20.2%	Provenance validation and constraints
JSON-LD	25	28.1%	Web-friendly RDF serialisation

Note: Percentages are calculated against the analytically coded implementation subset ($n = 89$) because not all 156 included publications reported sufficient technical detail for this coding category. Categories are non-exclusive where studies used multiple technologies, standards, or capabilities.

RDF, OWL, and SPARQL form the dominant technology stack for provenance representation and querying. GeoSPARQL supports spatial predicates, while SHACL enables constraint validation. Despite demonstrated benefits, adoption of reasoning and validation technologies remains limited due to performance and complexity concerns.

3.6. Existing Spatial Provenance Models

Spatial provenance modelling is anchored in three key standards: W3C PROV, ISO lineage and quality standards, and GeoSPARQL. The W3C PROV family provides a technology-neutral model for representing provenance on the Web, built around core concepts—entity, activity, and agent, and relations such as prov:used, prov:wasGeneratedBy, and prov:wasDerivedFrom [38,114,115]. PROV-O, the OWL/RDF ontology, forms the backbone for semantic provenance and is extended by GeoPROV to incorporate geospatial entities, operations, and parameters for SDSCs. ISO 19115 and ISO 19157 complement PROV by defining

metadata and data quality standards. ISO 19115 specifies lineage elements (LI_Lineage, LI_ProcessStep, LI_Source) for catalogue-level documentation, while ISO 19157 formalises quality measures such as completeness, positional accuracy, and logical consistency. These standards serve as the catalogue-facing layer, while PROV-O and GeoSPARQL enable machine-actionable provenance graphs. GeoSPARQL extends RDF/SPARQL with spatial vocabularies and topological functions [116], allowing queries that combine provenance paths with spatial predicates. Integrating these standards addresses heterogeneity in schemas and CRSs but remains challenged by tooling maturity and feature-level granularity at scale [34,38,91,92,117]. Table 5 illustrates the distribution of provenance models and frameworks adopted in spatial data supply chain research. It compares five categories: W3C PROV/PROV-O, ISO 19115/19157 lineage standards, custom domain ontologies (e.g., GeoPROV, BioProv), the open provenance model (OPM), and hybrid approaches combining multiple standards. The visualisation highlights W3C PROV as the dominant model, used in 67.4% of studies ($n = 60$), followed by ISO lineage standards at 44.9% ($n = 40$). Custom ontologies account for 22.5% ($n = 20$), while OPM and hybrid approaches represent 16.9% ($n = 15$) and 13.5% ($n = 12$), respectively. This figure emphasises the trend toward semantic interoperability, with PROV-based models forming the backbone of most implementations and ISO standards providing complementary, catalogue-lineage and quality structures.

Table 5. Adoption of provenance models and frameworks in SDSC.

Model/Framework	Studies (n)	Percentage	Key Characteristics
W3C PROV/PROV-O	60	67.4%	Generic provenance vocabulary, extensible
ISO 19115/19157 Lineage	40	44.9%	Geospatial-specific metadata standards
Custom Domain Ontologies	20	22.5%	Specialised vocabularies (GeoPROV, BioProv)
Open Provenance Model (OPM)	15	16.9%	Legacy provenance model, mapped to PROV
Hybrid Approaches	12	13.5%	Combination of multiple standards

Note: Percentages are calculated against the analytically coded implementation subset ($n = 89$) because not all 156 included publications reported sufficient technical detail for this coding category. Categories are non-exclusive where studies used multiple technologies, standards, or capabilities.

3.7. Provenance Requirements in Coastal Remote Sensing Workflows

Coastal remote sensing workflows are provenance-intensive [5–7,16,24,38,84,90,118–126] because they transform heterogeneous observations into EO-derived products that support scientific interpretation, risk assessment, and administrative decision-making [2,3]. Unlike static geospatial datasets, coastal EO products are shaped by rapidly changing environmental conditions, including tide, wave run-up, storm surge, turbidity, sediment transport, vegetation seasonality, and anthropogenic modification. A shoreline, flood extent, bathymetric surface, habitat polygon, or coastal-risk boundary is therefore not merely a mapped object; it is the outcome of a sensor-to-decision chain involving acquisition, preprocessing, model selection, parameterisation, validation, uncertainty assessment, catalogue publication, and downstream reuse [21,39,42,127].

For this reason, provenance in coastal remote sensing must capture more than dataset-level lineage. It must document the source sensor and platform, acquisition time, spatial and spectral resolution, preprocessing method, algorithm or AI model, training and validation data, environmental context, quality metrics, spatial reference system, processing software, responsible agent, and version history [128–131]. The following subsections translate these requirements into five major coastal remote sensing workflow categories.

3.7.1. Satellite-Derived Shorelines and Coastal Landform Monitoring

Satellite-derived shoreline and coastline products are among the most widely used EO derivatives in coastal observation [132]. They are commonly generated from medium-resolution satellite archives such as Landsat and Sentinel-2, as well as from very-high-resolution commercial imagery such as PlanetScope, WorldView, and GeoEye. These products support shoreline-change detection, erosion and accretion monitoring, coastal landform mapping, beach-state analysis, dune and cliff retreat assessment, and long-term coastal vulnerability studies [133]. However, the derived shoreline is not a neutral representation of the coast. It depends on the selected shoreline proxy, the image date and tidal stage, atmospheric correction, cloud and shadow masking, water-index formulation, thresholding method, spatial resolution, and validation source [21,39,42,127].

Provenance capture is therefore essential for interpreting satellite-derived shorelines across time and across administrative contexts. A shoreline extracted using NDWI, MNDWI, AI segmentation, or manual digitisation may represent different physical boundaries: an instantaneous wet/dry line, a vegetation edge, a tidal datum, a geomorphic scarp, or an administratively defined coastline [8,96,134–136]. GeoPROV should record the source scene identifier, acquisition timestamp, tide or water-level reference, shoreline proxy definition, preprocessing chain, extraction algorithm, threshold values, transect baseline, coordinate reference system, validation dataset, positional uncertainty, and output feature version. This enables users to query not only where a shoreline segment is located but also how it was produced, which environmental conditions influenced it, and whether it is comparable with earlier or later shoreline products.

For coastal landform monitoring, the same provenance logic applies to beaches, beach berms, beach scarps, sand spits, tidal flats, lagoons, estuaries, rocky cliffs, salt flats, dunes, marshes, and coastal artificial structures such as groynes, seawalls, dikes, jetties, and piers [128]. These features are often mapped from multi-temporal satellite imagery and then reused in geomorphological interpretation, coastal-risk mapping, and planning assessments. Provenance-aware shoreline and landform products therefore improve reproducibility, change-detection reliability, and the administrative defensibility of coastal monitoring outputs.

3.7.2. UAV Photogrammetry, LiDAR, and High-Resolution 3D Coastal Reconstruction

UAV photogrammetry and LiDAR have become central technologies for high-resolution coastal observation because they can generate centimetre-scale orthomosaics, dense point clouds, digital surface models, digital elevation models, beach-volume estimates, dune morphology models, cliff retreat surfaces, and post-storm damage assessments [137,138]. These products are especially valuable where satellite imagery lacks sufficient spatial resolution or where rapid event-based monitoring is required after erosion, flooding, storm surge, or infrastructure damage. UAV-based Structure-from-Motion and Multi-View Stereo workflows, however, introduce many processing decisions that directly affect positional accuracy, vertical accuracy, and feature interpretation.

A provenance-aware UAV workflow should capture flight plan, platform type, sensor model, camera calibration, flight altitude, image overlap, acquisition date and time, weather condition, ground-control-point distribution, RTK/PPK configuration, GNSS/IMU metadata, photogrammetric software, point-cloud generation settings, filtering method, interpolation method, coordinate reference system, vertical datum, and accuracy metrics such as RMSE [5,8,134]. For LiDAR workflows, additional provenance elements include scan angle, pulse density, return classification, boresight calibration, point-cloud classification, ground/non-ground filtering, water-surface correction, and vertical uncertainty. These ele-

ments determine whether a DEM or DSM is suitable for beach-volume analysis, cliff retreat monitoring, coastal engineering, flood modelling, or infrastructure exposure assessment.

Green LiDAR and topo-bathymetric LiDAR add a further provenance challenge because they connect terrestrial, intertidal, and shallow-water environments. In such workflows, depth penetration, water clarity, bottom reflectance, refraction correction, water-level reference, and bathymetric validation must be recorded alongside standard LiDAR metadata. GeoPROV can represent UAV and LiDAR products as spatially typed prov:Entity instances linked to acquisition activities, processing activities, responsible agents, quality assessments, and administrative uses. This allows coastal 3D products to remain traceable from raw imagery or point clouds through to derived terrain models and decision-support layers.

3.7.3. AI-Enabled Coastal Observation and Foundation-Model Provenance

AI-enabled coastal observation is increasingly used for shoreline extraction, land–water segmentation, coastal land-use/land-cover mapping, mangrove and wetland classification, flood-boundary detection, cloud removal, super-resolution, habitat mapping, and change detection. Deep learning models and emerging foundation models can improve automation and transferability across coastal environments, but they also create new provenance risks. Unlike conventional rule-based workflows, AI outputs are shaped by training data selection, annotation rules, class definitions, model architecture, model weights, hyperparameters, fine-tuning strategy, prompt configuration, postprocessing rules, software libraries, and computational environment [139,140].

For this reason, AI-generated coastal EO products require model-level and feature-level provenance. A shoreline extracted by a convolutional neural network, transformer model, or foundation model should be linked to the input image, training dataset, validation dataset, annotation protocol, model version, inference date, confidence score, thresholding rule, postprocessing operation, and uncertainty estimate. Without these details, users cannot determine whether an AI-derived shoreline, flood extent, habitat polygon, or land-cover class is transferable across regions, seasons, sensors, or tidal conditions.

GeoPROV can support AI workflow transparency by representing the AI model as a prov:Agent or software-associated prov:Entity, the inference process as a prov:Activity, and the AI-derived feature as a generated prov:Entity. This allows coastal managers and researchers to query which model generated a feature, which data were used for training and validation, what classification schema was applied, and how uncertainty was assessed. Such provenance is particularly important when AI-derived products inform coastal engineering, emergency response, erosion setbacks, habitat restoration, compliance monitoring, or administrative boundary interpretation [141].

3.7.4. Open-Water Dynamics, Satellite-Derived Bathymetry, and Water-Quality Products

Open-water and nearshore dynamics are central to coastal observation because many coastal products are influenced by water-column conditions, optical complexity, and hydrodynamic variability. EO-derived open-water products [21,28] include turbidity, total suspended matter, chlorophyll-a, coloured dissolved organic matter, sea-surface temperature, algal-bloom indicators, plume extent, water clarity, shallow-water bathymetry, and nearshore morphology [142–144]. These products may be generated from ocean-colour sensors, Sentinel-2 MSI, Sentinel-3 OLCI, Landsat, MODIS, thermal sensors, airborne or UAV hyperspectral data, in situ observations, and satellite-derived bathymetry algorithms [145–147].

The provenance requirements for these products differ from terrestrial mapping workflows. Open-water products require explicit documentation of atmospheric correction, sun-glint correction, BRDF treatment, water-column correction, turbidity condition, opti-

cally deep-water masking, bottom type, depth-control points, in situ matchup data, sensor calibration, vicarious calibration, algorithm version, temporal compositing, and uncertainty estimates [126]. These elements are crucial because coastal waters are optically complex: suspended sediment, chlorophyll, coloured dissolved organic matter, bottom reflectance, depth, tide, waves, and currents can all change the apparent signal recorded by the sensor [69,70,86].

Open-water provenance is also important for shoreline and coastal-risk interpretation. The apparent land–water boundary can shift because of tide, wave run-up, storm surge, sediment plumes, water turbidity, or seasonal hydrodynamics. A provenance model for coastal EO derivatives [16,39,130,131,134] must therefore link marine, intertidal, and terrestrial processing chains rather than treating them as isolated products. GeoPROV supports this requirement by connecting water-level observations, in situ validation, satellite acquisition metadata, processing parameters, uncertainty surfaces, and derived coastal features within a single machine-actionable provenance graph [148]. This improves the reliability of bathymetry, water-quality [21], habitat, flood, and shoreline products used in coastal science and management.

3.7.5. Coastal-Risk, Flood-Response, and Administrative Decision Products

Coastal remote sensing increasingly supports decision products rather than only scientific maps. EO-derived outputs are used for coastal hazard zoning, erosion-risk mapping, flood and storm-surge response, infrastructure exposure assessment, setback planning, insurance modelling, emergency management, ecosystem restoration, blue-carbon accounting, and land–sea administration. These products are often generated by integrating satellite imagery, SAR, UAV surveys, LiDAR DEMs, hydrodynamic models, tide gauges, IoT sensors, cadastral layers, infrastructure datasets, and planning boundaries. Their reliability depends on whether the full evidential chain can be traced from observation to decision.

Flood-response products illustrate the issue clearly. A mapped inundation extent may depend on SAR or optical imagery, event timestamp, tide level, rainfall or river discharge, DEM source and resolution, vertical datum, hydrodynamic model parameters, flood-detection algorithm, thresholding rule, validation observations, and uncertainty surface. Similarly, coastal-risk zones and erosion setbacks may depend on shoreline time-series, beach-profile change, projected sea-level rise, storm-surge assumptions, infrastructure layers, land tenure boundaries, and planning regulations. If these inputs and transformations are not documented, the final decision layer becomes difficult to audit, reproduce, or defend.

GeoPROV links scientific provenance with administrative accountability by representing downstream [35] uses as prov:Activity instances. For example, a shoreline vector may be generated from satellite imagery, demonstrated with UAV or GNSS data, published through a STAC/ISO catalogue, and then reused in DSAS-based erosion-rate analysis, setback planning, or cadastral boundary interpretation. GeoPROV can record each stage, including the responsible agent, input product, processing method, validation evidence, uncertainty measure, and administrative output. In this sense, provenance becomes a governance mechanism: it allows coastal decision-makers to understand not only what a coastal product shows but also how it was generated, whether it is fit for purpose, and how confidently it can be used in planning, regulation, or emergency response.

Taken together, these five workflow categories show why coastal remote sensing requires a provenance model that is spatially explicit, standards-aligned, feature-level, and operationally deployable. The key challenge is not simply to preserve a record of raw data sources but to maintain a machine-actionable chain linking sensor acquisition, preprocessing, AI or geoprocessing activity, derived coastal feature, validation evidence, uncertainty estimate, catalogue record, and downstream administrative use. This requirement directly

motivates GeoPROV as a lightweight semantic profile for provenance-aware coastal EO data supply chains.

3.8. EO-Derived Coastal Products, Enabling Technologies, and Administrative Impact

EO derivatives are the operational currency of coastal observation. They translate sensor measurements into spatial products that can be queried, compared, demonstrated, and used in decision-making. In provenance terms, an EO derivative should be treated as a generated prov:Entity whose meaning depends on the chain of prov:Activity instances that created it. This is particularly important in coastal environments, where the same raw observation may generate multiple products with different scientific and administrative meanings: an instantaneous waterline, a tidal-datum shoreline, a wetland boundary, a flood extent, a turbidity plume, or a coastal-risk zone.

Table 6 summarises key coastal EO derivative classes, their enabling observation technologies, provenance-critical elements, and coastal-administrative uses. The table illustrates why a generic dataset-level lineage record is insufficient for next-generation coastal observation. A derived shoreline, for instance, must be linked not only to the source image but also to the shoreline proxy, tide level, water-index threshold, cloud mask, validation transects, uncertainty estimate, and versioned output geometry. Similarly, UAV and LiDAR elevation products must retain links to ground control, sensor calibration, point-cloud filtering, vertical datum, interpolation, and accuracy assessment. These elements determine whether the derivative is suitable for coastal engineering, risk zoning, habitat assessment, cadastral update, or emergency response.

Table 6. EO-derived coastal products, enabling remote sensing technologies, provenance-critical elements, and coastal-administrative uses.

EO-Derived Coastal Product	Enabling Technologies	Provenance-Critical Elements	Coastal-Administrative Impact
Satellite-derived shoreline or coastline vectors	Landsat; Sentinel-2; PlanetScope; WorldView; GeoEye; CoastSat; DSAS; AI segmentation	Scene ID; acquisition date; tide stage; waterline proxy; NDWI/MNDWI or AI model; threshold; cloud/shadow mask; CRS; transect baseline; validation data; positional uncertainty	Shoreline monitoring; erosion-rate reporting; setback planning; coastal-risk zoning; land–sea boundary interpretation
UAV photogrammetry and LiDAR-derived DEM/DSM products	UAV RGB/multispectral/thermal imaging; Structure-from-Motion; UAV LiDAR; airborne LiDAR	Flight plan; altitude; overlap; camera calibration; GNSS/IMU data; ground control points; point-cloud density; filtering; interpolation; vertical datum; RMSE	Beach and dune volume estimation; cliff retreat monitoring; infrastructure exposure; engineering design; post-storm damage assessment
Satellite-derived bathymetry and nearshore morphology	Sentinel-2; Landsat; ICESat-2; airborne/topo-bathymetric LiDAR; machine-learning bathymetry models	Atmospheric correction; water-column correction; turbidity condition; optically deep-water mask; bottom type; depth-control points; model version; depth uncertainty	Navigation support; dredging assessment; coastal flood modelling; habitat mapping; nearshore sediment-dynamics analysis

Table 6. *Cont*.

EO-Derived Coastal Product	Enabling Technologies	Provenance-Critical Elements	Coastal-Administrative Impact
Water-quality and open-water dynamics products	MODIS; Sentinel-3 OLCI; Sentinel-2 MSI; Landsat; thermal sensors; in situ sensors; ocean-colour algorithms	Atmospheric correction; algorithm selection; in situ matchups; turbidity/chlorophyll/SST retrieval method; temporal compositing; uncertainty estimates	Pollution monitoring; algal-bloom response; fisheries and marine-park management; water-quality compliance; ecosystem health assessment
Mangrove, wetland, tidal-flat, seagrass, and habitat maps	Sentinel-1 SAR; Sentinel-2; Landsat; hyperspectral sensors; UAV imagery; AI classification	Training data; class definitions; seasonality; tidal condition; segmentation/classification model; accuracy assessment; confusion matrix; validation samples	Conservation planning; blue-carbon accounting; restoration monitoring; habitat protection; environmental impact assessment
Coastal flood, storm-surge, and inundation extents	SAR; optical satellites; DEMs; hydrodynamic models; tide gauges; IoT sensors; UAV surveys	Event timestamp; water-level source; DEM source/resolution; flood-detection algorithm; model parameters; validation observations; uncertainty surface	Emergency response; disaster mitigation; insurance and exposure modelling; evacuation planning; infrastructure resilience
Coastal land-use/land-cover and artificial-structure layers	VHR satellite imagery; UAV imagery; object-based image analysis; AI classification; cadastral and infrastructure datasets	Classification schema; training samples; object-extraction method; date of imagery; spatial resolution; administrative boundary version; accuracy metrics	Coastal development control; asset registry maintenance; seawall and groyne monitoring; compliance assessment; planning approvals

3.9. Provenance Integration, Querying, and Interoperability

This section synthesises how semantic web-based provenance is integrated, queried, and standardised within Spatial Data Supply Chains (SDSCs). It examines dominant workflow-integration patterns, querying and reasoning capabilities, and interoperability practices across geospatial domains. The review highlights persistent trade-offs between deployment practicality, granularity, and standards compliance that constrain operational SDSC provenance.

3.10. Integration with Geospatial Workflows

Provenance capture in SDSCs follows three dominant integration patterns (Table 7): embedded capture, service-oriented collection, and post hoc reconstruction, with a smaller set of hybrid approaches combining these strategies.

Table 7. Integration of architectural patterns.

Pattern	Studies (n)	Advantages	Limitations
Embedded Capture	32	High completeness, real-time	Tool modification required
Service-Oriented	28	Scalable, standardised	May miss minute details
Post hoc Reconstruction	21	Non-invasive, retroactive	Incomplete, error-prone
Hybrid Approaches	8	Balanced trade-offs	Complex implementation

Note: Percentages are calculated against the analytically coded implementation subset ($n = 89$) because not all 156 included publications reported sufficient technical detail for this coding category. Categories are non-exclusive where studies used multiple technologies, standards, or capabilities.

Embedded capture, adopted by 32 studies, integrates provenance recording directly within geoprocessing [5,8] tools and workflows. This approach offers high completeness and near real-time capture but typically requires modification of existing tools or execution environments. Service-oriented approaches (28 studies) rely on workflow orchestrators or middleware to collect provenance through standardised interfaces, enabling scalability across heterogeneous systems but often omitting fine-grained operational details. Post hoc reconstruction (21 studies) derives provenance retrospectively from logs and metadata, minimising system intrusion but frequently producing incomplete or error-prone lineage. Hybrid approaches (eight studies) attempt to balance these trade-offs, achieving broader coverage at the cost of increased architectural complexity.

Overall, no single integration pattern fully satisfies SDSC requirements. Embedded approaches favour completeness, service-oriented models prioritise scalability, and post hoc techniques maximise deployability. This fragmentation underscores the need for adaptive provenance frameworks capable of combining capture modes within a unified semantic model.

3.11. Querying and Reasoning Production

Querying support across reviewed studies is uneven (Table 8). Basic lineage traversal, tracing data origins and dependencies, is almost ubiquitous, supported by 94% of implementations. In contrast, advanced capabilities are substantially less common: temporal provenance queries appear in 42% of studies, spatially constrained provenance queries in 34%, and automated reasoning or inference in only 38%.

Table 8. Querying and reasoning capabilities.

Capability	Studies (n)	Percentage	Complexity Level
Basic Lineage Traversal	84	94.4%	Simple
Temporal Provenance Queries	37	41.6%	Moderate
Spatial Provenance Queries	30	33.7%	Moderate
Trust and Quality Assessment	28	31.5%	Complex
Automated Reasoning	34	38.2%	Complex
Cross-dataset Integration	25	28.1%	Complex

Note: Percentages are calculated against the analytically coded implementation subset ($n = 89$) because not all 156 included publications reported sufficient technical detail for this coding category. Categories are non-exclusive where studies used multiple technologies, standards, or capabilities.

Approximately one-third of studies expose provenance through SPARQL endpoints or linked-data catalogues, enabling flexible querying of lineage metadata. More advanced systems incorporate rule-based reasoning for trust and quality assessment, formal consistency checking, and explanation generation, often supported by interactive dashboards.

However, large graph sizes, complex spatial predicates, and reasoning overheads remain significant barriers to scalable querying in production SDSCs.

3.12. Interoperability and Standards Compliance

Interoperability through standards alignment is a prerequisite for SDSC deployment and is addressed in approximately 78% of reviewed studies (Table 9). Most implementations combine W3C PROV with established geospatial standards, including ISO 19115/19115-2 lineage, OGC web services, and EPCIS-style supply chain specifications [3,92].

Table 9. Standards compliance and interoperability.

Standard/Specification	Studies (n)	Compliance Level	Integration Challenges
OGC Web Services	45	High	Legacy system integration
DCAT/GeoDCAT-AP	30	Moderate	Semantic mapping complexity
STAC	11	Emerging	Limited provenance extensions
ISO 19100 Series	52	High	Semantic gap with PROV
Dublin Core	28	Moderate	Limited spatial semantics

Note: Counts are non-exclusive because a single study may align with multiple standards or specifications.

ISO 19100-series standards show the highest adoption, reflecting their role in catalogue-level metadata and regulatory compliance, while DCAT and GeoDCAT-AP are increasingly used to support cross-catalogue interoperability. Integration with STAC appears in a smaller but growing subset of recent studies, signalling the rise in cloud-native geospatial infrastructures. Despite these advances, semantic alignment between PROV-based models and ISO lineage remains inconsistently formalised, often relying on custom mappings that inhibit reuse.

The review consistently identifies a tension between semantic expressiveness and strict standards compliance. Rich provenance models enable granular reasoning and traceability but frequently diverge from established specifications, requiring bespoke extensions. Conversely, strict adherence to standards often limits the ability to represent dynamic workflows, feature-level lineage, and intermediate processing states.

4. Evidence-Based Gaps and Limitations

The reviewed literature demonstrates substantial progress in applying semantic web technologies to provenance in SDSCs, with widespread adoption of W3C PROV and ISO lineage standards complemented by ontology-driven integration and linked-data principles. Reported strengths include structured provenance modelling, interoperability across heterogeneous systems, and advanced querying capabilities that enhance transparency, trust, and reproducibility. Nevertheless, the synthesis reveals persistent and systematic gaps that limit the operational effectiveness of provenance in real-world SDSCs.

4.1. Recurring Evidence-Based Gaps

First, provenance granularity remains insufficient. Despite formal support for feature- and attribute-level provenance in PROV-based extensions, most implementations continue to operate at the dataset or workflow level. Fine-grained provenance is inconsistently modelled, computationally expensive to query at scale, and fragmented across domains, hindering reuse and cross-system integration [3,51,92].

Second, integration with heterogeneous geospatial workflows remains incomplete. Embedded capture mechanisms demonstrate high completeness but require intrusive tool modifications, while service-oriented and post hoc approaches prioritise scalability and deploya-

bility at the expense of detail. The absence of unified capture strategies results in provenance gaps across distributed service-oriented architectures and cloud-native pipelines [4,7,8].

Third, querying, reasoning, and usability constraints persist. While SPARQL and OWL enable lineage traversal, inference, and quality assessment, large provenance graphs introduce latency and storage bottlenecks. Moreover, explanation interfaces, visual analytics, and formal verification mechanisms remain underdeveloped, particularly for non-expert users and decision-makers [29,149,150].

Fourth, interoperability and standardisation remain fragmented. Although mappings between PROV, ISO 19115/19157, and OGC specifications exist, they are often informal or locally defined. Vendor-specific extensions and inconsistent identifier practices limit cross-domain exchange, and no widely accepted spatial provenance profile has yet emerged [3,92].

Finally, automation and scalability pose major challenges. Semantic execution engines and agent-based capture techniques reduce manual effort, but their adoption is constrained by high integration costs, domain-specific customisation, and performance limitations in streaming and big data scenarios. Reasoning and storage bottlenecks restrict practical deployment at graph scales of 10^7–10^9 triples.

4.2. Implications for GeoPROV

In the context of coastal observation, GeoPROV enables transparent lineage tracking across multi-source Earth observation workflows [129–131,151], supporting reproducibility and decision-making in applications such as shoreline monitoring, coastal hazard assessment, and marine ecosystem analysis. Furthermore, the framework supports provenance-aware remote sensing workflows [5,37], particularly in coastal observation contexts such as shoreline monitoring and environmental assessment [152].

These evidence-based gaps motivate the need for lightweight, standards-aligned semantic profiles rather than increasingly complex domain-specific ontologies. GeoPROV responds directly to the identified limitations by formalising spatially explicit PROV patterns while preserving compatibility with ISO metadata, GeoSPARQL, and emerging catalogue standards such as STAC. By minimising modelling overhead and supporting incremental adoption, GeoPROV aims to improve interoperability, granularity, and operational deployability without exacerbating scalability constraints.

4.3. Limitations of the Review

This review has several limitations. The literature search was restricted to English-language publications, introducing potential language bias. Although multiple academic databases were queried, the relevant grey literature and proprietary implementations may be under-represented. Moreover, the synthesis largely reflects conceptual and architectural contributions, as large-scale empirical deployments of semantic provenance in operational SDSCs remain limited. Finally, given the rapid evolution of semantic web technologies and geospatial standards, the findings represent a temporal snapshot that will require periodic updating to remain current.

5. GeoPROV: A Minimal Semantic Profile for Spatial Provenance

5.1. Design Rationale

Unlike existing PROV-O extensions, GeoPROV prioritises minimalism and interoperability by defining only essential spatial constructs, avoiding complex workflow modelling overhead. GeoPROV is proposed as a minimal semantic profile of the W3C PROV-O, explicitly designed to address the recurrent and systematically observed provenance limitations identified in Section 3 and synthesised in Table 10.

Table 10. Comparison of provenance models and standards identified in the literature review.

Feature	PROV-O	GeoPROV (Published)	GeoPROV Coastal EO Application Profile
Generic Provenance	✓	✓	✓
Spatial Entities	Limited	✓	✓
GeoSPARQL Support	Limited	✓	✓
SDSC Focus	No	Yes	Yes
Coastal EO Workflows	No	No	Yes

Rather than introducing a comprehensive domain ontology or an execution-level workflow model, GeoPROV targets a narrowly defined but critical gap: the absence of spatially explicit, feature-level, standards-aligned provenance that can be operationalised within real-world Spatial Data Supply Chains (SDSCs). The proposed GeoPROV framework builds upon previous provenance research while extending existing approaches through explicit representation of geospatial entities, spatial processing activities, observational datasets, quality metadata, and supply chain relationships. Table 11 compares GeoPROV with existing provenance frameworks. The contribution of this study lies in an EO-specific implementation profile that specifies how existing provenance constructs are mapped to observational resources, geoprocessing workflows and organisational responsibilities throughout the data supply chain. In this profile, source observations, intermediate datasets and derived products are mapped to prov:Entity; acquisition, preprocessing, transformation, quality assessment, publication and dissemination are mapped to prov:Activity; and the organisations responsible for these stages are mapped to prov:Agent. The proposed mappings connect processing inputs and outputs with their parameters, software and model versions, validation evidence and responsible organisations. They also extend the workflow description through organisational handovers to the final dissemination stage, linking computational derivation with responsibility for product delivery. The intended capability is therefore an integrated trace that supports both backward investigation from a disseminated product to its observations, processing history and responsible organisations—and forward dependency tracing from an input or processing step to affected downstream products. These capabilities arise from the coordinated application of existing semantics; this study does not claim new fundamental provenance classes and reasoning operators. Its incremental contribution is the proposed reusable mapping and implementation specification for EO data supply chains, with coastal workflows serving as the illustrative application rather than defining the limits of the profile.

The design of GeoPROV was directly informed by the evidence-based gaps identified in the systematic review. First, the predominance of dataset-level provenance across SDSCs highlighted the need for finer-grained traceability; consequently, GeoPROV supports feature-level entities and persistent identifiers to enable lineage tracking at the level of individual spatial features. Second, the review revealed weak and often inconsistent alignment between W3C PROV and ISO 19115/19157 lineage models, motivating explicit mappings that bridge machine-actionable provenance graphs with established geospatial metadata and quality standards. Third, the limited representation of spatial context within existing provenance frameworks led to the incorporation of GeoSPARQL, enabling geospatial semantics, spatial relationships, and location-aware provenance queries. Fourth, the lack of systematic provenance validation mechanisms identified in many studies is addressed through optional SHACL constraints that support completeness, consistency, and standards-conformance checking. Finally, the review found limited integration between provenance frameworks and modern Earth Observation cataloguing ecosystems; GeoPROV therefore incorporates STAC-compatible linking mechanisms to connect provenance records with EO

assets and cloud-native catalogue infrastructures. Collectively, these design decisions ensure that GeoPROV responds directly to the principal limitations identified in the literature while maintaining a lightweight, interoperable, and implementation-oriented profile for provenance-aware spatial data supply chains.

Table 11. Strengths and weaknesses in provenance automation and big data scalability.

Aspect	Strengths	Weaknesses
Provenance modelling frameworks	PROV widely extended with GeoSPARQL and DCAT, enabling expressive, interoperable models and end-to-end chains	Feature- and attribute-level detail inconsistent; domain fragmentation hinders reuse.
Workflow integration	Provenance-aware WPS/GIService chaining captures lineage at execution time.	Heterogeneous platforms limit plug-and-play integration.
Granularity and expressiveness	Typed workflows and QA-linked steps support error tracing	Dataset-level provenance dominates; fine-grained modelling burdensome at scale
Querying, reasoning, visualisation	SPARQL/OWL enable inference, fitness-for-use dashboards	Latency, usability, and explanation challenges persist
Interoperability and standardisation	ISO–PROV alignments and common models reduce silos	Vendor-specific extensions; lack of common geo-provenance profile
Automation	Semantic execution engines improve completeness	High integration effort; limited support for real-time pipelines.
Big data scalability	Linked-data lifecycle tooling enables enrichment	Storage and reasoning bottlenecks at large graph scales

GeoPROV is guided by three evidence-driven principles. Minimalism limits extensions to those strictly required for spatial provenance, addressing concerns over modelling complexity and scalability. Standards alignment maintains full compatibility with PROV-O while explicitly linking to ISO 19115/19157 and GeoSPARQL, resolving persistent informality in PROV–ISO mappings. Deployability emphasises incremental adoption within existing SDSCs without requiring architectural re-engineering. Grounded in the recurring limitations identified in Table 11—coarse provenance granularity, weak interoperability, and inconsistent standards integration—GeoPROV is positioned as a pragmatic, adoption-oriented profile rather than a speculative model.

The systematic review identified several recurring gaps that directly informed the design of GeoPROV. Most existing provenance implementations remain focused on dataset-level or workflow-level lineage, providing limited support for feature-level traceability. To address this limitation, GeoPROV introduces feature-level entities and persistent identifiers that enable provenance tracking of individual EO-derived coastal features. The review also revealed weak and inconsistent alignment between W3C PROV and ISO 19115/19157 lineage models, motivating explicit mappings between machine-actionable provenance graphs and established geospatial metadata and quality standards. Limited support for spatial semantics was addressed through the integration of GeoSPARQL, enabling spatially aware provenance representation and querying. The relative absence of provenance validation mechanisms in the literature led to the inclusion of optional SHACL constraints for provenance completeness, consistency, and standards-conformance checking. Finally, the limited integration of provenance frameworks with modern Earth Observation catalogues motivated STAC-compatible linking mechanisms that connect provenance records with EO

assets and cloud-native discovery infrastructures. Collectively, these capabilities position GeoPROV as a lightweight, interoperable, and standards-aligned provenance profile for EO-derived coastal products and spatial data supply chains.

5.2. Scope and Contribution

Within this explicitly constrained scope, GeoPROV introduces four complementary contributions, each corresponding to an empirically observed gap summarised in Table 11. First, it supports spatially typed prov:Entity instances, enabling datasets, collections, and individual features to be consistently associated with spatial semantics and persistent identifiers, thereby addressing the predominance of dataset-level provenance. Second, GeoPROV enables explicit modelling of core geospatial processing activities, including reprojection, resampling, and generalisation, allowing spatial transformations and their parameters to be represented transparently within provenance graphs. Third, it promotes feature-level persistent identifiers, supporting fine-grained lineage tracking across iterative processing, partial reprocessing, and versioning cycles, as required in large-scale operational systems. Fourth, GeoPROV defines optional SHACL constraints that enable validation of provenance completeness, consistency, and standards conformance without imposing mandatory reasoning or execution semantics, mitigating performance and scalability concerns.

GeoPROV is explicitly not intended to replace ISO metadata standards, workflow specification languages, or domain-specific ontologies. Instead, it functions as a semantic bridging layer, linking operational SDSC workflows with ISO-based metadata and quality frameworks through a machine-actionable PROV-O profile. As demonstrated by the review-to-design mapping in Table 11, GeoPROV constrains its ambition to what is empirically necessary, providing a standards-aligned pathway toward interoperable, provenance-centric SDSCs that reflect real operational dynamics rather than idealised linear pipelines.

5.3. GeoPROV Conceptual Model

GeoPROV extends the W3C PROV-O framework through a minimal set of spatially aware constructs to support feature-level provenance in Spatial Data Supply Chains (SDSCs). It retains PROV-O's entity–activity–agent model while introducing explicit spatial typing, persistent feature-level identifiers, and constrained extensions derived directly from the evidence-based gaps summarised in Table 11. Spatial datasets and features are modelled as typed prov:Entity instances, enabling lineage tracing at operational granularity. Geospatial processes (e.g., reprojection, resampling, aggregation) are represented as parameterised prov:Activity instances, addressing the informal or incomplete modelling of spatial transformations observed in prior work. Interoperability is ensured through explicit alignment with ISO 19115/19157 and GeoSPARQL, while optional SHACL constraints support validation without imposing scalability-limiting reasoning requirements. Collectively, GeoPROV functions as a lightweight semantic integration layer optimised for interoperability, deployability, and analytical fitness.

5.4. GeoPROV Pattern for EO-Derived Coastal Observation Products

For coastal remote sensing, GeoPROV can be operationalised as a provenance pattern that links raw observations, processing activities, EO derivatives, validation evidence, and administrative outputs. The pattern begins with the acquisition entity, represented as a prov:Entity linked to a STAC item or equivalent catalogue record. This entity records the sensor, platform, acquisition time, spatial footprint, spectral bands, spatial resolution, cloud cover, orbit or flight metadata, and relevant environmental context such as tide level or water condition. For UAV and LiDAR observations, the acquisition entity also

records flight plan, ground control configuration, GNSS/IMU metadata, point density, and calibration information.

The second component is the processing activity, modelled as a parameterised prov:Activity. In coastal workflows, this may include atmospheric correction, radiometric calibration, orthorectification, reprojection, water-index calculation, tidal normalisation, cloud and shadow masking, AI segmentation, bathymetric inversion, water-column correction, point-cloud classification, DEM generation, or flood-extent extraction. GeoPROV records the software, algorithm, model version, threshold, training data, processing date, responsible agent, and computational environment associated with each activity. This allows derived products to be interpreted as reproducible outcomes rather than detached map layers.

The third component is the EO-derived coastal product, represented as a spatially typed prov:Entity with a persistent identifier and GeoSPARQL-compatible geometry. Examples include shoreline segments, erosion-rate transects, DEM tiles, bathymetric grids, turbidity surfaces, mangrove polygons, flood boundaries, and infrastructure exposure features. Each product is linked to the activity that generated it through prov:wasGeneratedBy and to its source observations through prov:wasDerivedFrom. Feature-level identifiers allow partial updates, local reprocessing, and version comparisons to be traced without regenerating provenance for the entire dataset [153].

The fourth component is validation and quality assessment. GeoPROV links each derivative to validation datasets, field measurements, UAV/GNSS checkpoints, confusion matrices, RMSE values, positional accuracy measures, thematic accuracy measures, uncertainty surfaces, and ISO 19157 quality elements. This enables users to query not only where a coastal feature came from but also how reliable it is for a specific purpose.

The final component is administrative use. Coastal products frequently enter planning, risk, emergency, conservation, engineering, or cadastral workflows. GeoPROV can represent these downstream uses as prov:Activity instances that consume EO derivatives and generate administrative outputs such as erosion-risk zones, development setbacks, flood-response maps, restoration-priority areas, or coastal asset registers. This design connects scientific provenance with governance accountability, allowing users to determine which observation, algorithm, validation evidence, and uncertainty estimate supported a particular coastal decision.

Figure 2 summarises this GeoPROV-enabled sensor-to-decision chain for EO-derived coastal observation products. GeoPROV-enabled EO derivative chain for coastal observation. The figure illustrates how raw satellite, UAV/LiDAR, IoT, tide-gauge, and GNSS observations are transformed through preprocessing, AI/geoprocessing, EO-derived coastal feature generation, validation and quality assessment, and STAC/ISO catalogue publication before supporting coastal-administrative decision layers. The GeoPROV semantic provenance graph acts as a cross-cutting machine-actionable layer that records prov:Entity, prov:Activity, prov:Agent, GeoSPARQL geometry/CRS, ISO lineage and quality information, SHACL constraints, and STAC links across the full sensor-to-decision chain.

The systematic review revealed that provenance requirements vary substantially across coastal EO-derived products. Therefore, GeoPROV defines a core provenance model combined with product-specific provenance extensions to ensure that information influencing reliability, reproducibility, and uncertainty can be consistently recorded.

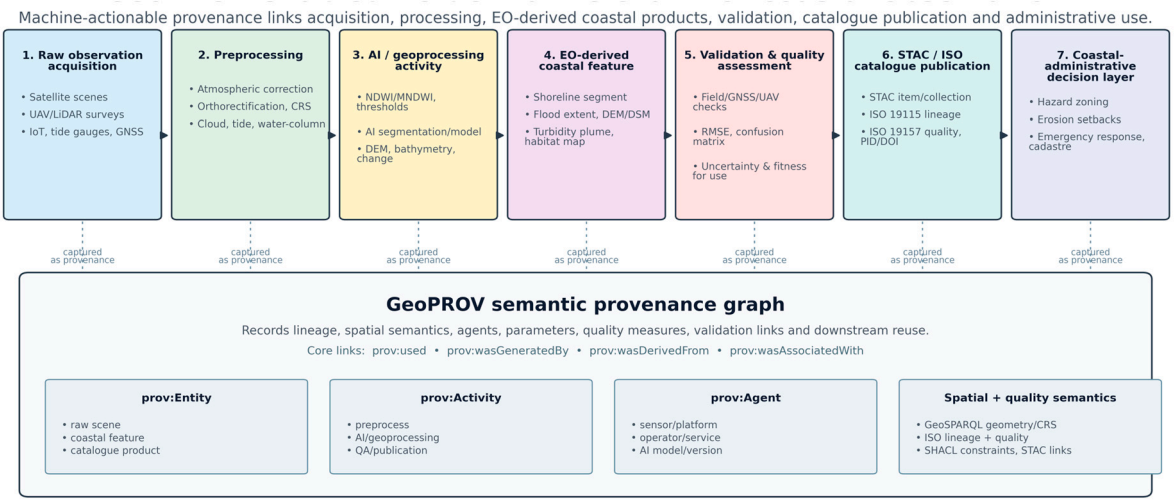


Figure 2. GeoPROV-enabled EO derivative chain for coastal observation.

5.5. Example Application: Sentinel-2/UAV Shoreline Monitoring

A provenance-aware shoreline monitoring workflow may begin with a Sentinel-2 Level-1C scene, supported by tide-gauge observations and optional UAV validation data [154]. The preprocessing stage applies atmospheric correction, cloud and shadow masking, orthorectification, reprojection, and water-index calculation [155]. A shoreline extraction activity then applies an NDWI or MNDWI threshold, an AI segmentation model, or a hybrid rule-based method to generate shoreline vectors [156]. The resulting shoreline segments are represented as spatially typed prov:Entity instances with persistent feature identifiers and GeoSPARQL-compatible geometries. Each segment is linked through prov:wasGeneratedBy to the extraction activity and through prov:wasDerivedFrom to the source scene, tide record, cloud mask, and validation dataset. ISO 19157 quality elements record positional accuracy, validation RMSE, completeness, and uncertainty. The final shoreline product can then be published through a STAC/ISO catalogue and reused in DSAS-based erosion-rate estimation, setback planning, flood-risk assessment, or land–sea boundary interpretation.

This example illustrates the central value of GeoPROV: it preserves the evidential chain from observation to administrative decision, allowing users to determine not only where a coastal feature is located but also how, when, by whom, and with what uncertainty it was produced.

The Sentinel-2/UAV shoreline monitoring example demonstrates the conceptual applicability and completeness of the GeoPROV framework. However, it should not be interpreted as formal validation. Future work should implement GeoPROV within operational Earth Observation infrastructures and evaluate interoperability, scalability, provenance completeness, and query performance using real-world datasets. GeoPROV is not intended to replace existing provenance frameworks. Rather, it serves as a lightweight integration profile that combines complementary capabilities from PROV-O, GeoSPARQL, SHACL, ISO 19115/19157 lineage concepts, and STAC metadata specifications. Its primary contribution lies in adapting and integrating these established standards to address provenance requirements specific to coastal EO-derived products.

5.6. GeoPROV Formal Specification

To support implementation and interoperability, GeoPROV is defined as a lightweight semantic profile extending established provenance and geospatial standards. The profile adopts PROV-O classes for entities, activities, and agents; GeoSPARQL for spatial representations; STAC for EO asset descriptions; and SHACL for constraint validation.

Core GeoPROV classes, properties, namespace definitions, and mappings to source standards are provided in Supplementary Materials. Machine-readable implementations including RDF/Turtle examples, SHACL validation shapes, GeoSPARQL queries, and JSON-LD serialisations are publicly available through the project repository.

6. Implications and Future Research Directions

The findings of this review highlight the need to treat provenance as a first-class, operational component of Spatial Data Supply Chains rather than as auxiliary documentation. While semantic web-based approaches demonstrate the potential of machine-actionable provenance to improve transparency, trust, and reproducibility, their impact is constrained by fragmented integration, limited feature-level granularity, and inconsistent standards alignment. GeoPROV addresses these issues through a lightweight semantic profile that enables spatially explicit provenance while remaining compatible with PROV-O, ISO 19115/19157, GeoSPARQL, and emerging catalogue standards, supporting interoperable and incrementally deployable provenance across SDSCs. For coastal observation, the main implication is that provenance must move downstream with EO derivatives rather than remaining attached only to raw scenes or processing reports. A satellite-derived shoreline, UAV-derived DEM, AI-generated flood boundary, or turbidity plume map may be reused many times across erosion monitoring, engineering design, ecological assessment, and coastal planning. If the derivative is separated from its acquisition conditions, model parameters, validation evidence, and uncertainty estimates, its administrative and scientific meaning becomes unstable. GeoPROV addresses this challenge by supporting feature-level provenance that connects the sensor-to-decision chain: acquisition, preprocessing, AI or geospatial analysis, derivative generation, quality assessment, catalogue publication, and administrative reuse.

Future research should therefore prioritise provenance-aware EO derivative pipelines for coastal systems. Three areas are especially important. First, AI-enabled coastal workflows require provenance models that capture training data lineage, annotation rules, model versioning, prompt design, uncertainty, and postprocessing decisions. Second, open-water and intertidal products require provenance models that explicitly represent tide, wave, turbidity, water-column correction, bathymetric assumptions, and in situ validation. Third, coastal-administrative applications require provenance links between EO-derived features and planning, cadastral, emergency-response, conservation, and infrastructure-management decisions. Addressing these areas would allow provenance to function not merely as a technical record but also as a mechanism for reproducible coastal science and accountable coastal governance.

The review identified substantial variation in provenance requirements across coastal EO-derived products. Consequently, GeoPROV adopts a core provenance model supported by product-specific extensions. For example, shoreline products require documentation of shoreline indicators, tidal datum, acquisition time, extraction methodology, and positional accuracy. SAR flood products require information on sensor configuration, polarisation, incidence angle, processing parameters, classification methods, and validation metrics. UAV and LiDAR products additionally require flight characteristics, ground control information, vertical datum specifications, and point-cloud processing workflows. Capturing these provenance elements improves product traceability, reproducibility, uncertainty assessment, and fitness-for-purpose evaluation.

Future research should prioritise automated, scalable provenance capture in cloud-native, streaming, and AI-driven geospatial pipelines, where manual and post hoc methods are inadequate. Key challenges include advancing feature-level provenance, improving scalable graph storage and SPARQL optimisation, and developing selective reasoning

and user-centred visualisation to support billion-triple analytics. Progress in these areas, supported by standardised profiles such as GeoPROV, is essential for operationalising provenance-centric spatial data governance in complex, data-intensive environments.

7. Conclusions

This study provided a systematic review of 156 publications on semantic web-based provenance modelling in Spatial Data Supply Chains (SDSCs). While standards such as W3C PROV and ISO 19115/19157 offer a strong foundation, their application in SDSCs remains fragmented, with most approaches limited to coarse, dataset-level provenance and ad hoc integrations that hinder interoperability, feature-level traceability, and scalability. These findings indicate that effective SDSC provenance requires coherent, lightweight semantic integration beyond existing technical capabilities.

In response, this paper introduced GeoPROV, a minimal semantic profile that extends PROV-O with spatial semantics while remaining aligned with established geospatial and catalogue standards. Grounded in evidence from the review, GeoPROV acts as an enabling layer for EO-derived coastal products, including satellite-derived shorelines, UAV and LiDAR elevation models, bathymetric grids, water-quality indicators, habitat classifications, flood extents, and coastal-administrative decision layers. Its value lies in connecting raw observation, processing activity, AI or geospatial model, derived feature, validation evidence, uncertainty estimate, and downstream use within a machine-actionable provenance graph. Advancing provenance-centric SDSCs will require collaboration among remote sensing scientists, coastal managers, standards bodies, and platform developers so that future coastal observation systems are not only data-rich, but also transparent, reproducible, interoperable, and administratively defensible.

Because the systematic review was conducted across multiple SDSC domains, the coastal Earth Observation examples presented in this study should be interpreted as representative application scenarios rather than a comprehensive review of coastal remote sensing systems. The review synthesised provenance requirements from both coastal EO applications and the broader geospatial provenance literature, which collectively informed the development of the GeoPROV profile for coastal EO-derived products.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18213737/s1.

Author Contributions: M.A.S.: Conceptualisation, Methodology, Investigation, Formal analysis, Data curation, Writing—original draft, Writing—review and editing, Visualisation, Project administration, Funding acquisition. P.K.L.: Supervision, Writing—review and editing, A.N.: Supervision, Writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data supporting the findings of this study are derived from publicly available scholarly publications identified through Scopus, Web of Science, IEEE Xplore, and SpringerLink, as described in the review methodology. No new primary empirical, participant-level, or proprietary geospatial datasets were generated or analysed during this study. The bibliographic records, screening decisions, PRISMA summary data, and coding classifications used to support the synthesis and tables are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to acknowledge the academic and research support provided by the School of Engineering and Technology at Central Queensland University. During the preparation of this manuscript, the authors used Microsoft 365 Copilot basic and language-editing support tools to assist with grammar, clarity, and structural refinement. The authors reviewed, edited, and verified all generated content and take full responsibility for the accuracy, originality, and integrity of the final manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SDSC	Spatial Data Supply Chain
PROV-O	Provenance Ontology (W3C)
RDF	Resource Description Framework
OWL	Web Ontology Language
SPARQL	SPARQL Protocol and RDF Query Language
GeoSPARQL	Geographic SPARQL
ISO	International Organization for Standardization
STAC	SpatioTemporal Asset Catalog
SHACL	Shapes Constraint Language
SDI	Spatial Data Infrastructure
EO	Earth Observation
WPS	Web Processing Service
OGC	Open Geospatial Consortium
PID	Persistent Identifier
AI	Artificial Intelligence
IoT	Internet of Things
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses

References

1. Varadharajulu, P.; Azeem Saqiq, M.; Yu, F.; McMeekin, D.A.; West, G.; Arnold, L.; Moncrieff, S. Spatial data supply chains. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 41–45. [CrossRef]
2. Sadiq, M.A.; West, G.; McMeekin, D.A.; Arnold, L.; Moncrieff, S. Provenance ontology model for land administration spatial data supply chains. In *Proceedings of the 11th International Conference on Innovations in Information Technology (IIT), Dubai, United Arab Emirates, 1–3 November 2015*; IEEE: New York, NY, USA, 2015; pp. 184–189. [CrossRef]
3. Sadiq, M.A.; McMeekin, D.; Arnold, L. Spatial data supply chain provenance modelling for next generation spatial infrastructures using semantic web technologies. *GEOProcessing* **2016**, *2016*, 157.
4. Scheider, S.; Ballatore, A. Semantic typing of linked geoprocessing workflows. *Int. J. Digit. Earth* **2018**, *11*, 113–138. [CrossRef]
5. Tullis, J.A.; Lanter, D.P.; Basu, A.; Cothren, J.D.; Shi, X.; Limp, W.F.; Linck, R.F.; Young, S.G.; Davis, J.; Alsumaiti, T.S. Geoprocessing, workflows, and provenance. In *Remote Sensing Handbook, Volume II*; CRC Press: Boca Raton, FL, USA, 2024; pp. 303–337.
6. Young, S.G. 13 Geoprocessing, Workflows. In *Remote Sensing Handbook, Volume II: Image Processing, Change Detection, GIS, and Spa-tial Data Analysis*; CRC Press: Boca Raton, FL, USA, 2024; p. 303.
7. Yue, P.; Guo, X.; Zhang, M.; Jiang, L.; Zhai, X. Linked Data and SDI: The case on Web geoprocessing workflows. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 245–257. [CrossRef]
8. Zhang, M.; Jiang, L.; Zhao, J.; Yue, P.; Zhang, X. Coupling OGC WPS and W3C PROV for provenance-aware geoprocessing workflows. *Comput. Geosci.* **2020**, *138*, 104419. [CrossRef]
9. Vancisin, T.; Clarke, L.; Orr, M.; Hinrichs, U. Provenance visualization: Tracing people, processes, and practices through a data-driven approach to provenance. *Digit. Scholarsh. Humanit.* **2023**, *38*, 1322–1339. [CrossRef]
10. Tullis, J.A.; Kar, B. Where is the provenance? Ethical replicability and reproducibility in GIScience and its critical applications. *Ann. Am. Assoc. Geogr.* **2021**, *111*, 1318–1328. [CrossRef]
11. Mai, G.; Huang, W.; Sun, J.; Song, S.; Mishra, D.; Liu, N.; Gao, S.; Liu, T.; Cong, G.; Hu, Y. On the opportunities and challenges of foundation models for geospatial artificial intelligence. *arXiv* **2023**, arXiv:2304.06798.
12. Rieke, M.; Bigagli, L.; Herle, S.; Jirka, S.; Kotsev, A.; Liebig, T.; Malewski, C.; Paschke, T.; Stasch, C. Geospatial IoT—The Need for Event-Driven Architectures in Contemporary Spatial Data Infrastructures. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 385. [CrossRef]
13. Vandenbroucke, D.; Olijslagers, M.; Boguslawski, R.; Borzachiello, M.; Perego, A.; Smith, R. *Architectures and Standards for Spatial Data Infrastructures and Digital Government*; European Commission, Joint Research Center: Brussels, Belgium, 2020.
14. Chafiq, T.; Azmi, R.; Fadil, A.; Mohammed, O. Investigating the potential of blockchain technology for geospatial data sharing: Opportunities, challenges, and solutions. *Geomatica* **2024**, *76*, 100026. [CrossRef]
15. Karimi, H.A. *Big Data: Techniques and Technologies in Geoinformatics*; CRC Press: Boca Raton, FL, USA, 2025; 410p.
16. Zhang, F.; Wang, Z.; Guo, R.; Qu, G. Earth observation data provenance: A blockchain-based solution. *IEEE Trans. Ind. Inform.* **2024**, *20*, 9548–9556. [CrossRef]

17. Van Hoang, T. Impact of integrated artificial intelligence and internet of things technologies on smart city transformation. *J. Tech. Educ. Sci.* **2024**, *19*, 64–73. [CrossRef]
18. Rane, N. Integrating Leading-Edge Artificial Intelligence (AI), Internet of Things (IOT), and Big Data Technologies for Smart and Sustainable Architecture, Engineering and Construction (AEC) Industry: Challenges and Future Directions. *SSRN Electron. J.* **2023**. [CrossRef]
19. Ahmed, I.; Zhang, Y.; Jeon, G.; Lin, W.; Khosravi, M.R.; Qi, L. A blockchain- and artificial intelligence-enabled smart IoT framework for sustainable city. *Int. J. Intell. Syst.* **2022**, *37*, 6493–6507. [CrossRef]
20. Joshi, G.; Ghamisi, P. In Metadata We Trust: SartOn the central role of metadata in Earth observation datasets [Software and Data Sets]. *IEEE Geosci. Remote Sens. Mag.* **2026**, *14*, 161–166. [CrossRef]
21. Dube, T. *Remote Sensing of Climate*; Elsevier: Amsterdam, The Netherlands, 2024; 458p.
22. Vogt, L.; Baum, R.; Bhatty, P.; Kohler, C.; Meid, S.; Quast, B.; Grobe, P. SOCCOMAS: A FAIR web content management system that uses knowledge graphs and that is based on semantic programming. *Database* **2019**, *2019*, baz067. [CrossRef] [PubMed]
23. Destefanis, T.; Guliyeva, S.; Boccardo, P.; Fissore, V. Advancing flood detection and mapping: A review of Earth observation services, 3D data integration, and AI-based techniques. *Remote Sens.* **2025**, *17*, 2943. [CrossRef]
24. Mitchell, S.N.; Lahiff, A.; Cummings, N.; Hollocombe, J.; Boskamp, B.; Field, R.; Reddyhoff, D.; Zarebski, K.; Wilson, A.; Viola, B.; et al. FAIR data pipeline: Provenance-driven data management for traceable scientific workflows. *Philos. Trans. A Math. Phys. Eng. Sci.* **2022**, *380*, 20210300. [CrossRef] [PubMed]
25. Basereh, M.; Caputo, A.; Brennan, R. Automatic transparency evaluation for open knowledge extraction systems. *J. Biomed. Semant.* **2023**, *14*, 12. [CrossRef] [PubMed]
26. Gonzalez-Gil, P.; Martinez, J.A.; Skarmeta, A.F. Lightweight Data-Security Ontology for IoT. *Sensors* **2020**, *20*, 801. [CrossRef] [PubMed]
27. Nüst, D.; Pebesma, E. Practical Reproducibility in Geography and Geosciences. *Ann. Am. Assoc. Geogr.* **2020**, *111*, 1300–1310. [CrossRef]
28. Saleem, T.J.; Chishti, M.A. *Big Data Analytics for Internet of Things*; Wiley: Hoboken, NJ, USA, 2021; p. 376.
29. Radosevic, N.; Duckham, M.; Saiedur Rahaman, M.; Ho, S.; Williams, K.; Hashem, T.; Tao, Y. Spatial data trusts: An emerging governance framework for sharing spatial data. *Int. J. Digit. Earth* **2023**, *16*, 1607–1639. [CrossRef]
30. Lalli, G. Defining interoperability: A universal standard. *J. ICT Stand.* **2025**, *13*, 139–156. [CrossRef]
31. Zhao, J.; Miles, A.; Klyne, G.; Shotton, D. Linked data and provenance in biological data webs. *Brief. Bioinform.* **2009**, *10*, 139–152. [CrossRef] [PubMed]
32. Ciccarese, P.; Soiland-Reyes, S.; Belhajjame, K.; Gray, A.J.; Goble, C.; Clark, T. PAV ontology: Provenance, authoring and versioning. *J. Biomed. Semant.* **2013**, *4*, 37. [CrossRef] [PubMed]
33. Hassoun, A.E.R.; Tanhua, T.; Lips, I.; Heslop, E.; Petihakis, G.; Karstensen, J. The European Ocean Observing Community: Urgent gaps and recommendations to implement during the UN Ocean Decade. *Front. Mar. Sci.* **2024**, *11*, 1394984. [CrossRef]
34. Closa, G.; Masó, J.; Zabala, A.; Pesquer, L.; Pons, X. A provenance metadata model integrating ISO geospatial lineage and the OGC WPS: Conceptual model and implementation. *Trans. GIS* **2019**, *23*, 1102–1124. [CrossRef]
35. Sadiq, M.A.; Langat, P.K.; Neupane, A. GeoPROV: A domain-specialised provenance model for spatial data supply chains. *ISPRS Int. J. Geo-Inf.* **2026**, *15*, 272. [CrossRef]
36. He, L.; Yue, P.; Di, L.; Zhang, M.; Hu, L. Adding Geospatial Data Provenance into SDI—A Service-Oriented Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 926–936. [CrossRef]
37. Closa, G.; Masó, J.; Proß, B.; Pons, X. W3C PROV to describe provenance at the dataset, feature and attribute levels in a distributed environment. *Comput. Environ. Urban Syst.* **2017**, *64*, 103–117. [CrossRef]
38. Kumar, P.; Srivastava, P.K.; Khan, M.L.; Arunachalam, A.; Roy, P.S.; Kumar, K. (Eds.) *Advanced Geospatial and Ground Based Techniques in Forest Monitoring*; Earth Observation Series; Elsevier: Amsterdam, The Netherlands, 2025. [CrossRef]
39. Janowicz, K.; McKenzie, G.; Gao, S.; Hu, Y.; Bhaduri, B. GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *Int. J. Geogr. Inf. Sci.* **2020**, *34*, 625–636. [CrossRef]
40. Vogt, L. FAIR data representation in times of eScience: A comparison of instance-based and class-based semantic representations of empirical data using phenotype descriptions as example. *J. Biomed. Semant.* **2021**, *12*, 20. [CrossRef] [PubMed]
41. Carlos-Martinez, H.; Tapia-McClung, R.; Moctezuma-Ochoa, D.A.; Alegre-Mondragón, A.J. (Eds.) *Recent Developments in Geospatial Information Sciences: Selected Papers from iGISc 2023*; Lecture Notes in Geoinformation and Cartography; Springer Nature Switzerland AG: Cham, Switzerland, 2024; 254p. [CrossRef]
42. Buglisi, S.; De Luca, L.; Lauria, M.; Quattrocchi, A. Spatio-Temporal Data Integration for Knowledge-Driven Analysis of Historic Built Heritage Conservation State. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2025**, *XLVIII-M-9-2025*, 187–194. [CrossRef]
43. Wei, Y.; Zhang, C.; Yue, P.; Li, H. Geospatial knowledge cube: Enhancing semantic query in geospatial data cubes with ontology and LLMs. *Int. J. Geogr. Inf. Sci.* **2026**, *40*, 1–32. [CrossRef]

44. Open Geospatial Consortium. *OGC GeoSPARQL—A Geographic Query Language for RDF Data*, Version 1.1; OGC Standard 22-047r1; Open Geospatial Consortium: Arlington, VA, USA, 2024.
45. Car, N.J.; Homburg, T.; Perry, M.; Knibbe, F.; Cox, S.J.D.; Abhayaratna, J.; Bonduel, M.; Cripps, P.J.; Janowicz, K. (Eds.) *OGC GeoSPARQL—A Geographic Query Language for RDF Data*, Version 1.1; OGC Standard 22-047r1; Open Geospatial Consortium: Arlington, VA, USA, 2024.
46. Castano, S. *Advances in Conceptual Modeling: ER 2012 Workshops: CMS, ECDM-NoCoDA, MoDIC, MORE-BI, RIGiM, SeCoGIS, WISM, Florence, Italy, October 15–18, 2012: Proceedings*; Lecture Notes in Computer Science; Springer: Berlin, Germany; New York, NY, USA, 2012; Volume 7518, p. 130.
47. Meroño-Peñuela, A.; Simperl, E.; Kurteva, A.; Reklos, I. KG.GOV: Knowledge Graphs as the Backbone of Data Governance in AI. *J. Web Semant*. **2024**, *85*, 100847. [CrossRef]
48. Ranatunga, S.; Ødegård, R.S.; Jetlund, K.; Onstein, E. Use of Semantic Web Technologies to Enhance the Integration and Interoperability of Environmental Geospatial Data: A Framework Based on Ontology-Based Data Access. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 52. [CrossRef]
49. Řezník, T.; Raes, L.; Stott, A.; De Lathouwer, B.; Perego, A.; Charvát, K.; Kafka, Š. Improving the documentation and findability of data services and repositories: A review of (meta) data management approaches. *Comput. Geosci*. **2022**, *169*, 105194. [CrossRef]
50. Di, L.; Shao, Y.; Kang, L. Implementation of geospatial data provenance in a web service workflow environment with ISO 19115 and ISO 19115-2 lineage model. *IEEE Trans. Geosci. Remote Sens*. **2013**, *51*, 5082–5089. [CrossRef]
51. Ureña-Cámara, M.A.; Nogueras-Iso, J.; Lacasta, J.; Ariza-López, F.J. A method for checking the quality of geographic metadata based on ISO 19157. *Int. J. Geogr. Inf. Sci*. **2019**, *33*, 1–27. [CrossRef]
52. Reddy, P.K.; Sureka, A.; Chakravarthy, S.; Bhalla, S. *Big Data Analytics: 5th International Conference, BDA 2017, Hyderabad, India, December 12–15, 2017, Proceedings*; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2017; p. 12.
53. Janowicz, K.; Haller, A.; Cox, S.J.D.; Le Phuoc, D.; Lefrançois, M. SOSA: A Lightweight Ontology for Sensors, Observations, Samples, and Actuators. *J. Web Semant*. **2019**, *56*, 1–10. [CrossRef]
54. Milon-Flores, D.F.; Bernard, C.; Gensel, J.; Giuliani, G.; Chatenoux, B.; Dao, H. Towards semantic enrichment of Earth Observation data: The LEODS framework. *AGILE GISci. Ser.* **2024**, *5*, 11. [CrossRef]
55. World Wide Web Consortium (W3C). *RDF 1.1 Concepts and Abstract Syntax*; W3C Recommendation; World Wide Web Consortium: Cambridge, MA, USA, 2014. Available online: https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/ (accessed on 20 October 2026).
56. Chansanam, W.; Nguyen, L.T.; Li, C.; Soo Guan, C.K. Semantic Knowledge Graphs for Intercultural and Ethnographic Diversity in the Greater Mekong Subregion. *J. Intercult. Commun*. **2025**, *25*, 11–24. [CrossRef]
57. Yaman, B.; Thompson, K.; Fahey, F.; Brennan, R. LinkedDataOps: Quality oriented end-to-end geospatial linked data production governance. *Semant. Web* **2024**, *15*, 555–581. [CrossRef]
58. Hyvönen, E.; Tuominen, J. 8-star Linked Open Data Model: Extending the 5-star Model for Better Reuse, Quality, and Trust of Data. In Proceedings of the SEMANTiCS 2024 (Posters and Demos), Amsterdam, The Netherlands, 17–19 September 2024.
59. Trujillo, J. *Advances in Conceptual Modeling—Applications and Challenges: ER 2010 Workshops ACM-L, CMLSA, CMS, DE@ER, FP-UML, SeCoGIS, WISM, Vancouver, BC, Canada, November 1–4, 2010, Proceedings*; Information Systems and Applications, incl. Internet/Web, and HCI; Springer: Berlin/Heidelberg, Germany, 2010; 296p.
60. Shotton, D.; Portwin, K.; Klyne, G.; Miles, A. Adventures in semantic publishing: Exemplar semantic enhancements of a research article. *PLoS Comput. Biol*. **2009**, *5*, e1000361. [CrossRef] [PubMed]
61. Cudre-Mauroux, P. *The Semantic Web–ISWC 2012: 11th International Semantic Web Conference, Boston, MA, USA, November 11–15, 2012, Proceedings, Part I*; Information Systems and Applications, incl Inter-net/Web, and HCI; Springer: Berlin/Heidelberg, Germany, 2012; 677p.
62. Simeone, D.; Cursi, S.; Acierno, M. BIM semantic-enrichment for built heritage representation. *Autom. Constr.* **2019**, *97*, 122–137. [CrossRef]
63. Höltgen, L.; Zentgraf, S.; Hagedorn, P.; König, M. Utilizing large language models for semantic enrichment of infrastructure condition data: A comparative study of GPT and Llama models. *AI Civ. Eng*. **2025**, *4*, 1–34. [CrossRef]
64. Sadiq, M.A.; Langat, P.K.; Neupane, A. Enterprise spatial data provenance knowledge infrastructure. *ISPRS Int. J. Geo-Inf*. **2026**, *15*, 182. [CrossRef]
65. Harvey, F.; Leung, Y. *Advances in Spatial Data Handling and Analysis: Select Papers from the 16th IGU Spatial Data Handling Symposium*; Advances in Geographic Information Science; Springer: Cham, Switzerland, 2015; 319p.
66. Congalton, R.G. Assessing positional and thematic accuracies of maps generated from remotely sensed data. In *Remote Sensing Handbook, Volume I*; CRC Press: Boca Raton, FL, USA, 2024; pp. 465–493.
67. Johnson, M.; Dembska, M.; Schindler, S.; Bonnarel, F. Historical record of the data and its origins: Data provenance. In *Data-Intensive Radio Astronomy: Bringing Astrophysics to the Exabyte Era*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 377–396.

68. Herschel, M.; Diestelkämper, R.; Ben Lahmar, H. A survey on provenance: What for? What form? What from? *VLDB J.* **2017**, *26*, 881–906. [CrossRef]
69. McGlinn, K.; Rutherford, M.A.; Gisslander, K.; Hederman, L.; Little, M.A.; O'Sullivan, D. FAIRVASC: A semantic web approach to rare disease registry integration. *Comput. Biol. Med.* **2022**, *145*, 105313. [CrossRef] [PubMed]
70. Pan, B.; Stakhanova, N.; Ray, S. Data Provenance in Security and Privacy. *ACM Comput. Surv.* **2023**, *55*, 1–35. [CrossRef] [PubMed]
71. Bukhari, T.T.; Oladimeji, O.; Etim, E.D.; Ajayi, J.O. Systematic review of metadata-driven data orchestration in modern analytics engineering. *Gyanshauryam Int. Sci. Ref. Res. J.* **2022**, *5*, 536–564.
72. Fisher, M.; Mascardi, V.; Rozier, K.Y.; Schlingloff, B.-H.; Winikoff, M.; Yorke-Smith, N. Towards a framework for certification of reliable autonomous systems. *Auton. Agents Multi-Agent Syst.* **2021**, *35*, 8. [CrossRef]
73. Triapitcin, I. *Knowledge Injection Method for Real-Time Decision Support*; Acta Universitatis Lappeenrantaensis, No. 1097; LUT University Press: Lappeenranta, Finland, 2023.
74. Olaya, P.; Kennedy, D.; Llamas, R.; Valera, L.; Vargas, R.; Lofstead, J.; Taufer, M. Building trust in earth science findings through data traceability and results explainability. *IEEE Trans. Parallel Distrib. Syst.* **2022**, *34*, 704–717. [CrossRef]
75. Klump, J.; Wyborn, L.; Wu, M.; Martin, J.; Downs, R.R.; Asmi, A. Versioning data is about more than revisions: A conceptual framework and proposed principles. *Data Sci. J.* **2021**, *20*, 12. [CrossRef]
76. Whalley, W.B. Enhancing the Digital Earth via digital decimal geolocation and the FAIR data principles. *Earth Sci. Syst. Soc.* **2024**, *4*, 10110. [CrossRef]
77. Graser, A.; Sutton, T.; Bernasocchi, M. The QGIS project: Spatial without compromise. *Patterns* **2025**, *6*, 101265. [CrossRef] [PubMed]
78. Wilkinson, S.R.; Aloqalaa, M.; Belhajjame, K.; Crusoe, M.R.; de Paula Kinoshita, B.; Gadelha, L.; Garijo, D.; Gustafsson, O.J.R.; Juty, N.; Kanwal, S. Applying the FAIR principles to computational workflows. *Sci. Data* **2025**, *12*, 328. [CrossRef] [PubMed]
79. Sudmanns, M.; Augustin, H.; Killough, B.; Giuliani, G.; Tiede, D.; Leith, A.; Yuan, F.; Lewis, A. Think global, cube local: An Earth Observation Data Cube's contribution to the Digital Earth vision. *Big Earth Data* **2023**, *7*, 831–859. [CrossRef]
80. Hojati, M.; Robertson, C.; Roberts, S.; Chaudhuri, C. GIScience research challenges for realizing discrete global grid systems as a Digital Earth. *Big Earth Data* **2022**, *6*, 358–379. [CrossRef]
81. Wang, X.; Wang, Y.; Ambite, J.L.; Appaji, A.; Lander, H.; Moore, S.M.; Rajasekar, A.K.; Turner, J.A.; Turner, M.D.; Wang, L.; et al. Enabling Scientific Reproducibility through FAIR Data Management: An ontology-driven deep learning approach in the NeuroBridge Project. *AMIA Annu. Symp. Proc.* **2022**, *2022*, 1135–1144. [PubMed]
82. Prudhomme, T.; De Colle, G.; Liebers, A.; Sculley, A.; Xie, P.K.; Cohen, S.; Beverley, J. A semantic approach to mapping the Provenance Ontology to Basic Formal Ontology. *Sci. Data* **2025**, *12*, 282. [CrossRef] [PubMed]
83. Parikh, P.P.; Minning, T.A.; Nguyen, V.; Lalithsena, S.; Asiaee, A.H.; Sahoo, S.S.; Doshi, P.; Tarleton, R.; Sheth, A.P. A semantic problem solving environment for integrative parasite research: Identification of intervention targets for Trypanosoma cruzi. *PLoS Neglect. Trop. Dis.* **2012**, *6*, e1458. [CrossRef] [PubMed]
84. Vos, R.A.; Biserkov, J.V.; Balech, B.; Beard, N.; Blissett, M.; Brenninkmeijer, C.; van Dooren, T.; Eades, D.; Gosline, G.; Groom, Q.J.; et al. Enriched biodiversity data as a resource and service. *Biodivers. Data J.* **2014**, *2*, e1125. [CrossRef] [PubMed]
85. Gierend, K.; Kruger, F.; Genehr, S.; Hartmann, F.; Siegel, F.; Waltemath, D.; Ganslandt, T.; Zeleke, A.A. Provenance Information for Biomedical Data and Workflows: Scoping Review. *J. Med. Internet Res.* **2024**, *26*, e51297. [CrossRef] [PubMed]
86. Klingner, C.M.; Denker, M.; Grün, S.; Hanke, M.; Oeltze-Jafra, S.; Ohl, F.W.; Radny, J.; Rotter, S.; Scherberger, H.; Stein, A.; et al. Research data management and data sharing for reproducible research—Results of a community survey of the German National Research Data Infrastructure Initiative Neuroscience. *eNeuro* **2023**, *10*. [CrossRef] [PubMed]
87. Hastings, J.; Chepelev, L.; Willighagen, E.; Adams, N.; Steinbeck, C.; Dumontier, M. The chemical information ontology: Provenance and disambiguation for chemical data on the biological semantic web. *PLoS ONE* **2011**, *6*, e25513. [CrossRef] [PubMed]
88. Nguyen, V.; Bodenreider, O.; Sheth, A. Don't Like RDF Reification? Making Statements about Statements Using Singleton Property. In Proceedings of the 23rd International Conference on World Wide Web, Seoul, Republic of Korea, 7–11 April 2014; pp. 759–770. [CrossRef] [PubMed]
89. Sahoo, S.S.; Nguyen, V.; Bodenreider, O.; Parikh, P.; Minning, T.; Sheth, A.P. A unified framework for managing provenance information in translational research. *BMC Bioinform.* **2011**, *12*, 461. [CrossRef] [PubMed]
90. Khan, F.Z.; Soiland-Reyes, S.; Sinnott, R.O.; Lonie, A.; Goble, C.; Crusoe, M.R. Sharing interoperable workflow provenance: A review of best practices and their practical application in CWLProv. *Gigascience* **2019**, *8*, giz095. [CrossRef] [PubMed]
91. Closa, G.; Masó, J.; Pesquer, L.; Pons, X. Auditing Remote Sensing Data Using Geospatial Provenance. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; pp. 4253–4256.
92. Closa, G.; Masó, J.; Julià, N.; Pons, X. Geospatial Queries on Data Collection Using a Common Provenance Model. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 139. [CrossRef]

93. Fonseca, F.; Rodríguez, M.A.; Levashkin, S. *GeoSpatial Semantics: Second International Conference, GeoS 2007, Mexico City, Mexico, November 29–30, 2007: Proceedings*; Information Systems and Applications, incl Internet/Web, and HCI 4853; Springer: New York, NY, USA, 2007; 289p.
94. Meersman, R.; Panetto, H.; Dillon, T.; Missikoff, M.; Liu, L.; Pastor, O.; Cuzzocrea, A.; Sellis, T. *On the Move to Meaningful Internet Systems: OTM 2014 Conferences: Confederated International Conferences: CoopIS and ODBASE 2014, Amantea, Italy, October 27–31, 2014. Proceedings*; Information Systems and Applications, incl Internet/Web, and HCI 8841; Springer: Berlin/Heidelberg, Germany, 2014; 804p.
95. Elshehaly, M.; Jianu, R.; Slingsby, A.; Andrienko, G.; Andrienko, N.; Rhyne, T.M. Designing for Collaboration: Visualization to Enable Human-LLM Analytical Partnership. *IEEE Comput. Graph. Appl.* **2025**, *45*, 107–116. [CrossRef] [PubMed]
96. Sahoo, S.S.; Bodenreider, O.; Hitzler, P.; Sheth, A.; Thirunarayan, K. Provenance Context Entity (PaCE): Scalable Provenance Tracking for Scientific RDF Data. In *Scientific and Statistical Database Management*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6187, pp. 461–470. [CrossRef] [PubMed]
97. Guitton, F.; Oehmichen, A.; Bossé, É.; Guo, Y. Honest Computing: Achieving demonstrable data lineage and provenance for driving data and process-sensitive policies. *Data Policy* **2024**, *6*, e84. [CrossRef]
98. Verma, R.; Shrivastava, P.; Merla, N. Tracing the path: Data lineage and its impact on data governance. *Int. J. Glob. Innov. Solut.* **2024**. [CrossRef]
99. Hofmeister, M.; Bai, J.; Brownbridge, G.; Mosbach, S.; Lee, K.F.; Farazi, F.; Hillman, M.; Agarwal, M.; Ganguly, S.; Akroyd, J. Semantic agent framework for automated flood assessment using dynamic knowledge graphs. *Data-Centric Eng*. **2024**, *5*, e14. [CrossRef]
100. Kale, A.; Nguyen, T.; Harris, F.C.; Li, C.; Zhang, J.; Ma, X. Provenance documentation to enable explainable and trustworthy AI: A literature review. *Data Intell*. **2023**, *5*, 139–162. [CrossRef]
101. Chen, C.; Huang, H.; Ross, K.E.; Cowart, J.E.; Arighi, C.N.; Wu, C.H.; Natale, D.A. Protein ontology on the semantic web for knowledge discovery. *Sci. Data* **2020**, *7*, 337. [CrossRef] [PubMed]
102. Sahoo, S.S.; Valdez, J.; Kim, M.; Rueschman, M.; Redline, S. ProvCaRe: Characterizing scientific reproducibility of biomedical research studies using semantic provenance metadata. *Int. J. Med. Inform.* **2019**, *121*, 10–18. [CrossRef] [PubMed]
103. Sahoo, S.S.; Valdez, J.; Rueschman, M. Scientific Reproducibility in Biomedical Research: Provenance Metadata Ontology for Semantic Annotation of Study Description. *AMIA Annu. Symp. Proc.* **2016**, *2016*, 1070–1079. [PubMed]
104. Gandon, F.; Guéret, C.; Villata, S.; Breslin, J.; Faron-Zucker, C.; Zimmermann, A. *The Semantic Web: ESWC 2015 Satellite Events: ESWC 2015 Satellite Events, Portorož, Slovenia, May 31–June 4, 2015, Revised Selected Papers*; Computer Communication Networks and Telecommunications 9341; Springer: Cham, Switzerland, 2015; 471p.
105. Soranno, P.A.; Bissell, E.G.; Cheruvelil, K.S.; Christel, S.T.; Collins, S.M.; Fergus, C.E.; Filstrup, C.T.; Lapierre, J.F.; Lottig, N.R.; Oliver, S.K.; et al. Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. *Gigascience* **2015**, *4*, 28. [CrossRef] [PubMed]
106. Li, H.; Yue, P.; Wu, H.; Teng, B.; Zhao, Y.; Liu, C. A question-answering framework for geospatial data retrieval enhanced by a knowledge graph and large language models. *Int. J. Digit. Earth* **2025**, *18*, 2510566. [CrossRef]
107. Liu, Y.; Wang, X.; Wang, Y.; Huang, F.; Huang, Y.; Li, Y.; Zhang, W.; Gong, S.; Mai, G.; Yao, Y.; et al. Representation learning for geospatial data. *Ann. GIS* **2025**, *31*, 557–583. [CrossRef]
108. Wang, L.; Ambite, J.L.; Appaji, A.; Bijsterbosch, J.; Dockes, J.; Herrick, R.; Kogan, A.; Lander, H.; Marcus, D.; Moore, S.M.; et al. NeuroBridge: A prototype platform for discovery of the long-tail neuroimaging data. *Front. Neuroinform.* **2023**, *17*, 1215261. [CrossRef] [PubMed]
109. Yue, Z.; Willey, C.D.; Hjelmeland, A.B.; Chen, J.Y. BEERE: A web server for biomedical entity expansion, ranking and explorations. *Nucleic Acids Res*. **2019**, *47*, W578–W586. [CrossRef] [PubMed]
110. Bolling, C.; Weidlich, M.; Holzhutter, H.G. SEE: Structured representation of scientific evidence in the biomedical domain using Semantic Web techniques. *J. Biomed. Semant*. **2014**, *5*, S1. [CrossRef] [PubMed]
111. Chalk, S.J. SciData: A data model and ontology for semantic representation of scientific data. *J. Cheminform*. **2016**, *8*, 54. [CrossRef] [PubMed]
112. Fu, G.; Batchelor, C.; Dumontier, M.; Hastings, J.; Willighagen, E.; Bolton, E. PubChemRDF: Towards the semantic annotation of PubChem compound and substance databases. *J. Cheminform*. **2015**, *7*, 34. [CrossRef] [PubMed]
113. Rejeb, A.; Keogh, J.G.; Martindale, W.; Dooley, D.; Smart, E.; Simske, S.; Wamba, S.F.; Breslin, J.G.; Bandara, K.Y.; Thakur, S. Charting past, present, and future research in the semantic web and interoperability. *Future Internet* **2022**, *14*, 161. [CrossRef]
114. Lebo, T.; Sahoo, S.; McGuinness, D.; Belhajjame, K.; Cheney, J.; Corsar, D.; Garijo, D.; Soiland-Reyes, S.; Zednik, S.; Zhao, J. *PROV-O: The PROV Ontology*; W3C Recommendation; World Wide Web Consortium (W3C): Cambridge, MA, USA, 2013. Available online: https://www.w3.org/TR/2013/REC-prov-o-20130430/ (accessed on 20 October 2026).
115. World Wide Web Consortium (W3C). *PROV-O: The PROV Ontology*; W3C Recommendation; World Wide Web Consortium: Cambridge, MA, USA, 2013.

116. Zhang, F.; Lu, Q.; Du, Z.; Chen, X.; Cao, C. A comprehensive overview of RDF for spatial and spatiotemporal data management. *Knowl. Eng. Rev.* **2021**, *36*, e10. [CrossRef]
117. Wiemann, S.; Bernard, L. Spatial data fusion in Spatial Data Infrastructures using Linked Data. *Int. J. Geogr. Inf. Sci.* **2015**, *30*, 613–636. [CrossRef]
118. Information Resources Management Association; Information Resources Management Association. *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, 2019; Volume 93.
119. Stevens, R.D.; Tipney, H.J.; Wroe, C.J.; Oinn, T.M.; Senger, M.; Lord, P.W.; Goble, C.A.; Brass, A.; Tassabehji, M. Exploring Williams-Beuren syndrome using myGrid. *Bioinformatics* **2004**, *20*, i303–i310. [CrossRef] [PubMed]
120. Altintas, I.; Berkley, C.; Jaeger, E.; Jones, M.; Ludäscher, B.; Mock, S. Kepler: An extensible system for design and execution of scientific workflows. In Proceedings of the 16th International Conference on Scientific and Statistical Database Management (SSDBM 2004), Santorini Island, Greece, 21–23 June 2004; pp. 423–424. [CrossRef]
121. Simmhan, Y.L.; Plale, B.; Gannon, D. A framework for collecting provenance in data-centric scientific workflows. In Proceedings of the 2006 IEEE International Conference on Web Services (ICWS'06), Chicago, IL, USA, 18–22 September 2006; pp. 427–436.
122. Paterson, T.; Law, A. An XML transfer schema for exchange of genomic and genetic mapping data: Implementation as a web service in a Taverna workflow. *BMC Bioinform.* **2009**, *10*, 252. [CrossRef] [PubMed]
123. Belhajjame, K.; Cheney, J.; Coppens, S.; Garijo, D.; Soiland-Reyes, S. PROV Specification: A Provenance Model for Open Distributed Systems. *IEEE Internet Comput.* **2013**, *17*, 15–26.
124. Missier, P.; Belhajjame, K.; Cheney, J. The W3C PROV family of specifications for modelling provenance metadata. *Future Gener. Comput. Syst.* **2013**, *29*, 1203–1215.
125. Leipzig, J.; Nust, D.; Hoyt, C.T.; Ram, K.; Greenberg, J. The role of metadata in reproducible computational research. *Patterns* **2021**, *2*, 100322. [CrossRef] [PubMed]
126. Omidi, H.; Sacco, L.; Hutter, V.; Irsiegler, G.; Claus, M.; Schobben, M.; Jacob, A.; Schramm, M.; Fiore, S. Towards provenance-aware Earth observation workflows: The openEO case study. In *Proceedings of the 2025 IEEE International Conference on eScience (eScience), Chicago, IL, USA, 15–18 September 2025*; IEEE: Piscataway, NJ, USA, 2025; pp. 58–66. [CrossRef]
127. Panda, S.S.; Rao, M.N.; Thenkabail, P.S.; Misra, D.; Fitzgerald, J.P. Remote sensing systems—Platforms and sensors: Aerial, satellite, UAV, optical, radar, and LiDAR. In *Remote Sensing Handbook, Volume I: Sensors, Data Normalization, Harmonization, Cloud Computing, and Accuracies*, 2nd ed.; Thenkabail, P.S., Ed.; CRC Press: Boca Raton, FL, USA, 2024; pp. 3–86. [CrossRef]
128. Garlan, T.; Almar, R.; Bergsma, E.W.J. Challenges and opportunities in predicting future beach evolution: A review of processes, remote sensing, and modeling approaches. *Remote Sens.* **2025**, *17*, 3360. [CrossRef]
129. Mai, G.; Xie, Y.; Jia, X.; Lao, N.; Rao, J.; Zhu, Q.; Liu, Z.; Chiang, Y.-Y.; Jiao, J. Towards the next generation of Geospatial Artificial Intelligence. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *136*, 104368. [CrossRef]
130. Wu, J.; Orlandi, F.; O'Sullivan, D.; Dev, S. LinkedGeoClimate: An Interoperable Platform for Climate Data Access Within Geographical Context. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 10352–10366. [CrossRef]
131. Tran, B.-H.; Aussenac-Gilles, N.; Comparot, C.; Trojahn, C. Semantic Integration of Raster Data for Earth Observation on Territorial Units. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 149. [CrossRef]
132. Melet, A.; Teatini, P.; Le Cozannet, G.; Jamet, C.; Conversi, A.; Benveniste, J.; Almar, R. Earth observations for monitoring marine coastal hazards and their drivers. *Surv. Geophys.* **2020**, *41*, 1489–1534. [CrossRef]
133. Mushtaq, F.; Alhems, L.M.; Farooq, M.; Tiwari, S.P. A comprehensive review of geomatics based coastal zone management in the Realm of Arabian Gulf, Saudi Arabia. *Discov. Appl. Sci.* **2025**, *7*, 140. [CrossRef]
134. Tan, H.; Yue, P.; Yu, D.; Boyi, S.; Cao, Z. A provenance-aware method for updating Earth Observation training data. In Proceedings of the 2022 10th International Conference on Agro-geoinformatics (Agro-Geoinformatics), Quebec City, QC, Canada, 11–14 July 2022; pp. 1–6.
135. Sikos, L.F.; Philp, D. Provenance-aware knowledge representation: A survey of data models and contextualized knowledge graphs. *Data Sci. Eng.* **2020**, *5*, 293–316. [CrossRef]
136. Groth, P.; Gil, Y. Tracking provenance in knowledge graphs. *IEEE Internet Comput.* **2012**, *16*, 32–39.
137. Kovanič, Ľ.; Topitzer, B.; Peťovský, P.; Blišťan, P.; Gergeľová, M.B.; Blišťanová, M. Review of photogrammetric and lidar applications of UAV. *Appl. Sci.* **2023**, *13*, 6732. [CrossRef]
138. Jessin, J.; Heinzlef, C.; Long, N.; Serre, D. A systematic review of UAVs for island coastal environment and risk monitoring: Towards a Resilience Assessment. *Drones* **2023**, *7*, 206. [CrossRef]
139. Ho, S.-W.; Wu, J.-Y.; Chen, Y.-W.; Yang, C.-K.; Tsai, W.-P. Leveraging artificial intelligence (AI) techniques for sustainable marine resources. *Blue Biotechnol.* **2026**, *3*, 4. [CrossRef]
140. Christofi, D.; Mettas, C.; Evagorou, E.; Stylianou, N.; Eliades, M.; Theocharidis, C.; Chatzipavlis, A.; Hasiotis, T.; Hadjimitsis, D. A review of open remote sensing data with GIS, AI, and UAV support for shoreline detection and coastal erosion monitoring. *Appl. Sci.* **2025**, *15*, 4771. [CrossRef]

141. Zhu, J.; Dang, P.; Cao, Y.; Lai, J.; Guo, Y.; Wang, P.; Li, W. A flood knowledge-constrained large language model interactable with GIS: Enhancing public risk perception of floods. *Int. J. Geogr. Inf. Sci.* **2024**, *38*, 603–625. [CrossRef]
142. Panda, U.S.; Pradhan, U.; Naik, S.; Kumaresan, D.; Pradhan, S.; Sahoo, R.K.; Ramanamurthy, M. Comprehensive Framework for Monitoring of Seas and Coastal Ecosystems Under Changing Environments. In *Monitoring Coastal Marine Ecosystems Under Climate Change: Machine Learning and Artificial Intelligence Techniques*; Springer: Berlin/Heidelberg, Germany, 2026; pp. 81–139.
143. Biliani, I.; Skamnia, E.; Economou, P.; Zacharias, I. A Novel Methodology to Correct Chlorophyll-a Concentrations from Satellite Data and Assess Credible Phenological Patterns. *Remote Sens.* **2025**, *17*, 1156. [CrossRef]
144. Ansari, M.; Knudby, A.; Amani, M.; Sawada, M. Retrieving inland water quality parameters via satellite remote sensing: Sensor evaluation, atmospheric correction, and machine learning approaches. *Remote Sens.* **2025**, *17*, 1734. [CrossRef]
145. Lopes, C.L.; Mendes, R.; Azevedo, L.; Sousa, M.C.; Ribeiro, A.S.; Ferreira, A.; Biguino, B.; Brito, A.C.; Dias, J.M. Remote Sensing Technologies in Aerospace: Applications and Opportunities. In *Advanced Optical Sensors for Aerospace Applications*; Springer: Singapore, 2025; pp. 285–333.
146. Sun, S.; Xue, Q.; Xing, X.; Zhao, H.; Zhang, F. Remote sensing image interpretation for coastal zones: A review. *Remote Sens.* **2024**, *16*, 4701. [CrossRef]
147. Trinh, H.L.; Lim, J.X.; Kieu, H.T.; Er, J.W.; Pang, D.S.C.; Law, A.W.-K. Multisensor fusion for monitoring shorelines and coastal environments: A review and perspective with satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2026**, *19*, 10910–10935. [CrossRef]
148. Graser, A.; Wachsenegger, A.; Doulkeridis, C.; Theodoropoulos, G.; Salehi, B.; Dragaschnig, M.; Antoniou, S.; Theodoridis, Y. Spatiotemporal Model Cards Enabling Future-Proof GeoAI Systems. *ACM Trans. Spat. Algorithms Syst.* **2026**, *12*, 1–39. [CrossRef]
149. Ding, H.; Zhai, J.; Deng, D.; Ma, S. The case for learned provenance graph storage systems. In *Proceedings of the 32nd USENIX Security Symposium (USENIX Security 23), Anaheim, CA, USA, 9–11 August 2023*; USENIX Association: Berkeley, CA, USA, 2023; ISBN 978-1-939133-37-3, pp. 3277–3294. Available online: https://www.usenix.org/conference/usenixsecurity23/presentation/ding-hailun-provenance (accessed on 27 August 2026).
150. Janowicz, K.; Hitzler, P.; Li, W.; Rehberger, D.; Schildhauer, M.; Zhu, R.; Shimizu, C.; Fisher, C.; Cai, L.; Mai, G. Know, Know Where, KnowWhereGraph: A densely connected, cross-domain knowledge graph and geo-enrichment service stack for applications in environmental intelligence. *AI Mag.* **2022**, *43*, 30–39. [CrossRef]
151. Wang, S.; Huang, X.; Liu, P.; Zhang, M.; Biljecki, F.; Hu, T.; Fu, X.; Liu, L.; Liu, X.; Wang, R.; et al. Mapping the landscape and roadmap of geospatial artificial intelligence (GeoAI) in quantitative human geography: An extensive systematic review. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *128*, 103734. [CrossRef]
152. Homoky, W.B.; Weber, T.; Berelson, W.M.; Conway, T.M.; Henderson, G.M.; van Hulten, M.; Jeandel, C.; Severmann, S.; Tagliabue, A. Quantifying trace element and isotope fluxes at the ocean-sediment boundary: A review. *Philos. Trans. A Math. Phys. Eng. Sci.* **2016**, *374*, 20160246. [CrossRef] [PubMed]
153. Lymburner, L.; Bunting, P.; Lucas, R.; Scarth, P.; Alam, I.; Phillips, C.; Ticehurst, C.; Held, A. Mapping the multi-decadal mangrove dynamics of the Australian coastline. *Remote Sens. Environ.* **2020**, *238*, 111185. [CrossRef]
154. Simmhan, Y.L.; Plale, B.; Gannon, D. Karma2: Provenance management for data-driven workflows. *Int. J. Web Serv. Res.* **2008**, *5*, 1–22.
155. Auge, T.; Feistel, S.; Ekaputra, F.J.; Klettke, M.; Jürgensmann, S.; Michels, E.; Waltersdorfer, L. Towards an Integrated Provenance Framework: A Scenario for Marine Data. In Proceedings of the 2024 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Vienna, Austria, 7–12 July 2024; pp. 597–601.
156. Sun, W.; Chen, C.; Liu, W.; Yang, G.; Meng, X.; Wang, L.; Ren, K. Coastline extraction using remote sensing: A review. *GISci. Remote Sens.* **2023**, *60*, 2243671. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.