Article

Not peer-reviewed version

Rethinking Spatial Data Quality: A Socio-Technical and Lifecycle Framework

Tomaž Podobnikar *

Posted Date: 11 March 2026

doi: 10.20944/preprints202603.0798.v1

Keywords: spatial data quality (SDQ); geospatial information systems (GIS); data governance; data lifecycle management; artificial intelligence (AI); visual analytics; spatial data integration; data provenance and lineage; perceived data quality; socio-technical systems

Article

Rethinking Spatial Data Quality: A Socio-Technical and Lifecycle Framework

Tomaž Podobnikar [1,2]

[1] Ministry of Natural Resources and Spatial Planning, 1000 Ljubljana, Slovenia; tomaz.podobnikar@gov.si
[2] Faculty of Information Studies in Novo mesto, 8000 Novo mesto, Slovenia

Abstract

Spatial data quality (SDQ) is commonly assessed through technical verification. However, empirical evidence demonstrates that perceived data quality often diverges from objectively measured quality due to cognitive, institutional, and lifecycle-related factors. This paper proposes a multi-layered SDQ framework that integrates technical admissibility, process and lifecycle stewardship, visual and interpretive diagnostics, and governance indicators to enable holistic quality assessment within a socio-technical system. Rather than treating quality elements in isolation, the framework supports the diagnosis of emergent quality states and associated risk patterns. The framework is demonstrated through two empirical cases: validation of planned land use data using the OPIAvalid toolkit, and semantic conflation of multiple digital elevation models (DEMs) with heterogeneous lineage. Results show that governance failures, specification misuse, and degradation of lineage can undermine trust and decision-making even when datasets formally comply with ISO-based indicators. Visual spatial forensics and lineage-aware integration proved essential for detecting undocumented methodological shortcuts and restoring justified trust in authoritative data. Artificial intelligence is positioned as a diagnostic and explanatory support, assisting in anomaly detection, prioritization, and communication of quality risks, while deterministic validation and expert judgment remain mandatory. Overall, the framework shifts SDQ management from isolated technical validation toward lifecycle-oriented, transparent, and sustainable data governance.

Keywords: spatial data quality (SDQ); geospatial information systems (GIS); data governance; data lifecycle management; artificial intelligence (AI); visual analytics; spatial data integration; data provenance and lineage; perceived data quality; socio-technical systems

1. Introduction

In the geospatial domain, spatial data quality (SDQ) is fundamental to the reliability of digital representations used in high-stakes decision-making, environmental governance, and sustainable development. Errors or misinterpretations in spatial data can propagate directly into planning failures, legal disputes, and policy misjudgments. Historically, SDQ was treated as a reactive concern, becoming visible primarily through failure, when cartographic products misrepresent reality, analytical outputs contradicted observable conditions, or decisions produced unintended consequences.

This reactive paradigm gradually evolved into operational definitions centered on the principle of fitness for use, shifting attention from intrinsic data properties toward suitability for specific application contexts [1]. Contemporary geospatial information science conceptualizes SDQ through multidimensional elements such as positional accuracy, thematic completeness, logical consistency, and temporal quality. These dimensions are formally specified and evaluated within international standards, most notably ISO 19157-1, which defines quality elements, evaluation procedures, and reporting mechanisms for geographic data [2].

Despite this formalization, a persistent and increasingly visible trust gap remains. Datasets that demonstrate full technical compliance with established standards often fail to support effective decision-making in practice. Conversely, datasets with known technical limitations are frequently trusted, reused, and defended in high-stakes contexts due to institutional provenance, visual coherence, or historical acceptance. This discrepancy exposes a fundamental limitation of traditional dataset-centric quality assessment: technical correctness alone cannot explain whether spatial data are perceived as credible, defensible, or operationally actionable within real governance environments.

This limitation has intensified with the proliferation of spatial analytics, digital twins, and AI-assisted decision-making. In these highly integrated ecosystems, SDQ is no longer a downstream concern addressed after data acquisition or integration. Instead, it becomes a precondition for accountability and trust. Errors propagate rapidly across interconnected systems, while visually plausible outputs may conceal structural, semantic, or topological flaws that remain undetected by conventional quality metrics.

Recent critiques argue that existing standards must evolve to remain relevant, particularly in the context of big data, building information modeling (BIM), and AI-driven workflows [3]. These critiques emphasize that conventional quality models cannot operate reliably at the instance level, ensure traceability across heterogeneous sources, or support lifecycle-aware accountability. Together, these developments indicate that SDQ cannot be understood solely as a post-production audit of technical attributes, but must be conceptualized as a socio-technical property emerging from the interaction of technical compliance, human cognition, and institutional governance.

In response to this challenge, this paper proposes a holistic framework that redefines SDQ as a socio-technical, lifecycle-oriented system property rather than a static attribute of an isolated dataset. We argue that trust does not reside in data itself. Instead, it emerges from sustained alignment between measurement practices, data representation, institutional provenance, and explanatory context. By integrating geospatial standards with insights from cognitive science, governance studies, and visual analytics, this work establishes conceptual foundations for visual spatial forensics, circular data stewardship, and emergent trust-based quality assessment. Unlike existing SDQ models that treat trust as a byproduct of technical compliance or metadata completeness, the proposed framework explicitly conceptualizes trust as an emergent socio-technical property arising from interactions across the data lifecycle.

2. From Technical Compliance to Socio-Technical Spatial Data Quality

While technical standards and quantitative metrics are essential for establishing baseline admissibility of spatial data, they are insufficient for explaining how SDQ is interpreted, trusted, and acted upon in real-world decision-making. Formally compliant datasets may nevertheless fail in practice, while technically limited datasets may remain authoritative, revealing a persistent gap between actual and perceived quality.

This section argues that SDQ cannot be understood as a purely technical property of datasets but must be conceptualized as a socio-technical outcome shaped by human cognition, institutional context, and governance practices across the data lifecycle. Section 2.1 reviews the technical and epistemological foundations of SDQ. Section 2.2 introduces the critical distinction between actual and perceived spatial data quality, which provides the analytical lens for the remainder of the section. Sections 2.3 and 2.4 examine cognitive and institutional mechanisms that systematically shape perceived quality, while Section 2.5 synthesizes these mechanisms by conceptualizing trust as an emergent socio-technical property.

Together, these perspectives motivate the need for a lifecycle-oriented, trust-aware SDQ framework, which is introduced conceptually here and formalized in Section 3. Figure 1 summarizes the key dimensions discussed in Sections 2.1–2.5.

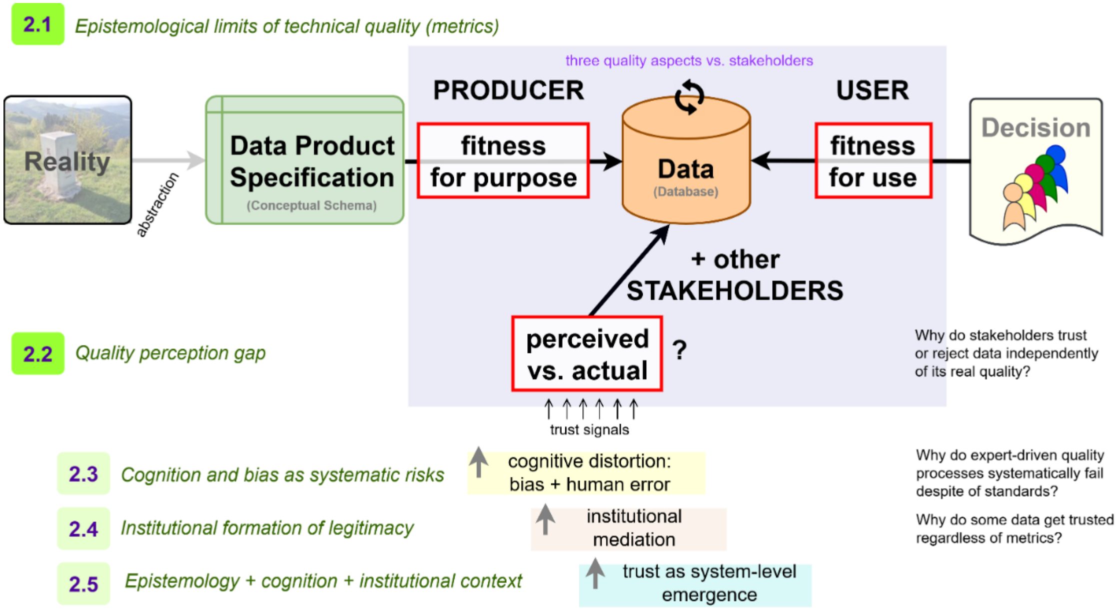


Figure 1. Scheme of the main points of the Sections 2.1–2.5. The three aspects of SDQ in relation to stakeholders are in red frames.

2.1. Technical and Epistemological Foundations of Spatial Data Quality

Classical foundations of data quality are rooted in technical and managerial traditions such as total data quality management (TDQM) and Six Sigma, which conceptualize quality as conformance to predefined requirements [4,5]. In this context, the term total refers to company-wide quality management rather than to an absolute notion of quality. Within these paradigms, quality is treated as an objective and measurable property of a data artifact. While this perspective remains essential for ensuring baseline consistency, comparability, and interoperability, it is increasingly insufficient for explaining how spatial data function as knowledge in complex decision-making environments. In particular, it does not account for how abstraction choices, modeling assumptions, and intended use shape what counts as "quality" in practice.

Contemporary data quality scholarship extends this view by situating quality within an epistemological framework, in which data are understood as representations of selected aspects of geographic reality rather than neutral reflections of it. From this perspective, SDQ emerges from the interaction between observable data characteristics, such as spatial resolution, scale, positional uncertainty, and semantic classification, and the interpretive contexts through which geospatial data acquire meaning. Quality is therefore not an intrinsic attribute of a data artifact but a relational and knowledge-dependent construct, shaped by modeling assumptions, abstraction choices, analytical purpose, and interpretive frames. Ultimately, what matters is not only whether spatial data conform to formal specifications, but whether they support justified spatial claims and defensible decisions within a given geographic, temporal, and institutional context.

This epistemological tension is reflected in two complementary perspectives that dominate modern quality frameworks. From the producer's perspective, quality is defined as fitness for purpose. It measures the extent to which data conforms to the goals, assumptions, and requirements (including legislative ones), encoded in a data product specification (DPS). This view is formally codified in ISO 19157-1:2023, which defines quality as the degree to which a dataset appropriately represents selected aspects of reality within a defined universe of discourse [2]. Quality of geographic data is evaluated through measurable elements such as completeness, logical consistency, positional accuracy, thematic accuracy, and temporal quality, which together describe intrinsic technical quality [6].

From the user or business perspective, quality is defined as fitness for use. In this context, data are evaluated according to their suitability for a specific decision-making situation, regardless of whether that situation aligns with the producer's original intent [7]. This distinction is particularly salient in the context of big data and open data, where spatial datasets are routinely reused, repurposed, and integrated across domains and scales [8]. Under these conditions, data quality is not a static property intrinsic to a dataset, but a relational and context-dependent construct [6]. A dataset may be technically sound from the producer's perspective, yet inadequate from the user's perspective if it lacks the contextual, temporal, or semantic information required for a new analytical or operational task [9].

Technical compliance therefore constitutes a necessary but insufficient condition for SDQ. It establishes a baseline of admissibility, but it does not explain whether spatial data will be interpreted correctly, trusted appropriately, or used responsibly once they enter broader socio-technical systems.

2.1.1. Generic Example: Fitness for Purpose vs. Fitness for Use (Producer–User Misalignment)

A national mapping agency produces a land use/land cover dataset optimized for national reporting. The dataset is defined by a minimum mapping unit of 1 ha, standardized thematic classes, and a five-year update cycle, as specified in the DPS. Quality evaluation confirms high completeness, logical consistency, and thematic accuracy within the declared universe of discourse, fully satisfying ISO 19157-1 requirements. When the same dataset is reused for regional flood-risk modeling, however, it fails to represent small hydrological features and recent urban development that are critical to the new application. Although the dataset remains technically compliant and fully fit for the producer's declared purpose, its spatial resolution, update frequency, and thematic structure are insufficient for the hydrological modeling task.

This example illustrates how a dataset may satisfy formal technical quality requirements while remaining unsuitable for a different analytical or decision-making context. It therefore highlights the distinction between fitness for purpose, defined by the producer's specification, and fitness for use, which depends on the requirements of a specific analytical or decision-making context.

2.2. Actual and Perceived Quality

A critical but often implicit distinction in SDQ discourse is the difference between actual (objective) and perceived SDQ [10] (Figure 2). Actual quality refers to measurable compliance with explicitly defined specifications, such as accuracy thresholds, logical consistency, or topological rules derived from a DPS or standards such as ISO 19157-1. In contrast, perceived quality reflects a stakeholder's belief that a dataset is reliable, authoritative, or appropriate for a given purpose [11]. Stakeholders may include data producers, custodians, stewards, and users from diverse professional and institutional backgrounds, each applying different interpretive frames when assessing quality.

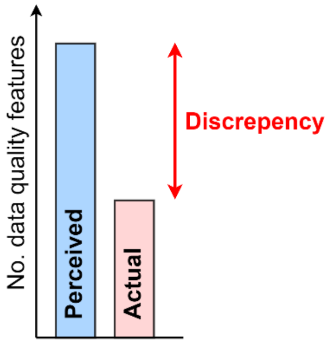


Figure 2. Perceived vs. actual data quality and discrepancy as data quality gap.

Perceived quality is shaped by trust signals that are independent of formal technical evidence. These signals include institutional provenance, historical acceptance, reputational authority, and visual coherence. In complex systems, validation processes are often opaque, fragmented, or delegated to external actors – stakeholders. As a result, datasets whose actual quality is poorly understood may be widely trusted, while technically robust datasets may be questioned or rejected if they conflict with established practices or appear visually implausible.

This divergence between actual and perceived quality is not always accidental. It is often intentional – driven by specific purposes – and becomes particularly salient in high-stakes or adversarial contexts.

Analogous to distinctions in assessment science between clinical and forensic settings, technically compliant evaluations may fail when decision-making priorities shift from accuracy toward persuasion, defensibility, or risk avoidance [12]. In spatial planning, legal disputes, or environmental regulation, for example, datasets may satisfy formal technical criteria (e.g., where dataset might pass all the official technical tests) while lacking the contextual validity required for accountability or institutional scrutiny (e.g., being correct and useful for law or environmental protection) [13].

Empirical evidence suggests that perceived data quality often exceeds actual quality. This occurs partly because stakeholders underestimate the complexity of enforcing technical specifications, and partly due to cognitive biases such as overconfidence and the Dunning–Kruger effect. This effect occurs when individuals with limited technical expertise overestimate data reliability [14]. Together, these dynamics further widen the gap between technical compliance and justified trust, reinforcing the need to distinguish measured quality from perceived credibility within SDQ assessment.

2.2.1. Generic Example: Perceived Quality Exceeding Actual Quality

An officially endorsed elevation model is widely used in infrastructure planning due to its institutional provenance and visually smooth surface representation. Formal inspection later reveals undocumented resampling and interpolation steps that locally violate stated vertical accuracy thresholds. Despite measurable deviations from the declared specification, the dataset continues to be trusted and reused because validation processes are opaque, and institutional authority substitutes for verification. This example demonstrates how perceived quality can systematically exceed actual quality when trust signals override technical evidence.

2.3. Cognitive Bias and Human Error

Human involvement in spatial data production, validation, and use inevitably introduces cognitive biases that cannot be fully mitigated through technical standards alone. Well-documented effects such as confirmation bias, overconfidence, and adversarial allegiance influence how data quality is assessed, interpreted, and communicated [15–17]. These biases should not be understood as individual failings, but as predictable systemic risks emerging from the interaction of human cognition, institutional pressures, and technical system design. Cognitive limitations such as unreliable memory, spurious pattern recognition, and erroneous causal attribution are well-established sources of unintentional error, particularly in complex spatial interpretation and validation tasks [18,19].

The Swiss Cheese Model of system failure conceptualizes error as an emergent outcome of misaligned defensive layers rather than as isolated mistakes [20]. In spatial data infrastructures, this misalignment often manifests through complex validation workflows, ambiguous specifications, and organizational incentives that reward formal compliance over critical scrutiny [21]. Under such conditions, quality assurance risks becoming performative – focused on procedural completion, rather than epistemic, oriented toward understanding and mitigating uncertainty. Empirical studies further indicate that resistance to error reporting and corrective feedback is common in professional environments, reinforcing normalization of deviance instead of institutional learning [22].

doi:10.20944/preprints202603.0798.v1

Recognizing cognitive bias as a first-class quality risk factor reframes SDQ assurance from a purely technical task into a system design challenge [23]. Research on human–AI collaboration demonstrates that expert judgment is systematically shaped by automation bias, confirmation effects, and authority cues, even in high-stakes analytical contexts [24,25]. These findings confirm that bias is not incidental, but a structural property of complex decision environments involving both human and algorithmic actors.

In spatial data production and validation, such biases can normalize implausible patterns, discourage critical inspection of authoritative datasets, and reinforce premature trust in technically compliant outputs. Cognitive bias therefore explains how quality risks emerge at the operational level. Whether these risks are detected, corrected, or institutionalized, however, depends largely on governance structures, accountability mechanisms, and the transparency of provenance and metadata. Without independent verification pathways and explicit responsibility for challenge and review, biased interpretations may persist even within formally standardized quality systems.

2.3.1. Generic Example: Cognitive Bias and Performative Quality Assurance

A contractor delivers a spatial dataset that passes all required validation scripts and checklist-based quality assurance procedures under tight time constraints. Reviewers, familiar with the production workflow and operating under institutional pressure to certify delivery, do not challenge subtle anomalies in spatial distributions. Subsequent integration with independent datasets reveals systematic inconsistencies traceable to unexamined preprocessing assumptions. This example illustrates how cognitive bias, combined with organizational incentives, can transform quality assurance from an epistemic safeguard into a performative exercise.

2.4. Institutional Context, Authority and Legitimacy

SDQ is embedded within institutional contexts that shape how quality is defined, interpreted, and enforced. Standards are not neutral instruments, but products created, negotiated products of professional cultures that prioritize specific epistemic values and accountability regimes. In this sense, spatial data standards function as institutionalized social agreements that align expectations between producers, regulators, and users.

This perspective aligns with the concept of the co-production of science and social order, whereby technical knowledge and institutional authority mutually reinforce one another [26]. Legal-administrative contexts emphasize auditability and traceability, engineering cultures prioritize performance and efficiency, and scientific communities often tolerate probabilistic uncertainty. These institutional logics influence not only how quality metrics are defined, but also how uncertainty, deviation, and ambiguity are interpreted in practice.

Datasets produced under recognized governance frameworks, such as those aligned with the UN-GGIM Integrated Geospatial Information Framework [27] or the INSPIRE Directive [28] are often trusted by default and effectively black-boxed [29]. In such cases, trust derives less from auditable technical evidence than from perceived legitimacy, institutional continuity, and accountability. Conversely, technically rigorous datasets lacking institutional endorsement may face skepticism, particularly in regulatory or adversarial contexts.

These dynamics explain how trust in spatial data is initially established and stabilized through authority and legitimacy. However, they do not explain why trust may persist despite demonstrable technical shortcomings, nor why it can collapse abruptly even when formal quality compliance is maintained.

2.4.1. Generic Example: Institutional Authority Sustaining Trust Despite Technical Limitations

A spatial dataset produced within a national spatial data infrastructure and aligned with INSPIRE requirements is routinely accepted in regulatory processes. Independent analysis later identifies outdated source data and classification ambiguities that reduce analytical suitability for current applications. Nevertheless, the dataset remains authoritative in formal decision-making due to its institutional legitimacy and regulatory embedding. This example illustrates how governance structures can stabilize trust even when technical limitations are known.

2.5. Trust as an Emergent Socio-Technical Outcome

Trust in spatial data must be understood not as an institutional attribute or a direct consequence of technical certification. Rather, it is emergent socio-technical property arising from the interaction of technical admissibility, human cognition, interpretive plausibility, and institutional governance across the data lifecycle.

From a systems perspective, trust does not derive from individual quality indicators or formal compliance with standards [30]. Instead, it emerges from the alignment of multiple, interacting dimensions, including technical evidence, visual and semantic coherence, accountability mechanisms, and consistent performance over time. This perspective challenges linear quality models that implicitly treat trust as a cumulative result of improving technical metrics.

Consequently, trust cannot be reduced to a single score, label, or certification [31]. The contributing dimensions are non-compensatory: strong performance in one dimension (e.g., positional accuracy) cannot reliably offset weaknesses in another (e.g., undocumented lineage or opaque governance). This explains why datasets that pass all formal quality checks may still fail in legal, policy, or planning contexts, while others with known technical limitations remain widely accepted.

A key mechanism supporting the emergence and stabilization of trust is data lineage: the documented history of data origins, transformations, assumptions, and methodological choices. Lineage enables reconstruction of how spatial knowledge claims were produced and constrained, providing the evidentiary basis for responsible interpretation and reuse [32,33]. In its absence, trust is often sustained by institutional reputation rather than verifiable evidence, reinforcing gaps between perceived and actual quality.

Viewed in this way, trust becomes a dynamic socio-technical contract rather than a static property of data. It persists only as long as technical evidence, institutional practice, and human interpretation remain coherently aligned. This understanding provides the conceptual foundation for the lifecycle-oriented, forensic-aware SDQ framework developed in the subsequent sections.

2.5.1. Generic Example: Trust Collapse Despite Formal Quality Compliance

A regional land use change dataset is produced through a documented workflow and achieves high positional accuracy while formally completing with ISO 19157-1 quality elements. During environmental impact assessment, however, analysts identify spatial patterns that contradict known geomorphological constraints and recent field observations. Subsequent investigation reveals that key modeling assumptions used during preprocessing were not captured in the dataset’s lineage documentation. Although the dataset remains technically compliant with its declared specifications, the absence of transparent methodological provenance undermines interpretive plausibility and accountability. This example illustrates how trust can collapse when technical compliance, epistemic transparency, and institutional assurance fail to remain aligned.

2.6. Synthesis: What the Examples Reveal

Collectively, the preceding examples demonstrate that SDQ failures rarely originate from isolated technical deficiencies. Instead, they emerge from misalignments between technical compliance, human interpretation, and institutional governance (Table 1).

Table 1. Illustrative examples demonstrating why technical SDQ metrics alone are insufficient for comprehensive spatial data quality assessment.

Subsection	Quality Perspective	Dataset Context	Formal Technical Status	Observed Breakdown	Core Insight
2.1.1	Producer–user misalignment (fitness for purpose vs. fitness for use)	National land use/land cover dataset reused for flood-risk modeling	Fully compliant with ISO 19157-1 within the declared universe of discourse	Coarse mapping units and limited thematic detail prevent representation of hydrological features	Technical compliance does not guarantee suitability beyond the original design intent
2.2.1	Perceived vs. actual quality	Official elevation model used in infrastructure planning	Declared accuracy and visual coherence accepted without verification	Undocumented resampling violates local accuracy assumptions	Perceived reliability may exceed actual quality when institutional authority substitutes for verification
2.3.1	Cognitive Bias and human error	Contractor-produced dataset delivered under time and budget constraints	Passed all checklist-based QA/QC procedures	Systematic inconsistencies detected only after integration with independent data	Cognitive and organizational biases can transform quality assurance into a performative process
2.4.1	Institutional authority and legitimacy	INSPIRE-aligned national dataset used in regulatory processes	Formally standardized and institutionally endorsed	Known limitations remain unchallenged in formal decision-making contexts	Institutional legitimacy can stabilize trust despite known technical shortcoming
2.5.1	Emergent trust failure	High-quality land use change dataset used in environmental assessment	Fully compliant metrics and documented workflow	Implausible spatial patterns reveal missing modeling assumptions in lineage documentation	Trust emerges only when technical evidence, epistemic transparency, and governance remain aligned

3. A Holistic Framework for Spatial Data Quality

The analysis in previous section clearly demonstrates that spatial data quality cannot be adequately explained through technical compliance alone. Persistent divergence between actual and perceived quality shaped by cognitive bias, institutional authority, and governance practices, shows that SDQ functions as a socio-technical, lifecycle-dependent system property rather than a static dataset attribute. These findings motivate the need for an explicit structural model capable of explaining how quality is produced, interpreted, stabilized, or undermined across contexts of use.

Based on these insights, this section formalizes a holistic framework that conceptualizes SDQ as a multi-layered, non-linear, and emergent system (Figure 3). Rather than proposing new technical metrics or operational procedures, the framework. provides an analytical architecture that explains where quality originates, how it is maintained or degraded over time, and why technically correct data may still succeed or fail in practice.

The framework is organized into six analytically distinct layers, progressing from normative definition and technical admissibility through process and lifecycle stewardship, toward visual–interpretive assessment, governance, and emergent system-level quality states. Each layer addresses a specific class of mechanisms identified in previous section, while their interactions explain observed gaps between formal compliance, perceived credibility, and operational trust.

Operating at the level of conceptual integration, the framework organizes the diverse dimensions of SDQ identified in the previous section into a coherent structure applicable across production, reuse, integration, and governance contexts. Each subsection introduces one layer, clarifying its function, scope, and limitations, and explaining how it interacts with other layers throughout the data lifecycle.

The framework is intentionally layered and non-linear. Early layers establish necessary conditions for quality, such as purpose definition and technical admissibility, while later layers explain how quality evolves through reuse, interpretation, governance, and institutional practice. The final subsection does not add an additional layer, but it synthesizes these interactions by characterizing emergent quality states, which are configurations of risk, credibility, and trust that cannot be inferred from individual indicators alone.

Unlike the previous section, which examined conceptual tensions and explanatory limits in existing quality paradigms, this section proposes an explicit structural model that integrates those insights into a stable reference framework. The emphasis here is on structural completeness, internal consistency, and explanatory power.

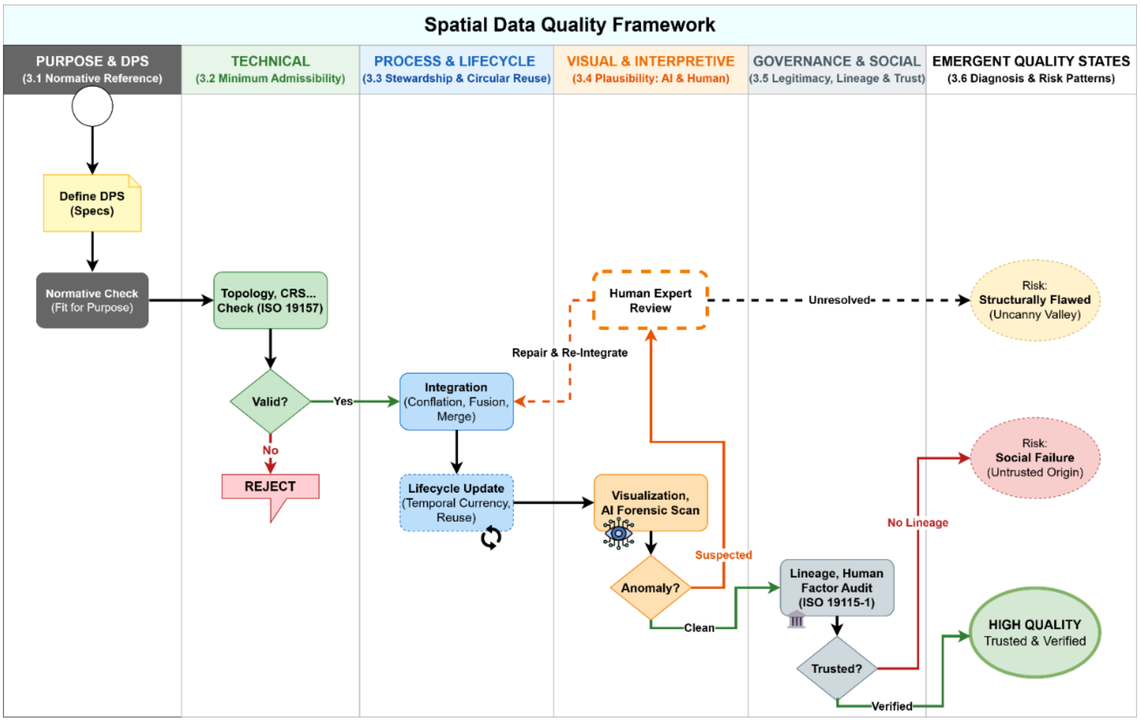


Figure 3. A simplified representation of the proposed multi-layered framework for SDQ as a socio-technical system.

Within the proposed framework, ISO standards and community principles can play complementary but distinct roles. ISO 19157-1 defines the technical admissibility layer by formalizing product-level quality elements and evaluation procedures, while ISO 19115-1 [34] operationalizes responsibility and transparency through structured metadata and lineage declarations. The FAIR principles extend these standards by emphasizing findability, interoperability, and reuse across systems and communities, whereas the TRUST principles articulate the institutional and organizational conditions – transparency, responsibility, user focus, sustainability, and technical robustness – under which quality claims remain credible over time [35,36]. In the conceptual diagram, ISO 19157-1 anchors technical conformance, ISO 19115-1 spans process and governance through metadata and lineage, FAIR supports cross-layer reuse and interoperability, and TRUST frames the governance layer that stabilizes trust and legitimacy. Together, these instruments reinforce the interpretation of SDQ as an emergent socio-technical property rather than a single measurable attribute.

3.1. Purpose and Data Product Specification: The Normative Quality Reference

Any meaningful assessment of SDQ requires a stable normative reference. Within the proposed framework, this reference is purpose, understood not as retrospective usefulness but as the explicit design intent declared at the time of data production. Purpose specifies which aspects of reality are represented, the level of abstraction applied, and the assumptions and constraints under which representation is considered valid. It therefore establishes the conditions under which all subsequent quality claims can be meaningfully interpreted.

Crucially, purpose is not inferred from downstream use or user expectations. It is treated as a producer-declared commitment that anchors quality assessment to a documented baseline rather than to shifting operational demands. This distinction is essential in contemporary data ecosystems, where spatial datasets are routinely reused, integrated, and repurposed far beyond their original context of production.

The DPS operationalizes purpose by translating design intent into a formal quality reference that defines the universe of discourse, spatial and temporal scope, content, representation model, quality requirements, and evaluation criteria. Within the framework, the DPS functions as the primary normative anchor: quality is first assessed as conformance to declared design intent, not as generalized adequacy for unspecified uses.

Beyond its technical role, the DPS functions as a boundary object across producers, validators, and users. It mediates expectations and responsibilities across institutional and disciplinary boundaries, remaining sufficiently formal to support standardization while interpretable enough to accommodate heterogeneous practices. Importantly, later quality dimensions do not replace the DPS. Rather, they interrogate assumptions, revealing where they remain valid, where they are stretched through reuse, and where they fail under new conditions.

Within the framework, the DPS establishes the reference point against which all subsequent quality dimensions are interpreted:

- Technical quality evaluates conformance to DPS-defined criteria;
- Process and lifecycle quality examine how DPS assumptions persist, evolve, or erode over time;
- Visual analytics and interpretive quality expose patterns and inconsistencies not captured by formal metrics;
- Governance and social indicators (trust) determine how DPS compliance is recognized, challenged, or legitimized in institutional contexts.

By fixing purpose explicitly at the entry point, the framework preserves analytical clarity while allowing reuse, reinterpretation, and institutional dynamics to be examined without destabilizing the original quality reference. Once purpose has been formally articulated through the DPS, quality assessment can proceed to the first evaluative layer: verification of technical conformance to that declared design intent.

3.2. Technical Quality: Minimum Admissibility and Specification Conformance

Technical quality constitutes the foundation layer of the framework, defining the minimum conditions under which spatial data are admissible for interpretation, integration, or reuse. Anchored in the DPS, this layer evaluates whether a dataset conforms to its declared design intent using standardized, reproducible, and largely deterministic criteria. Here, admissibility refers strictly to technical and structural coherence with the declared specification, not to legal validity, institutional endorsement, or suitability for a specific decision context.

This layer typically corresponds to the internationally established quality elements formalized in ISO 19157-1, including positional accuracy, thematic accuracy, completeness, logical consistency, temporal quality, and metadata quality. These indicators operationalize quality as conformance to specification, answering a narrow but essential question [37]: Does the dataset meet the conditions it claims to satisfy?

A defining feature of this foundation layer is the presence of non-negotiable constraints. Coordinate reference systems (CRSs) must be correctly defined and applied; schemas must conform to specification; topological rules must be satisfied; fundamental geometric and logical relationships must hold. These properties are binary in nature and cannot be meaningfully delegated to probabilistic inference or AI-based interpretation, as violations undermine the internal coherence of the data itself.

At the same time, the framework explicitly positions technical quality as necessary but not sufficient. Although ISO-based indicators are necessary for establishing baseline correctness, they are insensitive to risks arising from contextual reuse, semantic reinterpretation, cross-scale integration, or institutional dynamics. Technically compliant datasets may still mislead or fail operationally when applied beyond their declared scope.

Therefore, the framework treats ISO 19157-1 not as a comprehensive quality model, but as a baseline layer upon which additional evaluative dimensions can be coherently built. Technical quality assessment relies on quantitative, empirical, and statistical methods applicable to both vector and raster data. Within the framework, technical quality indicators are grouped by intrinsic properties, such as:

- Positional accuracy indicators assess geometric fidelity through measures such as RMSE, CE90, or LE90, consistency of CRSs, alignment with authoritative basemaps, topological correctness (absence of gaps, overlaps, and slivers), and distortion patterns in projected geometries;
- Attribute accuracy indicators evaluate the correctness and semantic coherence of non-spatial attributes, including error rates, controlled vocabularies, domain enforcement, and plausibility rules that reflect real-world logic (e.g., wetlands not occurring on steep slopes);
- Completeness indicators address feature and attribute presence, spatial coverage, expected density patterns, and temporal continuity, including the identification of missing tiles, absent attributes, or gaps in time series;
- Logical consistency indicators verify internal coherence through topology rules, schema validation, constraint satisfaction, and rule-based integrity checks that ensure geometric and semantic validity;
- Temporal quality indicators assess timeliness, temporal granularity, update regularity, and temporal drift, distinguishing between current, outdated, and misaligned observations;
- Metadata and lineage completeness indicators ensure ISO-compliant documentation, traceable provenance, and an unbroken lineage chain, enabling downstream interpretation and reuse;
- Resolution consistency indicators examine scale compatibility, spatial and temporal resolution alignment, and fitness of resolution relative to the declared purpose;

In this layer, metadata completeness is evaluated strictly as a technical prerequisite for interpretability and reuse, not as a guarantee of institutional transparency, trust, or governance maturity, which are addressed separately in Section 3.5.

Together, these indicators establish a minimum technical truth: whether the dataset is internally coherent, correctly structured, and faithful to its declared specification. Only once this foundation is secured does it become meaningful to assess process dynamics, interpretive quality, governance, trust lifecycle and sustainability, etc., which build upon, but do not replace this baseline layer.

3.3. Process Quality and Lifecycle: Data Stewardship and Circular Reuse

In this framework, process quality refers to how spatial data are produced, maintained, reused, and governed across their lifecycle, rather than to computational processing steps. Process and lifecycle therefore extend quality assessment beyond static product properties to the operational pathways that connect data acquisition, integration, updating, and long-term reuse. Within the proposed framework, this dimension captures how quality is preserved, degraded, or enhanced through data handling practices rather than through design intent or technical validation alone.

Conventional spatial data management has largely followed a linear “acquire–use–replace” model [38], in which new data are repeatedly collected to address perceived quality limitations in existing datasets [39]. Although this approach may appear to ensure technical currency, it often leads to fragmented data landscapes composed of multiple, partially overlapping versions with inconsistent assumptions, resolutions, and uncertainty characteristics. Repeated acquisition does not necessarily improve data quality. In many cases, it increases unpredictability at fine scales and undermines long-term comparability. This is particularly evident for stable or slowly changing spatial phenomena, such as digital elevation models (DEMs), cadastral boundaries, or reference frameworks, where wholesale dataset replacement discards accumulated institutional knowledge and obscures the evolution of data quality over time.

A circular approach to spatial data stewardship offers a more sustainable and operationally robust alternative. Rather than treating new data as replacements, circular strategies emphasize incremental improvement through integration, conflation, fusion, or merge and targeted correction

[39]. New data sources are used to diagnose limitations, refine specific components, and contextualize existing datasets, allowing quality improvements to be applied selectively where they are most needed. This stepwise enhancement approach preserves epistemic continuity, reduces unnecessary duplication of data acquisition efforts, and supports predictable quality evolution across successive reuse cycles. For many operational contexts, incremental correction has proven to be more sustainable and reliable than repeatedly regenerating entire data layers.

In many practices, therefore, spatial datasets seldom follow a linear lifecycle. Data produced for a specific purpose are routinely reused, integrated, and reinterpreted across new analytical, institutional, and technological contexts. This empirical reality challenges project-bound models of quality assurance and necessitates a shift toward lifecycle stewardship, in which data quality is continuously reassessed and refined as part of ongoing use. From this perspective, circularity is not merely an efficiency strategy, but a mechanism for maintaining trust and coherence as data move across contexts and over time.

Circular stewardship depends critically on metadata and lineage. Reliable documentation of data origins, transformations, and validation steps is a prerequisite for meaningful reuse and responsible integration. Lineage enables users to understand not only what the data represent, but how and why they take their current form. At the same time, integration, conflation, fusion, and merge can function as forensic processes: inconsistencies between datasets often expose undocumented assumptions, methodological changes, or integrity issues, thereby reconstructing missing lineage retrospectively. Process and lifecycle quality therefore require lineage-aware frameworks that acknowledge uncertainty and incompleteness rather than obscuring them behind formal compliance. This shift toward circular stewardship is further justified by the well-known 1:10:100 rule of data quality, which highlights how the cost and impact of unresolved data errors increase by orders of magnitude as they propagate across successive stages of the data lifecycle [40].

Beyond technical considerations, circular lifecycle stewardship also strengthens institutional integrity and accountability. When datasets are expected to persist across multiple reuse cycles, undocumented deviations, shortcuts, or selective interpretations become long-term liabilities rather than short-term conveniences. Iterative validation, cross-domain integration, and transparent documentation increase the likelihood that inconsistencies, whether accidental or intentional, are detected over time. In this way, process and lifecycle quality act as a self-correcting mechanism that reinforces trust across organizational and temporal boundaries.

From a sustainability perspective, circular spatial data management aligns directly with global and regional policy objectives by promoting resource efficiency, institutional learning, and continuity of evidence. It supports the principles underlying the Sustainable Development Goals (SDGs) [41] by reducing redundant data acquisition and preserving long-term comparability, reflects the INSPIRE Directive’s emphasis on reuse and shared responsibility, and is consistent with the UN-GGIM vision of authoritative, maintainable, and fitness for purpose geospatial information infrastructures. By shifting the focus from short-term compliance toward lifecycle stewardship, process and lifecycle quality transform SDQ from a reactive control mechanism into a proactive governance instrument, embedding trust, transparency, and sustainability within the data ecosystem itself.

Finally, while process and lifecycle quality explain how spatial data evolve through workflows and institutional practices, they do not address whether resulting spatial patterns remain plausible and coherent as representations of real-world phenomena. This limitation motivates the interpretive and visual quality assessment discussed in the following subsection.

3.4. Visual and Interpretive Quality: Spatial Visual Analytics and Human–AI Cognition

Visual and interpretive quality introduce a perceptual and diagnostic layer of SDQ assessment, focused on evaluating whether observed spatial patterns are plausible, coherent, and methodologically consistent with both real-world processes and declared production methods. Within the proposed framework, this layer operationalizes spatial visual analytics as a structured diagnostic process that combines human expertise with AI-supported pattern analysis to detect inconsistencies that may elude formal technical metrics.

In this context, spatial visual analytics is extended beyond exploratory visualization toward visual spatial forensics, where visual anomalies are treated as diagnostic signals of potential production shortcuts, lineage gaps, or governance failures rather than as purely visual artifacts. This reframing positions visual assessment as a form of quality intelligence that complements, rather than replaces, standards-based validation.

Crucially, visual quality assessment is neither an aesthetic judgment nor a substitute for statistical validation or ISO-based indicators. Instead, it examines how data behave when rendered, compared, and interrogated across scales, focusing on pattern coherence, environmental logic, semantic consistency, and methodological traceability, rather than cartographic style or design preference [42,43]. Its function is diagnostic and hypothesis-generating, identifying signals that warrant further technical, procedural, or institutional investigation.

From a conceptual standpoint, visual quality can be defined as the extent to which spatial patterns align with scientific expectations, physical constraints, and declared methodologies, as observable through visual inspection and visual analytics [44]. This includes realistic gradients and boundaries, semantic consistency between attributes and spatial form, topological coherence across scales, temporal plausibility of change patterns, and recognizable methodological "fingerprints" associated with specific processing techniques. When such expectations are violated, visual anomalies serve not as conclusions but as structured hypotheses directing attention to potential quality risks.

Visual assessment is particularly effective as an early-warning system. Common forensic indicators include unnaturally regular geometries, abrupt discontinuities at administrative boundaries without environmental justification, repetitive pixel or tile artifacts, excessive smoothing, artificial symmetry, or visual signatures inconsistent with the stated production method. These patterns more often signal undocumented transformations, inappropriate automation, or misaligned assumptions embedded in production workflows than isolated technical errors.

The diagnostic strength of this layer arises from the complementary roles of human and AI cognition. Human experts contribute contextual reasoning, domain knowledge, and the recognition of environmental implausibility, while AI-based methods provide systematic, repeatable detection of non-natural regularities, repetitive structures, and statistical deviations in spatial textures and patterns. Empirical research on human–AI collaboration shows that effective performance emerges when cognitive roles are clearly differentiated and algorithmic outputs are treated as diagnostic inputs rather than authoritative judgments [24,25]. In this role, AI functions as an independent counter-check to human interpretation, not as a replacement for expert judgement.

Many forms of low-quality automation or inappropriate data manipulation leave distinctive visual traces including grid artifacts, halo effects, artificial textures or false detail at inappropriate resolutions. Such signatures have been widely documented in the remote sensing quality literature and can be robustly detected through a combination of expert visual inspection and algorithmic analysis [45,46].

To improve reproducibility and reduce subjectivity, visual and interpretive quality can be supported by quantifiable indicators derived from image and spatial analysis, such as texture metrics, edge density, spatial entropy, fractal measures, and boundary contrast indices. These measures do not replace expert inspection, but formalize and stabilize it, enabling systematic comparison across datasets and supporting AI-assisted anomaly detection while preserving interpretability.

Visual and interpretive quality interacts closely with data integration and lineage. During integration, inconsistencies between datasets frequently expose divergences between declared specifications and actual production methods. Lineage documentation allows such discrepancies to be traced back to specific processing decisions, assumptions, or undocumented shortcuts, turning integration into a forensic trigger that reveals quality issues invisible when datasets are analyzed in isolation.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 11 March 2026 doi:10.20944/preprints202603.0798.v1

Beyond technical diagnostics, visual quality also has important social and institutional implications. Visually implausible or overly "perfect" datasets may indicate deeper organizational pathologies, such as excessive automation, insufficient field validation, or the masking of uncertainty through presentation choices. Because visual evidence carries disproportionate persuasive power in decision-making contexts, visual credibility risks must be treated as quality risks in their own right.

By introducing visual spatial forensics as an explicit quality dimension, the framework moves beyond an exclusive reliance on formal technical testing, which can unintentionally obscure real-world inconsistencies [47]. Visual plausibility provides an independent line of evidence, consistent with forensic principles emphasizing corroboration across multiple sources [48]. When a dataset is technically compliant yet visually inconsistent with environmental or methodological logic, the limitation lies not in visual assessment, but in the contextual sensitivity of the technical metrics themselves.

While visual and interpretive quality assess plausibility and methodologically credibility, they do not by themselves explain why datasets are ultimately trusted, accepted, or rejected in institutional and policy contexts. This transition from plausibility to legitimacy is addressed in the governance and social dimensions of SDQ discussed in the following subsection.

3.5. Governance and Social Indicators: Legitimacy, Lineage, and Trust as Epistemic Memory

Governance and social indicators constitute the institutional and socio-technical layer of the proposed framework, addressing aspects of SDQ that cannot be inferred from technical correctness or visual plausibility alone. This layer illustrates how trust in spatial data is produced, maintained, and disputed through organizational structures, social practices, and accountability mechanisms [49]. By doing so, it links technical validation to real-world use in policy, spatial planning, and legal decision-making.

These indicators capture the conditions under which spatial data are produced, maintained and reused, shaping not only internal quality management but also external perceptions of credibility and legitimacy. Datasets associated with stable institutions, explicit mandates, and sustained funding are often trusted by default, whereas technicaly sound datasets lacking institutional anchoring may face skepticism. Trust, therefore, is not a direct function of accuracy but an emergent property of governance maturity and institutional transparency.

Human factors play a decisive role at this level, not by introducing errors per se, but by shaping how errors are interpreted, communicated, and institutionally resolved. Expert judgment, organizational norms, and shared cognitive frames influence whether quality concerns are escalated, normalized, or dismissed. While expert knowledge enables contextual reasoning and innovation, it also introduces risks of interpretive drift and collective blind spots. Concepts such as naïve geography illustrate how everyday spatial assumptions can shape expert interpretation in GIS environments [50]. Governance quality is thus reflected not in the absence of error, but in the presence of mechanisms that reliably detect, document, and correct it through institutional learning. Socio-technical research emphasizes that recurring human errors signal deficiencies in system design, organizational culture, or incentive structures and should be addressed through preventive governance mechanisms rather than individual blame [22].

Errors in spatial data production are frequently unintentional, arising from limitations in perception, memory or reasoning [51]. Governance structures determine whether such errors are corrected or instead normalized and propagated. From a socio-technical perspective, integrity failures are best addressed through design rather than attribution of fault. Robust governance creates conditions in which undocumented shortcuts or deviations from specifications become detectable, regardless of intent. Beyond unintentional error, spatial data are also vulnerable to negligent or deliberate distortions driven by political, economic, or reputational incentives [52]. Effective governance therefore requires structural safeguards, including independent validation and institutional cultures that reward critical review over superficial compliance [27].

A central feature of this layer is provenance and lineage transparency, which functions as the forensic trail of institutional responsibility [53]. Although lineage is traditionally understood as documentation of data sources, processing steps, and transformations, it is frequently incomplete, simplified, or fragmented during reuse. As a result, lineage operates less as a precise historical record and more as a form of epistemic memory [54], which is a socially mediated account how spatial knowledge has been constructed, and modified over time.

Within this governance context, metadata standards, particularly ISO 19115-1, play a role that extends beyond technical documentation. Rather than merely describing datasets, ISO 19115-1 metadata institutionalize responsibility by formally declaring purpose, assumptions, constraints, lineage, usage limitations, and points of contact. In doing so, metadata function as governance instruments that shape how datasets can be trusted, challenged, or defended across organizational, legal, and policy contexts.

Importantly, metadata completeness under ISO 19115-1 does not guarantee technical correctness, but it strongly influences institutional credibility. Explicit documentation of uncertainty, methodological choices, and known limitations enables critical scrutiny and responsible reuse. Conversely, missing, outdated, or generic metadata obscure responsibility, encourage overconfidence, and amplify interpretive risk, particularly in downstream integration and AI-assisted analysis. In this framework, ISO 19115-1 metadata are therefore treated not as a quality dimension in themselves, but as a governance mechanism that stabilizes trust by preserving epistemic memory across time, institutions, and reuse cycles.

This epistemic memory is essential for sustaining trust. Open and inspectable lineage signals a willingness to acknowledge uncertainty. Conversely, opaque workflows (including undocumented revisions, selective disclosure, or black-boxed processing) erode trust even when formal specifications are met. Lineage thus acts as connective tissue linking the technical, process, interpretive, and governance layers across the data lifecycle [55]. Its importance becomes particularly evident during data integration, where inconsistencies often reveal divergences between declared specifications and actual production practices. In this sense, integration functions as a forensic trigger, exposing undocumented methodological choices or institutional shortcuts that remain invisible when datasets are used in isolation [56].

When such lineage gaps are discovered, the framework favors circular stewardship over disposal. Rather than discarding problematic datasets, they are recycled through integration and incremental improvement, treating lineage preservation as a prerequisite for sustainable reuse. This approach also encompasses ethical considerations, including spatial biases such as the systematic under-representation of marginalized areas, which remain invisible to conventional technical metrics [57]. Addressing these biases requires institutional mechanisms that promote inclusive data collection, participatory validation, and respect for data sovereignty, consistent with international frameworks such as SDGs monitoring and UN-GGIM guidance.

The growing use of AI amplifies governance challenges. Just as generative models may produce plausible but ungrounded outputs, spatial data systems risk reinforcing inherited assumptions when lineage and metadata are incomplete. In such contexts, both human analysts and automated systems may infer coherence where documentation is lacking. Governance therefore determines whether AI-assisted workflows enhance transparency and auditability, or instead amplify black-box opacity and misplaced confidence [58,59].

Ultimately, governance and social indicators transform quality assessment from a technical exercise into a question of credibility, legitimacy, and accountability. At this stage, the quality question shifts from “Is the data correct?” to “Is the data trustworthy and appropriate for this context?” [49]. By explicitly integrating human factors, lineage, and ISO 19115-1 metadata as instruments of epistemic memory, this layer ensures that quality assessment reflects how spatial data actually function within complex institutional and decision-making environments.

doi:10.20944/preprints202603.0798.v1

3.6. Emergent Quality States: Integrated Diagnosis and Risk Patterns

This subsection does not introduce an additional quality layer but synthesizes the preceding ones by describing emergent system-level quality states. SDQ does not reside in isolated indicators but emerges from interactions among purpose definition, technical admissibility, visual–interpretive plausibility, lifecycle stewardship, and governance maturity across the data lifecycle. These interactions give rise to recognizable quality states, such as systemic patterns of reliability, risk, and credibility, that cannot be inferred from individual metrics or checklist-based evaluations alone. Within the proposed framework, quality is therefore conceptualized as an emergent property of a socio-technical system rather than as an additive or static score [51].

This perspective is particularly relevant in contemporary spatial data infrastructures, open data ecosystems, and AI-assisted geospatial workflows, where datasets are routinely reused, integrated, and repurposed far beyond their original design intent. In such environments, technically correct data may fail operationally, legally, or politically, while socially trusted datasets may remain analytically constrained. Quality assessment must therefore shift from isolated attributes to interaction patterns, focusing on how different quality dimensions reinforce, compensate for, or contradict one another.

When technical, visual, and governance layers interact, they frequently generate conflicting signals that reveal latent integrity risks. These configurations demonstrate that no single quality dimension is dominant. Instead, distinct quality states emerge from their alignment or tension. Common patterns include:

- Technically sound but socially weak: High positional or thematic accuracy combined with low institutional accountability, unclear lineage, or limited community trust results in datasets that are formally compliant yet contested or underused. This pattern is often seen in externally commissioned or black-boxed datasets that lack transparent data product specifications.
- Socially legitimate but technically incomplete: Strong institutional endorsement or community validation paired with outdated resolution, limited attribute depth, or known analytical constraints can lead to widespread adoption despite restricted analytical reliability.
- Complete but epistemically fragile: High completeness without adequate metadata or lineage documentation reduces interpretability and reuse potential. In this case, uncertainty is not absent but undocumented, increasing the risk of downstream misinterpretation and integration failure.
- Visually convincing but structurally flawed: High visual coherence may mask hidden topological, semantic, or methodological inconsistencies, creating false confidence and delaying error detection. This pattern underscores the necessity of combining visual forensics with formal validation [60].

From a system's perspective, these emergent quality states act as risk signatures. They indicate whether a dataset is likely to support robust decision-making, invite contestation, or conceal latent failures that may only surface during integration or high-stakes use. Importantly, most quality failures do not result from overt fabrication but from methodological shortcuts, undocumented assumptions, or incentive-driven distortions embedded in production and validation workflows [52].

Detecting such risks requires forensic interpretation, understood as the systematic combination of independent evidence streams rather than reliance on any single authoritative indicator [61]. Within the proposed framework, this includes:

- Cross-dataset pattern matching, comparing spatial, temporal, and semantic structures across independent datasets representing the same phenomena;
- Lineage reconstruction through integration, where partially overlapping datasets expose contradictions that allow elements of the production workflows to be inferred retrospectively;
- Statistical indicators revealing hidden processing practices,, such as distributional anomalies, unexpectedly low variance, repeated spatial motifs, or autocorrelation patterns indicative of resampling, smoothing, tiling, or copy–paste operations;

- Socio-institutional analysis, examining incentives, accountability mechanisms, organizational pressures, and governance contexts surrounding data production and validation.

The central insight of this section is that SDQ cannot be meaningfully reduced to a composite score without interpretive context. A dataset may satisfy all technical criteria and still fail socially. It may be also visually persuasive yet epistemically fragile, or it may be institutionally trusted despite known analytical limitations. Emergent quality states make these configurations explicit and render their associated risks visible.

Recognizing such states transforms quality assessment from a static control exercise into a diagnostic and governance-oriented practice. Rather than asking only whether data meet predefined thresholds, the framework examines how quality is produced, sustained, and challenged over time. This shift supports earlier detection of integrity risks, more transparent communication of limitations, and more responsible reuse of spatial data in high-stakes decision-making contexts, thereby directly reinforcing sustainability, accountability, and trust.

3.6.1. Composite Quality Index

To operationalize quality states, the framework supports the introduction of a Composite Quality Index (*CQI*) based on the core quality dimensions defined in Sections 3.2–3.5. To ensure that technical soundness acts as a foundation for all subsequent assessments, the *CQI* is calculated using a multiplicative gate for technical quality combined with a weighted average of the remaining dimensions:

$$CQI = TQ/100 \cdot (W_L \cdot LQ + W_V \cdot VQ + W_G \cdot GQ) \quad (1)$$

where technical quality (*TQ*) can represent ISO 19157-based admissibility criteria (Section 3.2). This component acts as a prerequisite: if technical admissibility is zero, the entire index is zero, regardless of other scores. Process and lifecycle quality (*LQ*) captures stewardship, update practices, and lineage continuity (Section 3.3), visual and interpretive quality (*VQ*) reflects forensic plausibility signals and pattern coherence (Section 3.4), and governance and trust quality (*GQ*) represents institutional reliability, provenance completeness, and transparency (Section 3.5). All components are independently normalized to a 0–100 scale, where:

- 0 indicates unacceptable quality;
- 50 indicates risk or partial compliance;
- 100 indicates full adequacy for the declared purpose.

The weights must satisfy $W_L + W_V + W_G = 1$, ensuring the secondary dimensions form a complete weighted average. Technical quality is thus treated as a strict prerequisite for meaningful higher-level assessment, mathematically encoded as a multiplier to prevent compensation by other factors when fundamental admissibility criteria are not met.

Efficient computation of the individual components can be achieved through tiered and reusable indicators instead of extensive re-evaluation. *TQ* can be derived directly from standardized validation outputs (e.g., pass/fail ratios, error densities, threshold exceedance rates). *LQ* can be estimated from metadata completeness scores, update latency, version continuity, and the presence of verifiable lineage elements. *VQ* can be operationalized through lightweight visual diagnostics, such as detection of systematic artifacts, spatial autocorrelation anomalies, or inconsistencies between declared uncertainty and observed spatial patterns. *GQ* can be approximated using structured provenance indicators, including documentation depth, institutional accountability, auditability, and reuse history.

These indicators are intentionally designed to be computationally inexpensive, partially automatable, and incrementally refinable, allowing early risk signaling without requiring full reprocessing of datasets. The current implementation of the OPIAvalid toolkit (described in the Discussion section) primarily addresses *TQ* and partially *VQ*. However, the proposed framework enables future expansion toward integrated, risk-aware composite scoring. Thus, quality assessment evolves from a static control mechanism into an instrument supporting decision-making, prioritization, transparency, and responsible reuse throughout the spatial data lifecycle.

4. Discussion

The proposed framework demonstrates that SDQ cannot be reduced to isolated technical metrics, but instead emerges from the interaction among technical validation, human interpretation, institutional context, and lifecycle governance. This perspective becomes particularly salient in light of the growing use of artificial intelligence (AI), including generative AI, in geospatial workflows [62]. While these technologies promise increased efficiency and analytical reach, they also intensify long-standing tensions between automation and explainability, probabilistic inference and deterministic spatial constraints, and computational scalability and institutional accountability.

Importantly, AI-based methods do not resolve classical quality problems. Rather they tend to amplify both strengths and weaknesses of existing data practices. Spatial data remain governed by non-negotiable constraints, such as CRSs, topology, and logical consistency that require explicit rule-based validation and cannot be inferred statistically. Consequently, AI cannot function as an autonomous arbiter of SDQ. Its contribution is most effective when positioned as a supportive analytical layer that assists in anomaly detection, prioritization of quality risks, and communication of uncertainty, while remaining grounded in transparent standards, declared purpose, and inspectable lineage.

Accordingly, AI-related considerations appear throughout this discussion implicitly rather than as a separate analytical dimension. They surface in issues of specification interpretation, lifecycle stewardship, visual plausibility, governance, and trust, without displacing the central role of standards, human judgment, and institutional responsibility. This framing shifts the focus from whether AI should be used, but on how socio-technical quality frameworks can constrain, contextualize, and govern its use in ways that preserve credibility, accountability, and long-term sustainability of spatial data infrastructures. These considerations are particularly relevant for emerging applications such as Digital Twins, AI-assisted spatial planning, and automated decision-support systems.

Table 2 summarizes how deterministic validation and AI-based methods (including generative AI) can support different SDQ dimensions within the proposed framework. The table reflects the layered structure of SDQ introduced in Sections 3.1–3.5. Deterministic, standards-based validation remains essential for evaluating core technical properties of spatial datasets, such as completeness, logical consistency, positional accuracy, and metadata compliance, where reproducible rule-based checks can be formally defined. AI and machine-learning approaches play a complementary role by assisting in the detection of complex patterns, prioritization of verification efforts, and interpretation of quality risks that cannot be captured through deterministic rules alone. In this framework, AI therefore functions primarily as a diagnostic and explanatory instrument that supports expert assessment and governance processes, rather than replacing institutional quality control, formal validation procedures, or domain expertise.

Table 2. Role of deterministic validation and AI methods (including generative) across SDQ dimensions. Quality dimensions shown in bold correspond to the extended SDQ components emphasized in the proposed framework.

Quality dimension (ISO/framework)	**Deterministic validation (standards/rules)**	**AI/ML support**	**Primary role of AI**
Completeness (ISO 19157)	☑	⚠	Detection of spatial patterns in missing features, clustering of omission areas, prioritization for review
Logical consistency (ISO 19157)	☑	✕	None – strictly rule-based (topology, CRS, schema), mostly binary results

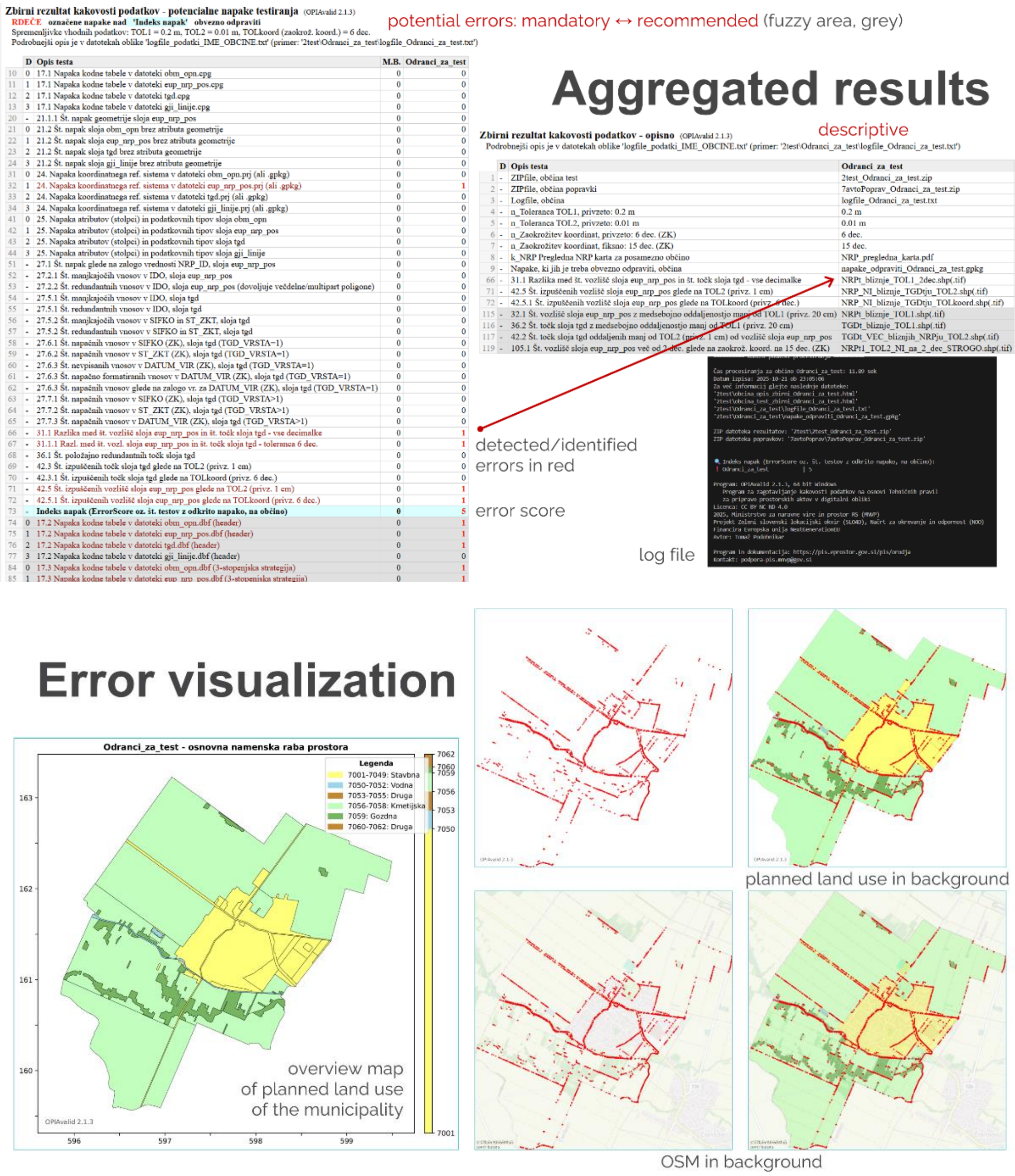
Positional accuracy (ISO 19157)	☑	⚠	Detection of spatial outlier, stratification of positional errors, prioritization of verification efforts
Thematic accuracy (ISO 19157)	☑	☑	Classification support, anomaly detection, identification of semantic inconsistencies
Temporal quality (ISO 19157)	☑	☑	Change detection, temporal trend analysis, identification of temporal drift
Process compliance – specification conformance (ISO 19157)	☑	⚠	Detection of recurring workflow deviations or systematic misinterpretation of specifications
Metadata completeness (ISO 19115-1)	☑	⚠	Detection of inconsistencies between metadata declarations and dataset characteristics, automated text analysis
Lineage & provenance (ISO 19115-1)	⚠	⚠	Reconstruction hints, workflow explanation, identification of undocumented processing assumptions
Conceptual & purpose-dependent quality (3.1)	✕	⚠	Identification of potential purpose mismatches and implicit design assumptions
Process & lifecycle quality (3.3)	⚠	☑	Anticipation of lifecycle risks, identification of weak points in complex production or integration workflows
Visual & interpretive plausibility (3.4)	✕	☑	Detection of implausible spatial patterns, scale inconsistencies, and structural anomalies
Governance & trust indicators (3.5)	✕	⚠	Explainability support, synthesis of governance context, identification of transparency gaps
Usability – fitness for purpose (3.2)	✕	☑	Context-aware interpretation and decision-oriented summaries of dataset suitability

Legend: ☑ appropriate or required; ⚠ supportive or explorative; ✕ not applicable.

The discussion that follows draws on practical experience gained through OPIAvalid, an automated, standards-aware toolkit that operationalizes the multi-layered SDQ framework proposed in this paper. The OPIAvalid is a freely available, standalone desktop application that currently integrates over 100 deterministic and diagnostic tools for evaluating SDQ across technical, process, and governance dimensions [63]. It is embedded within the national Spatial Information System of the Ministry of Natural Resources and Spatial Planning (Slovenia), where it supports validation of authoritative datasets, particularly planned land use (PLU) data.

Input data consist primarily of vector layers and their associated specifications, while outputs include structured quality reports distinguishing mandatory and recommended corrections, supported by summary statistics, logs, and dedicated spatial layers for each detected error type. The toolkit combines deterministic rule-based validation (e.g., topology, schema conformance, completeness) with diagnostic and visual outputs that support interpretation, prioritization, and correction of quality risks. This hybrid approach enables systematic detection of both formal specification misuse and context-dependent inconsistencies, thereby bridging measured and perceived data quality (Figure 4) [10]. Repeatable and inspectable workflows further support transparent governance and institutional learning across the data lifecycle.

To complement the vector-based, specification-driven example, the discussion also draws on a raster-based implementation focused on DEM production through multi-source data conflation [39]. Understanding data conflation is also referred to as understanding integration or fusion. The resulting authoritative DEM datasets, obtainable from the Surveying and Mapping Authority (Slovenia), have served as a national standard at multiple resolutions since 2005 and have been integrated into the European-scale model and the improved Google Earth terrain representation. This demonstrates how lifecycle-aware integration and incremental improvement can enhance data quality across heterogeneous data sources [64]. Together, these cases demonstrate that SDQ emerges through sustained stewardship, transparency and governance, rather than static compliance alone.



Zbirni rezultat kakovosti podatkov - potencialne napake testiranja (OPIAvalid 2.1.3)
RDEČE označene napake nad 'Indeks napak' obvezno odpraviti
Spremenljivke vhodnih podatkov: TOL1 = 0.2 m, TOL2 = 0.01 m, TOLkoord (zaokrož. koord.) = 6 dec.
Podrobnejši opis je v datotekah oblike 'logfile_podatki_IME_OBCINE.txt' (primer: '2test\Odranci_za_test\logfile_Odranci_za_test.txt')

	D	Opis testa	M.B.	Odranci_za_test
10	0	17.1 Napaka kodne tabele v datoteki obm_opn.cpg	0	0
11	1	17.1 Napaka kodne tabele v datoteki eup_nrp_pos.cpg	0	0
12	2	17.1 Napaka kodne tabele v datoteki tgd.cpg	0	0
13	3	17.1 Napaka kodne tabele v datoteki gji_linije.cpg	0	0
20	-	21.1.1 Št. napak geometrije sloja eup_nrp_pos	0	0
21	0	21.2 Št. napak sloja obm_opn brez atributa geometrije	0	0
22	1	21.2 Št. napak sloja eup_nrp_pos brez atributa geometrije	0	0
23	2	21.2 Št. napak sloja tgd brez atributa geometrije	0	0
24	3	21.2 Št. napak sloja gji_linije brez atributa geometrije	0	0
31	0	24. Napaka koordinatnega ref. sistema v datoteki obm_opn.prj (ali .gpkg)	0	0
32	1	24. Napaka koordinatnega ref. sistema v datoteki eup_nrp_pos.prj (ali .gpkg)	0	1
33	2	24. Napaka koordinatnega ref. sistema v datoteki tgd.prj (ali .gpkg)	0	0
34	3	24. Napaka koordinatnega ref. sistema v datoteki gji_linije.prj (ali .gpkg)	0	0
41	0	25. Napaka atributov (stolpci) in podatkovnih tipov sloja obm_opn	0	0
42	1	25. Napaka atributov (stolpci) in podatkovnih tipov sloja eup_nrp_pos	0	0
43	2	25. Napaka atributov (stolpci) in podatkovnih tipov sloja tgd	0	0
44	3	25. Napaka atributov (stolpci) in podatkovnih tipov sloja gji_linije	0	0
51	-	27.1 Št. napak glede na zalogo vrednosti NRP_ID, sloja eup_nrp_pos	0	0
52	-	27.2.1 Št. manjkajočih vnosov v IDO, sloja eup_nrp_pos	0	0
53	-	27.2.2 Št. redundantnih vnosov v IDO, sloja eup_nrp_pos (dovoljuje večdelne/multipart poligone)	0	0
54	-	27.5.1 Št. manjkajočih vnosov v IDO, sloja tgd	0	0
55	-	27.5.1 Št. redundantnih vnosov v IDO, sloja tgd	0	0
56	-	27.5.2 Št. manjkajočih vnosov v SIFKO in ST_ZKT, sloja tgd	0	0
57	-	27.5.2 Št. redundantnih vnosov v SIFKO in ST_ZKT, sloja tgd	0	0
58	-	27.6.1 Št. napačnih vnosov v SIFKO (ZK), sloja tgd (TGD_VRSTA=1)	0	0
59	-	27.6.2 Št. napačnih vnosov v ST_ZKT (ZK), sloja tgd (TGD_VRSTA=1)	0	0
60	-	27.6.3 Št. nevpisanih vnosov v DATUM_VIR (ZK), sloja tgd (TGD_VRSTA=1)	0	0
61	-	27.6.3 Št. napačno formatiranih vnosov v DATUM_VIR (ZK), sloja tgd (TGD_VRSTA=1)	0	0
62	-	27.6.3 Št. napačnih vnosov glede na zalogo vr. za DATUM_VIR (ZK), sloja tgd (TGD_VRSTA=1)	0	0
63	-	27.7.1 Št. napačnih vnosov v SIFKO (ZK), sloja tgd (TGD_VRSTA>1)	0	0
64	-	27.7.2 Št. napačnih vnosov v ST_ZKT (ZK), sloja tgd (TGD_VRSTA>1)	0	0
65	-	27.7.3 Št. napačnih vnosov v DATUM_VIR (ZK), sloja tgd (TGD_VRSTA>1)	0	0
66	-	31.1 Razlika med št. vozlišč sloja eup_nrp_pos in št. točk sloja tgd - vse decimalke	0	1
67	-	31.1.1 Razl. med št. vozl. sloja eup_nrp_pos in št. točk sloja tgd - toleranca 6 dec.	0	1
68	-	36.1 Št. položajno redundantnih točk sloja tgd	0	0
69	-	42.3 Št. izpuščenih točk sloja tgd glede na TOL2 (privz. 1 cm)	0	0
70	-	42.3.1 Št. izpuščenih točk sloja tgd glede na TOLkoord (privz. 6 dec.)	0	0
71	-	42.5 Št. izpuščenih vozlišč sloja eup_nrp_pos glede na TOL2 (privz. 1 cm)	0	1
72	-	42.5.1 Št. izpuščenih vozlišč sloja eup_nrp_pos glede na TOLkoord (privz. 6 dec.)	0	1
73	-	**Indeks napak (ErrorScore oz. št. testov z odkrito napako, na občino)**	**0**	**5**
74	0	17.2 Napaka kodne tabele v datoteki obm_opn.dbf (header)	0	1
75	1	17.2 Napaka kodne tabele v datoteki eup_nrp_pos.dbf (header)	0	1
76	2	17.2 Napaka kodne tabele v datoteki tgd.dbf (header)	0	1
77	3	17.2 Napaka kodne tabele v datoteki gji_linije.dbf (header)	0	0
84	0	17.3 Napaka kodne tabele v datoteki obm_opn.dbf (3-stopenjska strategija)	0	1
85	1	17.3 Napaka kodne tabele v datoteki eup_nrp_pos.dbf (3-stopenjska strategija)	0	1

Zbirni rezultat kakovosti podatkov - opisno (OPIAvalid 2.1.3)
Podrobnejši opis je v datotekah oblike 'logfile_podatki_IME_OBCINE.txt' (primer: '2test\Odranci_za_test\logfile_Odranci_za_test.txt')

	D	Opis testa	Odranci_za_test
1	-	ZIPfile, občina test	2test_Odranci_za_test.zip
2	-	ZIPfile, občina popravki	7avtoPoprav_Odranci_za_test.zip
3	-	Logfile, občina	logfile_Odranci_za_test.txt
4	-	n_Toleranca TOL1, privzeto: 0.2 m	0.2 m
5	-	n_Toleranca TOL2, privzeto: 0.01 m	0.01 m
6	-	n_Zaokrožitev koordinat, privzeto: 6 dec. (ZK)	6 dec.
7	-	n_Zaokrožitev koordinat, fiksno: 15 dec. (ZK)	15 dec.
8	-	k_NRP Pregledna NRP karta za posamezno občino	NRP_pregledna_karta.pdf
9	-	Napake, ki jih je treba obvezno odpraviti, občina	napake_odpraviti_Odranci_za_test.gpkg
66	-	31.1 Razlika med št. vozlišč sloja eup_nrp_pos in št. točk sloja tgd - vse decimalke	NRPt_bliznje_TOL1_2dec.shp(.tif)
71	-	42.5 Št. izpuščenih vozlišč sloja eup_nrp_pos glede na TOL2 (privz. 1 cm)	NRP_NI_bliznje_TGDtju_TOL2.shp(.tif)
72	-	42.5.1 Št. izpuščenih vozlišč sloja eup_nrp_pos glede na TOLkoord (privz. 6 dec.)	NRP_NI_bliznje_TGDtju_TOLkoord.shp(.tif)
115	-	32.1 Št. vozlišč sloja eup_nrp_pos z medsebojno oddaljenostjo manj od TOL1 (privz. 20 cm)	NRPt_bliznje_TOL1.shp(.tif)
116	-	36.2 Št. točk sloja tgd z medsebojno oddaljenostjo manj od TOL1 (privz. 20 cm)	TGDt_bliznje_TOL1.shp(.tif)
117	-	42.2 Št. točk sloja tgd oddaljenih manj od TOL2 (privz. 1 cm) od vozlišč sloja eup_nrp_pos	TGDt_VEC_bliznjih_NRPju_TOL2.shp(.tif)
119	-	105.1 Št. vozlišč sloja eup_nrp_pos več od 2 dec. glede na zaokrož. koord. na 15 dec. (ZK)	NRPt1_TOL2_NI_na_2_dec_STROGO.shp(.tif)

```
Čas procesiranja za občino Odranci_za_test: 11.89 sek
Datum izpisa: 2025-10-21 ob 23:05:06
Za več informacij glejte naslednje datoteke:
'2test\obcina_opis_zbirni_Odranci_za_test.html'
'2test\obcina_test_zbirni_Odranci_za_test.html'
'2test\Odranci_za_test\logfile_Odranci_za_test.txt'
'2test\Odranci_za_test\napake_odpraviti_Odranci_za_test.gpkg'

ZIP datoteka rezultatov: '2test\2test_Odranci_za_test.zip'
ZIP datoteka popravkov: '7avtoPoprav\7avtoPoprav_Odranci_za_test.zip'


🔍 Indeks napak (ErrorScore oz. št. testov z odkrito napako, na občino):
❗ Odranci_za_test          | 5

Program: OPIAvalid 2.1.3, 64 bit Windows
   Program za zagotavljanje kakovosti podatkov na osnovi Tehničnih pravil
   za pripravo prostorskih aktov v digitalni obliki
Licenca: CC BY NC ND 4.0
2025, Ministrstvo za naravne vire in prostor RS (MNVP)
Projekt Zeleni slovenski lokacijski okvir (SLO4D), Načrt za okrevanje in odpornost (NOO)
Financira Evropska unija NextGenerationEU
Avtor: Tomaž Podobnikar

Program in dokumentacija: https://pis.eprostor.gov.si/pis/orodja
Kontakt: podpora-pis.mnvp@gov.si
```

Figure 4. Selected outputs of the OPIAvalid toolkit. Top: Aggregated results in various forms with potential errors identified in red. Below: Overview map of the PLU data and potential error visualizations in georeferenced vector and raster forms.

4.1. Data Product Specification

A representative example of purpose misalignment arises from the interpretation of DPSs in national PLU datasets [63]. In this case, the DPS defined the thematic content and technical structure of multiple datasets, including a prescribed set of attribute fields. The specification listed these fields in a fixed sequence, beginning with the feature identifier, reflecting the producer's design intent that the attribute order be consistent across all delivered datasets.

However, this ordering requirement was not explicitly stated in normative language, as it was assumed to be self-evident. As a result, several data producers delivered datasets in which the attribute fields were present but arranged in different orders. From a purely technical perspective, the datasets were complete and syntactically valid. However, downstream processes relying on fixed field positions (e.g., automated validation, integration, or legacy systems) failed or produced inconsistent results.

Within the proposed framework, this situation is not classified as a technical data defect, nor as a legitimate variation in implementation. Instead, it is identified as a misuse of the specification, caused by an implicit assumption embedded in the DPS rather than by a failure of measurement or data capture. The apparent quality problem thus originates from an unarticulated design constraint and its inconsistent interpretation, rather than from non-compliance with declared quality elements.

This example illustrates how quality failures may emerge even when data satisfy declared technical criteria and underscores the importance of explicitly documenting design assumptions in the DPS. It also demonstrates why quality assessment must distinguish between production errors and specification ambiguities, reinforcing the role of purpose and design intent as the primary reference point for interpreting quality outcomes.

4.2. Technical Quality

Technical quality validation remains an indispensable prerequisite for spatial data use, but its primary limitation lies in scope rather than rigor. Rule-based and standardized methods reliably detect violations of declared specifications, but they provide limited insight into the causes, consequences, or downstream impact of detected errors.

Empirical evidence from contour line digitization for DEM generation illustrates this limitation. Even when authoritative source data and clear specifications are used, human-mediated digitization introduces random, systematic, and gross errors that vary across operators (Figure 5) [64]. While such deviations are readily identified through standardized positional accuracy measures [65], technical validation alone cannot explain their causes or downstream propagation.

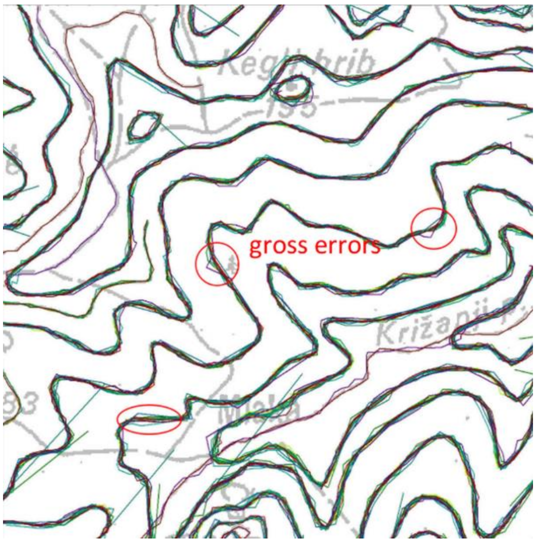


Figure 5. Visualization of the digitized contour lines validated using numerous measures.

Where uncertainty distributions are available or can be inferred, stochastic techniques such as Monte Carlo simulation extend technical quality assessment by evaluating the sensitivity of derived products to positional, thematic, or temporal uncertainty [39,66]. In DEM production, for example, such simulations enable analysis of how digitization error, interpolation strategy, and resolution interact spatially. These methods complement deterministic validation by exposing uncertainty structures, but they do not fully replace formal compliance checks.

Within the proposed framework, technical quality functions as a context-blind control layer that establishes minimum admissibility and must be completed by process, interpretive, and governance dimensions.

4.3. Process and Lifecycle

Process and lifecycle considerations determine whether technically admissible spatial data remain reliable, interpretable, and sustainable over time. Practical experience demonstrates that data quality outcomes depend less on individual validation events than on how datasets are produced, integrated, updated, and reused across successive cycles. In this respect, lifecycle management acts as a stabilizing mechanism that balances short-term compliance with long-term usability.

A central distinction emerges between static, predefined products and dynamically generated data products derived from the same underlying sources. For example, authoritative DEMs at 12.5 m, 25 m, and 100 m resolution can be produced either as fixed pre-defined products [64] (Figure 6) or generated on demand and specific requirements, both through rule-based integration of multiple heterogeneous sources. Both approaches rely on semantic spatial conflation, enabling systematic improvement through selective integration of newer or higher-quality inputs while preserving epistemic continuity [67,68]. This contrasts with repeated full re-acquisition strategies, which often lead to fragmented datasets, inconsistent uncertainty structures, and reduced comparability, especially at fine scales.

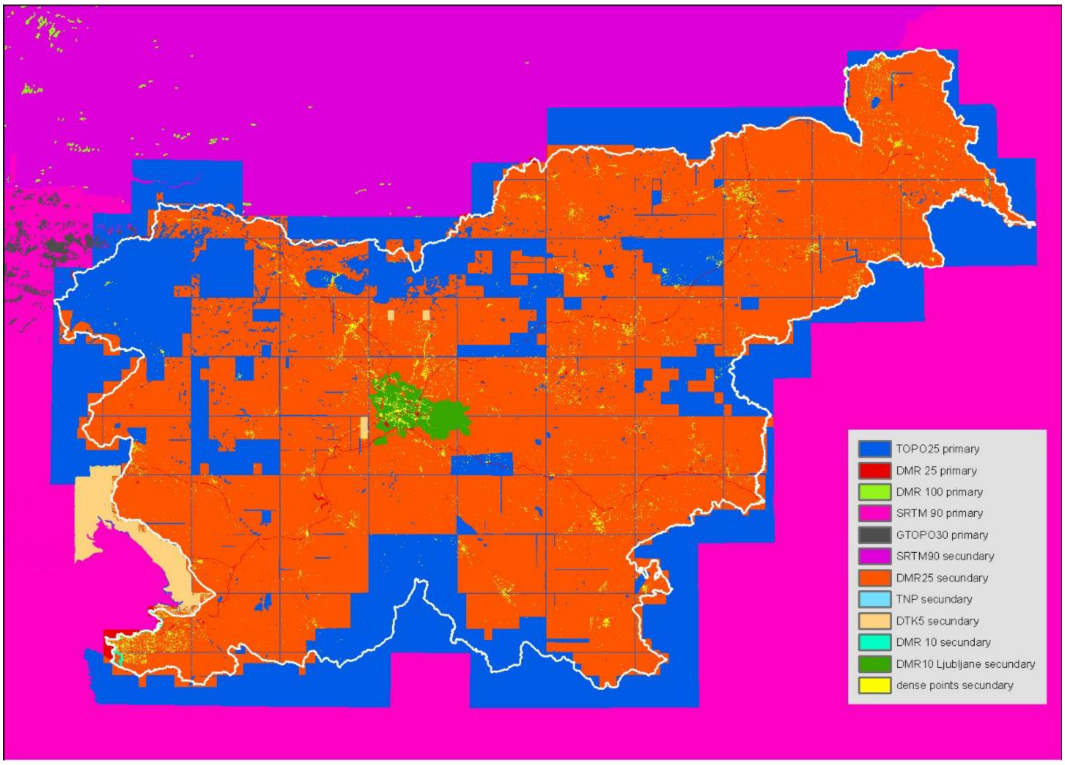


Figure 6. Conflation/integration of heterogeneous data for DEM generation: Raster-based mapping of dominant sources ranked by quality (Slovenia).

From a sustainability perspective, incremental improvement through integration and conflation is more robust than repeated replacement. For relatively stable spatial phenomena, such as terrain, cadastral frameworks, or PLU, repeated data acquisition introduces unnecessary variability and cost while discarding accumulated institutional knowledge. Stepwise correction of identified deficiencies, guided by quality diagnostics and documented through lineage, supports predictable quality evolution and aligns with circular economy principles. This outcome reflects the 1:10:100 quality cost rule (Section 3.3), where delayed detection of quality issues leads to exponential remediation costs (Figure 7).

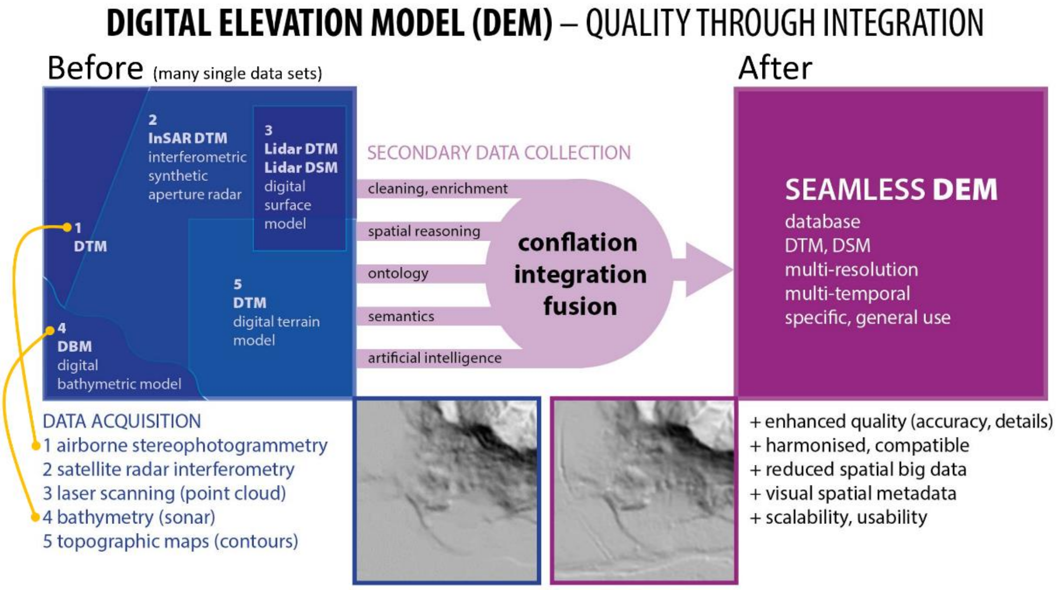


Figure 7. Process-aware lifecycle for stepwise multi-source DEM generation.

Process-aware lifecycle management further addresses structural inefficiencies commonly observed in geospatial projects, including fragmented data silos, duplicated datasets, inconsistent schemas, scale mismatches, and excessive preprocessing overhead. Shifting emphasis from repeated acquisition and ad hoc preprocessing toward standardized, interoperable, and analytics-ready data products enables more effective use of computational resources and human expertise. In this context, the role of preprocessing is not merely technical cleansing, but the creation of reusable, harmonized data layers that support multiple analytical, cartographic, and decision-making workflows.

Empirical spatial error modeling illustrates how lifecycle-aware integration enhances quality predictability. Predictive spatial error models derived from terrain parameters and known uncertainty sources produce error patterns that closely match independently observed errors, with residual discrepancies primarily attributable to interpolation artifacts or uncorrected gross errors. Such models are particularly valuable for spatial data conflation, where understanding local error behavior supports informed weighting, selective correction, and risk-aware integration of multiple datasets [69] (Figure 8).

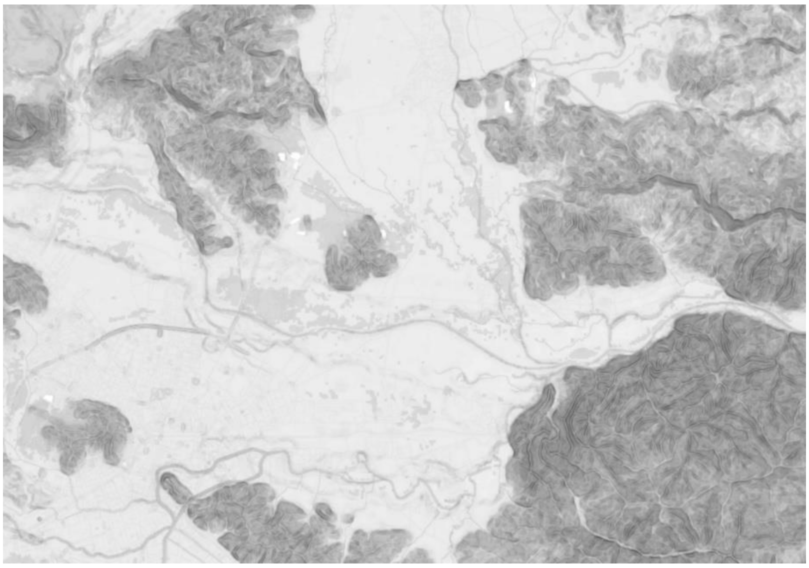


Figure 8. Spatial error model of the DEM, Ljubljana, Slovenia (area of 20 x 15 km): Darker shades indicate errors up to 8 m.

Metadata and lineage are essential to this process-oriented view. Explicit documentation of transformations, interpolation methods, generalization strategies, and update cycles enables responsible reuse and critical reinterpretation of spatial data products. When embedded within catalog-based infrastructures and aligned with FAIR principles, lifecycle documentation ensures that data remain findable, interpretable, interoperable, and reusable across institutional and temporal boundaries.

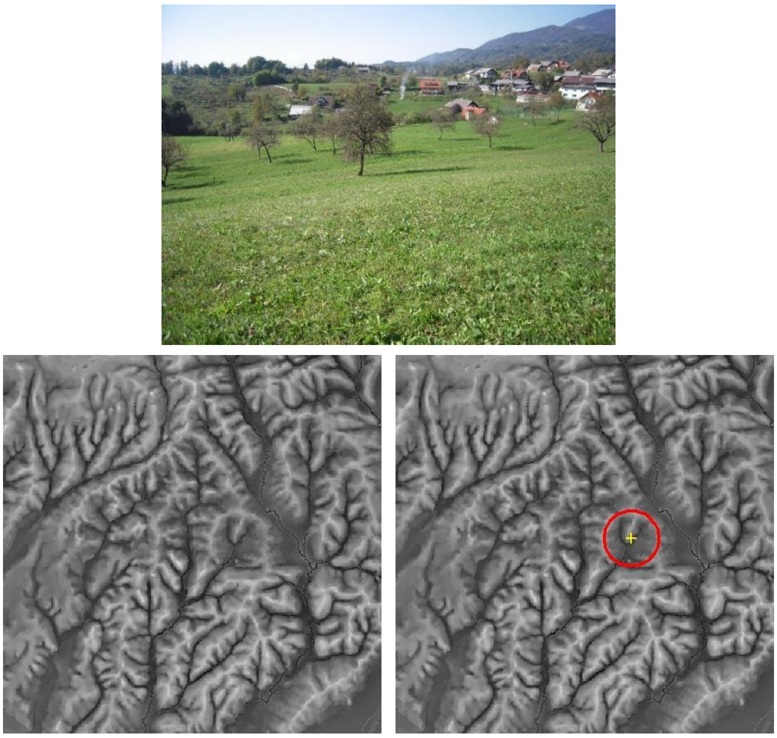
Recent advances in AI-assisted data engineering further reinforce this paradigm shift [70]. While deterministic preprocessing remains mandatory for spatial constraints, AI-supported pipelines increasingly assist in schema inference, anomaly detection, and prioritization of integration tasks. However, their effectiveness depends on being embedded within lineage-aware and governance-sensitive workflows rather than replacing structured lifecycle management.

Overall, process and and quality function as the connective layer transforming isolated validation into sustainable data stewardship. By emphasizing incremental improvement, documented integration, and predictable quality evolution, this dimension directly supports long-term sustainability, reduces data redundancy, and enables spatial data infrastructures to function as coherent, reusable ecosystems rather than collections of disposable products.

4.4. Visual and Interpretative Quality

Visual analytics complements deterministic quality assessment by serving as an early-warning mechanism that flags spatial anomalies for targeted investigation rather than providing formal validation outcomes. In practice, visually detected irregularities often indicate deeper inconsistencies in data production, preprocessing, or undocumented methodological choices that remain undetected by standards-based metrics alone.

Advanced visualization techniques, such as analytical shading (hillshading), multidirectional visibility index (MVI) [71], residual surface mapping, and contrast-enhanced overlays have proven particularly effective in signaling systematic artifacts in PLU and elevation datasets [72] (Figure 9). For example, circular or linear terrain patterns highlighted through MVI may indicate interpolation artifacts, sensor-induced biases, or remnants of legacy mapping practices rather than genuine geomorphological features. Similarly, visually plausible high-resolution LiDAR-derived DTMs may exhibit grid artifacts or “ghosting” effects around built structures, suggesting inappropriate filtering or parameterization not recorded in lineage documentation.



doi:10.20944/preprints202603.0798.v1

Figure 9. Enhancing geomorphological feature detection in Tunjice, Slovenia (above): A comparison of the MVI visualization (left) and an automated GIS-based circular detection method also based on the MVI (right) with radius of 715 m (area of 12.5 x 12.5 km). Visual quality is assessed using analog principles.

Data integration further enhances the forensic value of visual analysis. When datasets produced under different assumptions or workflows are combined, visual inconsistencies often arise that reveal misuse of specifications or undocumented shortcuts [73]. In this sense, integration acts as a stress test for SDQ. It supports informed decisions on whether datasets should be corrected, constrained in use, or excluded from further processing.

Simulation-based visualization strengthens visual forensics by providing a context for anomalies within declared uncertainty [74]. Monte Carlo-derived error realizations and comparative error fields allow analysts to differentiate persistent structural patterns from noise-driven artifacts, thereby strengthening the interpretability of visual findings [75] (Figure 10). Machine learning and AI-based methods can assist by systematically scanning large datasets for visually implausible patterns and prioritizing areas for expert review. However, interpretive judgment remains the responsibility of domain specialists.

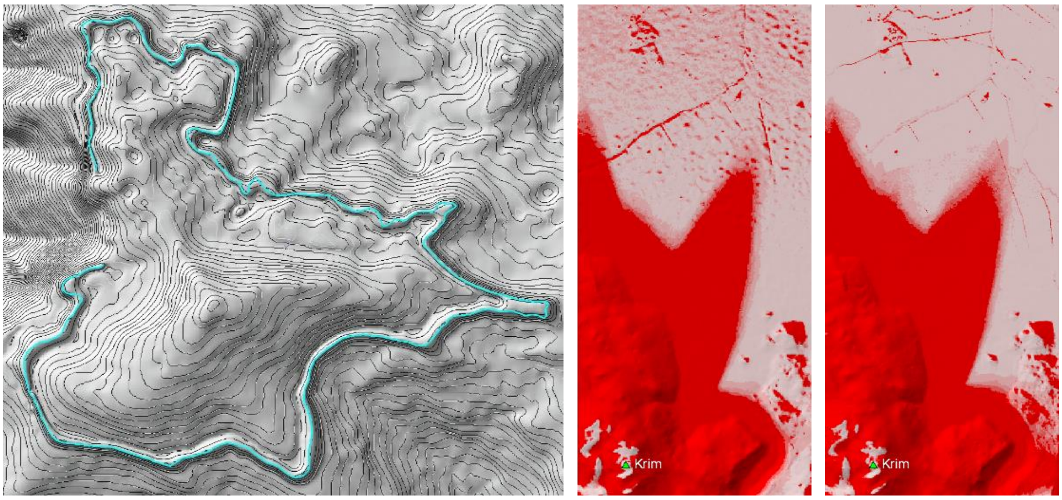


Figure 10. Visualization of potential DEM errors: (Left) Analytical shading of a DEM interpolated from contour lines containing a vertical attribute error. (Right) Viewshed simulation using Monte Carlo methods on two different DEMs, highlighting the higher precision of the rightmost model, especially in the plain areas.

Overall, spatial visual analytics strengthens SDQ management by exposing hidden assumptions, guiding corrective action, and reinforcing accountability through evidence. When coupled with lineage-aware workflows and deterministic validation, it provides a pragmatic and scalable mechanism for safeguarding trust and responsible reuse in complex spatial data environments.

4.5. Governance and Social Indicators

The semantic conflation of multiple DEM data sources of varying quality [39] illustrates how lineage functions as a governance mechanism for trust calibration, rather than as auxiliary documentation. After World War II, products with high-quality elevation attributes were generated using different sources, such as photogrammetry, LiDAR, InSAR, digitized contour lines, geodetic points, and auxiliary terrain information. Each source is characterized by distinct acquisition methods, temporal contexts, resolutions, and error structures. Transparent lineage is therefore essential for interpreting declared accuracy and managing institutional trust.

When these sources were conflated to create the integrated DEM of Slovenia, weighting schemes were explicitly based on production metadata and landscape characteristics such as terrain roughness, slope, geological structure, vegetation cover, and terrain skeleton. Lineage information used to additionally determine these weights was reconstructed not only from formal metadata, but also from project documentation, scientific publications, and expert knowledge derived from direct involvement in earlier DEM production (Figure 11). This illustrates how lineage functions as epistemic memory, preserving knowledge that would otherwise be lost during reuse.

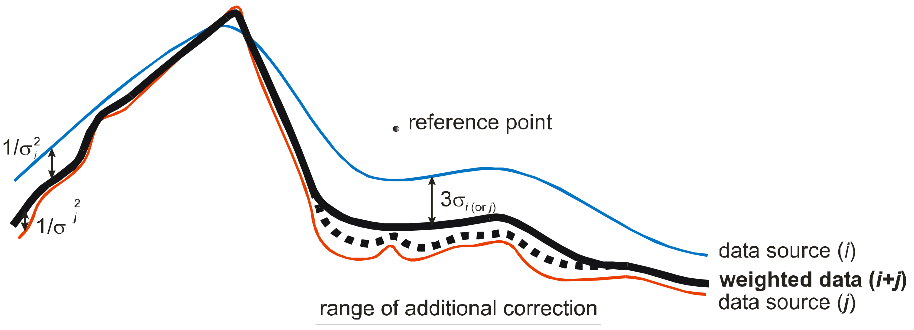


Figure 11. Integrated DEM generated using multiple DEM data sources with different weights.

When lineage was complete and coherent, the integration process produced consistent, reliable quality improvements across terrain types. On the contrary, incomplete or generic lineage can lead to misplaced trust in formally compliant datasets, resulting in overestimated local accuracy and unintended uncertainty propagation. This confirms that governance failures in SDQ rarely originate from technical deficiencies alone, but from the erosion of epistemic memory during data reuse.

Within the proposed framework, metadata and lineage therefore mediate between technical quality and institutional credibility. Rather than guaranteeing correctness, they preserve the conditions for critical interpretation, independent verification, and responsible reuse. In this sense, lineage-aware DEM integration supports sustainable spatial data stewardship by enabling incremental improvement, reducing redundant acquisition, and maintaining justified trust in authoritative elevation data across successive integration cycles.

4.6. Emergent Quality States in Practice

The primary outcome of the proposed multi-layered SDQ framework is a shift from isolated quality checks toward integrated diagnosis of emergent quality states and associated risk patterns. Rather than treating quality elements independently, the framework enables cumulative interpretation of technical, process, visual, and governance signals, translating complex conditions into operationally meaningful outcomes.

The discussion that follows draws on practical experience gained through OPIAvalid, an automated, standards-aware toolkit that operationalizes the multi-layered SDQ framework proposed in this paper. The toolkit enables stakeholders to evaluate real datasets against the DPS and associated quality requirements in a transparent and reproducible manner. By allowing municipalities, data producers, and system managers to test datasets against standardized validation procedures, OPIAvalid helps clarify how the DPS is interpreted in practice. This iterative interaction between specification, implementation, and validation not only improves the quality of individual datasets but also contributes to the progressive refinement of the DPS itself.

The first outcome is demonstrated through the implementation of the OPIAvalid toolkit for PLU data [10]. In this case, the dominant quality risk did not arise from technical non-compliance but from a governance failure rooted in the practical interpretation of the DPS. Although datasets formally satisfied ISO 19157-based technical indicators, inconsistencies emerged between the declared specification, its implementation in municipal workflows, and the expectations of downstream users. As a result, perceived dataset reliability diverged sharply from objectively measured quality.

By enforcing early source-level validation, multiple independent quality measures per attribute, and lineage-aware diagnostics, OPIAvalid exposes such discrepancies at the point of data production. This reduces ambiguity in DPS interpretation and strengthens institutional coordination

doi:10.20944/preprints202603.0798.v1

among municipalities, system managers, and administrative units. In practice, the toolkit functions not only as a validation instrument but also as a learning interface between specification, data production, and governance, enabling stakeholders to test how formal requirements operate on real datasets. The tangible outcome is more predictable decision-making in building permit procedures, demonstrating how governance-related quality risks can directly affect sustainability and public trust even when technical quality appears formally admissible.

The second outcome is illustrated through the integration of multiple DEM datasets of varying quality via semantic conflation [39]. For the newest DEMs, initial trust was placed almost exclusively in authoritative metadata. This led to error propagation when undocumented production shortcuts were present. Visual forensic analysis revealed systematic artifacts that were largely invisible to non-expert users, prompting a reassessment of producer reliability and a recalibration of integration weights based on observed quality signals. This case demonstrates how visual and governance indicators jointly influence outcomes and how failure to detect such signals can result in integrated products of lower quality than their best inputs. Incorporating visual forensics and independent validation restored control over uncertainty propagation and substantially improved the reliability of the resulting DEM products.

5. Conclusions

This study demonstrates that spatial data quality (SDQ) is not a static technical attribute, but a dynamic, socio-technical, and lifecycle-dependent system property. Limiting quality assessment to technical verification alone, which is still common in many spatial data projects, fails to capture critical risks related to interpretation, governance, and long-term sustainability. The proposed multi-layered framework addresses this limitation by integrating technical admissibility, process and lifecycle stewardship, visual and interpretive diagnostics, and governance indicators into a coherent diagnostic structure.

A central finding is that perceived data quality very often differs from objectively measured (technical) quality, with direct consequences for spatial planning, land management, and policy decisions. This discrepancy is driven less by technical non-compliance than by cognitive biases, institutional practices, and incomplete or degraded lineage. By explicitly incorporating visual spatial forensic and governance-related indicators, the framework enables early detection of methodological shortcuts, undocumented assumptions, and trust misalignments that remain invisible to conventional standards-based validation.

Operational implementation through the OPIAvalid toolkit confirms that automated, standards-aware validation can substantially improve transparency, repeatability, and collaboration among stakeholders without displacing expert judgment. Automation supports experts by reducing subjective ambiguity, enabling multiple independent quality signals, and preserving lineage as epistemic memory. Applications in planned land use (PLU) validation and multi-source digital elevation model (DEM) integration demonstrate tangible benefits for decision predictability, cost efficiency, and institutional trust.

More broadly, the framework repositions SDQ as a strategic asset within a circular data economy. Incremental improvement through integration, reuse, and lineage-aware governance is more sustainable; technically, economically, and institutionally than repeated data replacement. While AI and generative methods can enhance diagnostics, prioritization, and communication of quality risks, their use must remain constrained by deterministic validation, transparent metadata, and accountable governance, particularly in legal and policy-sensitive contexts.

In conclusion, treating SDQ as an integrated socio-technical outcome rather than a checklist-based control transforms quality management from a reactive verification task into a proactive governance instrument. When SDQ can be systematically diagnosed, communicated, and incrementally improved, spatial decisions no longer rely on implicit assumptions, but on transparent, explainable, and trustworthy evidence.

Funding: This research was funded by the Green Slovenian Location Framework (GreenSLO4D, No. 2550-22-0013) project (2021–2026) under the Recovery and Resilience Plan (development area: Digital Transformation), funded by the European Union through the NextGenerationEU instrument.

Data Availability Statement: The data and software package supporting this study are available on Zenodo at https://doi.org/10.5281/zenodo.18904636, accessed on 7 March 2026). The repository contains the OPIAvalid toolkit (v2.2.1, 64-bit Windows) implementing the spatial data quality management framework described in this paper, together with a Slovenian-language user manual and example test datasets. The software is distributed under a customized license based on the Apache-2.0 structure. Additional information about the toolkit is available at https://pis.eprostor.gov.si/pis/orodja.

Acknowledgments: The author gratefully acknowledges Jurij Mlinar and his team for their valuable support, as well as all participating stakeholders for their feedback during the development and testing of the OPIAvalid toolkit.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
BIM	Building information modeling
CRS	Coordinate reference system
CQI	Composite quality index
DEM	Digital elevation model
DPS	Data product specification
FAIR	Findable, accessible, interoperable, reusable
GIS	Geographic information system
ISO	International Organization for Standardization
ML	Machine learning
OPIAvalid	Automated toolkit for spatial data validation (formerly NRPvalid)
PLU	Planned land use
SDGs	Sustainable Development Goals
SDQ	Spatial data quality
TDQM	Total data quality management
TRUST	Transparency, responsibility, user focus, sustainability, and technology)
UN-GGIM	United Nations Committee of Experts on Global Geospatial Information Management

References

1. Goodchild, M.F. *Fundamentals of Spatial Data Quality*; Devillers, R., Jeansoulin, R., Eds.; London, 2006; 312 ISBN 978-1-905209-56-9.
2. *ISO 19157-1:2023(En), Geographic Information — Data Quality — Part 1: General Requirements*; **2023**.
3. López, F.; González, P.; Masó, J.; Zabala Torres, A.; Pascual, A.; Vergara, G.; García-Balboa, J. Geospatial Data Quality (ISO 19157-1): Evolve or Perish. *Rev. Cartográfica* **2020**, 129–154. https://doi.org/10.35424/rcarto.v0i100.692.
4. Crosby, P.B. *Quality Is Free: The Art of Making Quality Certain*; McGraw-Hill, 1979; 326 ISBN 978-0-07-014512-2.
5. Juran, J.M. *Juran on Quality by Design: The New Steps for Planning Quality Into Goods and Services*; Simon and Schuster, 1992; 552 ISBN 978-0-02-916683-3.
6. Batini, C.; Scannapieco, M. *Data and Information Quality: Dimensions, Principles and Techniques*; Springer: Cham, 2016; 528 ISBN 978-3-319-24104-3.
7. Strong, D.M.; Lee, Y.W.; Wang, R.Y. Data Quality in Context. *Commun. ACM* **1997**, *40*, 103–110. https://doi.org/10.1145/253769.253804.

8. Kitchin, R. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*; SAGE Publications Ltd: 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom, 2014 ISBN 978-1-4462-8748-4.
9. Goodchild, M.F. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* **2007**, *69*, 211–221. https://doi.org/10.1007/s10708-007-9111-y.
10. Podobnikar, T. Bridging Perceived and Actual Data Quality: Automating the Framework for Governance Reliability. *Geosciences* **2025**, *15*. https://doi.org/10.3390/geosciences15040117.
11. Wang, R.Y.; Strong, D.M. Beyond Accuracy: What Data Quality Means to Data Consumers. *J. Manag. Inf. Syst.* **1996**, *12*, 5–33. https://doi.org/10.1080/07421222.1996.11518099.
12. Heilbrun, K.; Grisso, T.; Goldstein, A.M.; LaDuke, C. Foundations of Forensic Mental Health Assessment. In *Forensic assessments in criminal and civil law: A handbook for lawyers*; Best practices in forensic mental health assessment; Oxford University Press: New York, NY, US, 2013; pp. 1–14 ISBN 978-0-19-976685-7.
13. El-Shenawy, O.E. Traditional Psychological Tests Usage in Forensic Assessment. *Forensic Leg. Investig. Sci.* **2017**, *3*, 1–5. https://doi.org/10.24966/FLIS-733X/100020.
14. Kruger, J.; Dunning, D. Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *J. Pers. Soc. Psychol.* **1999**, *77*, 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121.
15. Neal, T.M.S.; Slobogin, C.; Saks, M.J.; Faigman, D.L.; Geisinger, K.F. Psychological Assessments in Legal Contexts: Are Courts Keeping "Junk Science" Out of the Courtroom? *Psychol. Sci. Public Interest* **2019**, *20*, 135–164. https://doi.org/10.1177/1529100619888860.
16. Kahneman, D. *Thinking, Fast and Slow*; Thinking, fast and slow; Farrar, Straus and Giroux: New York, NY, US, 2011; p. 499; 499 ISBN 978-0-374-27563-1.
17. Murrie, D.C.; Boccaccini, M.T.; Guarnera, L.A.; Rufino, K.A. Are Forensic Experts Biased by the Side That Retained Them? *Psychol. Sci.* **2013**, *24*, 1889–1897. https://doi.org/10.1177/0956797613481812.
18. Buonomano, D. *Brain Bugs: How the Brain's Flaws Shape Our Lives*; Erscheinungsort nicht ermittelbar, 2012; 322 ISBN 978-0-393-34222-2.
19. Bell, J., & Holroyd, J. (2009). Review of Human Reliability Assessment Methods. Health and Safety Laboratory. - References - Scientific Research Publishing Available online: https://www.scirp.org/reference/referencespapers?referenceid=3719591 (accessed on 8 February 2026).
20. Reason, J. *Human Error*; 1st ed.; Cambridge University Press, 1990 ISBN 978-0-521-30669-0.
21. Dror, I.E. Cognitive and Human Factors in Expert Decision Making: Six Fallacies and the Eight Sources of Bias. *Anal. Chem.* **2020**, *92*, 7998–8004. https://doi.org/10.1021/acs.analchem.0c00704.
22. Dekker, S. *Safety Differently: Human Factors for a New Era, Second Edition*; 2nd ed.; CRC Press: Boca Raton, 2014; 312 ISBN 978-0-429-16297-8.
23. Norman, D.A. *The Design of Everyday Things*; Rev. and expanded edition.; MIT press: Cambridge (Mass.), 2013 ISBN 978-0-465-05065-9.
24. Reverberi, C.; Rigon, T.; Solari, A.; Hassan, C.; Cherubini, P.; Cherubini, A. Experimental Evidence of Effective Human–AI Collaboration in Medical Decision-Making. *Sci. Rep.* **2022**, *12*, 14952. https://doi.org/10.1038/s41598-022-18751-2.
25. Romeo, G.; Conti, D. Exploring Automation Bias in Human–AI Collaboration: A Review and Implications for Explainable AI. *AI Soc.* **2025**. https://doi.org/10.1007/s00146-025-02422-7.
26. *States of Knowledge: The Co-Production of Science and the Social Order*; Jasanoff, S., Ed.; Routledge: London, 2004; 336 ISBN 978-0-203-41384-5.
27. UNSD — UN-GGIM Available online: https://ggim.un.org/UN-IGIF/ (accessed on 21 January 2026).
28. *INSPIRE: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community*; **2007**.
29. Latour, B. *Science in Action: How to Follow Scientists and Engineers Through Society*; Harvard University Press, 1987; 292 ISBN 978-0-674-79291-3.
30. Ramchurn, S.D.; Huynh, D.; Jennings, N.R. Trust in Multi-Agent Systems. *Knowl. Eng. Rev.* **2004**, *19*, 1–25. https://doi.org/10.1017/S0269888904000116.

31. Stern, M.J.; Coleman, K.J. The Multidimensionality of Trust: Applications in Collaborative Natural Resource Management. *Soc. Nat. Resour.* **2015**, *28*, 117–132. https://doi.org/10.1080/08941920.2014.945062.
32. O'Shee, K.D. What Is Data Lineage? Tracking Data Through Enterprise Systems. *Graph Database Anal.* 2025.
33. Foote, K.D. Data Lineage and Data Quality: How They Intersect. *Dataversity* **2022**.
34. *ISO 19115-1 Geographic Information — Metadata — Part 1: Fundamentals.*;
35. Lin, D.; Crabtree, J.; Dillo, I.; Downs, R.R.; Edmunds, R.; Giaretta, D.; De Giusti, M.; L'Hours, H.; Hugo, W.; Jenkyns, R.; et al. The TRUST Principles for Digital Repositories. *Sci. Data* **2020**, *7*, 144. https://doi.org/10.1038/s41597-020-0486-7.
36. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, Ij.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3*, 160018. https://doi.org/10.1038/sdata.2016.18.
37. Veregin, H. Data Quality Parameters. In *Geographical Information Systems*; Scientific Research Publishing, 1999; pp. 177–189.
38. Masser, I. All Shapes and Sizes: The First Generation of National Spatial Data Infrastructures. *Int. J. Geogr. Inf. Sci.* **1999**, *13*, 67–84. https://doi.org/10.1080/136588199241463.
39. Podobnikar, T. Production of Integrated Digital Terrain Model from Multiple Datasets of Different Quality. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 69–89. https://doi.org/10.1080/13658810412331280130.
40. Olson, J.E. *Data Quality: The Accuracy Dimension*; San Francisco, 2003; 312 ISBN 978-1-55860-891-7.
41. *Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations, General Assembly, 2015.
42. Andrienko, N.; Andrienko, G. *Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach*; Berlin, Heidelberg, 2006; 718 ISBN 978-3-540-31190-4.
43. MacEachren, A.M.; Gahegan, M.; Pike, W.; Brewer, I.; Cai, G.; Lengerich, E.; Hardistry, F. Geovisualization for Knowledge Construction and Decision Support. *IEEE Comput. Graph. Appl.* **2004**, *24*, 13–17. https://doi.org/10.1109/MCG.2004.1255801.
44. Hundt, M.; Siirak, N.M.; Wildnei, M. Visual Analytics for Detecting Behaviour Patterns in Geo-Temporal Data. In Proceedings of the 2014 IEEE Conference on Visual Analytics Science and Technology (VAST); October 2014; pp. 355–356.
45. Ieremeiev, O.; Lukin, V.; Okarma, K.; Egiazarian, K. Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images. *Remote Sens.* **2020**, *12*. https://doi.org/10.3390/rs12152349.
46. Selva, M. The Quality of Remote Sensing Optical Images from Acquisition to Users. *Remote Sens.* **2021**, *13*. https://doi.org/10.3390/rs13071295.
47. Sanders, J.D.; Katz, S. The Overuse and Misuse of Psychological Testing: Why Less Is More. *Am J Fam L* **2013**, *26*, 221–226.
48. Archer, R.P.; Baum, L.J. Forensic Uses of Clinical Assessment Instruments. In *APA handbook of psychology and juvenile justice*; APA handbooks in psychology®; American Psychological Association: Washington, DC, US, 2016; pp. 425–443 ISBN 978-1-4338-1967-4.
49. Elwood, S.; Goodchild, M.F.; Sui, D.Z. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Ann. Assoc. Am. Geogr.* **2012**, *102*, 571–590. https://doi.org/10.1080/00045608.2011.595657.
50. Egenhofer, M.J.; Mark, D.M. Naive Geography. In Proceedings of the Spatial Information Theory A Theoretical Basis for GIS; Frank, A.U., Kuhn, W., Eds.; Springer: Berlin, Heidelberg, 1995; pp. 1–15.
51. Dekker, S. *The Field Guide to Understanding "Human Error"*; Farnham, Surrey, England ; Burlington, VT, USA, 2014; 248 ISBN 978-1-4724-3905-5.
52. Ioannidis, J.P.A. Why Most Published Research Findings Are False. *PLOS Med.* **2005**, *2*, e124. https://doi.org/10.1371/journal.pmed.0020124.
53. Haklay, M. How Good Is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environ. Plan. B Plan. Des.* **2010**, *37*, 682–703. https://doi.org/10.1068/b35097.

54. Lackey, J. Memory as a Generative Epistemic Source. *Philos. Phenomenol. Res.* **2005**, *70*, 636–658. https://doi.org/10.1111/j.1933-1592.2005.tb00418.x.
55. Yang, X.; Blower, J.D.; Bastin, L.; Lush, V.; Zabala, A.; Masó, J.; Cornford, D.; Díaz, P.; Lumsden, J. An Integrated View of Data Quality in Earth Observation. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **2013**, *371*, 20120072. https://doi.org/10.1098/rsta.2012.0072.
56. Fischer, J.; Egli, L.; Groth, J.; Barrasso, C.; Ehrmann, S.; Figgemeier, H.; Henzen, C.; Meyer, C.; Müller-Pfefferkorn, R.; Rümmler, A.; et al. Approaches and Tools for User-Driven Provenance and Data Quality Information in Spatial Data Infrastructures. *Int. J. Digit. Earth* **2023**, *16*, 1510–1529. https://doi.org/10.1080/17538947.2023.2198778.
57. Hecht, B.; Stephens, M. A Tale of Cities: Urban Biases in Volunteered Geographic Information. *Proc. Int. AAAI Conf. Web Soc. Media* **2014**, *8*, 197–205. https://doi.org/10.1609/icwsm.v8i1.14554.
58. 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. https://doi.org/10.1080/17538947.2023.2200042.
59. Ethics Guidelines for Trustworthy AI | Shaping Europe’s Digital Future Available online: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed on 31 January 2026).
60. MacEachren, A.M.; Robinson, A.; Hopper, S.; Gardner, S.; Murray, R.; Gahegan, M.; Hetzler, E. Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know. *Cartogr. Geogr. Inf. Sci.* **2005**, *32*, 139–160. https://doi.org/10.1559/1523040054738936.
61. Reason, J. *Managing the Risks of Organizational Accidents*; Routledge: London, 2016; 272 ISBN 978-1-315-54354-3.
62. Abdulmuttalib, H.; Al Doori, M.; Barsi, Á.; Blaschke, T.; Gao, Y.; Kugler, Z.; Lang, S.; Szabo, G.; Tiede, D. Remote Sensing Data Quality in the Era of AI. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *X-3–2024*, 1–11. https://doi.org/10.5194/isprs-annals-X-3-2024-1-2024.
63. Ministry of Natural Resources and Spatial Planning, ORODJA (Tools) Available online: https://pis.eprostor.gov.si/pis/orodja (accessed on 9 February 2026).
64. Surveying and Mapping Authority, Digitalni modeli višin (Digial elevation models) Available online: https://www.e-prostor.gov.si/podrocja/drzavni-topografski-sistem/digitalni-modeli-visin/ (accessed on 10 February 2026).
65. Podobnikar, T.; Vidiček, M. Positional and Thematic Digitisation Error Assessment in Practices. *Proc. Ninth Int. Symp. Spat. Accuracy Assess. Nat. Resour. Environ. Sci.* **2010**, 179. https://doi.org/10.13140/RG.2.2.29444.27527.
66. Heuvelink, G.B.M. Error Propagation in Environmental Modelling With GIS. In; Taylor & Francis: London, Bristol, 1998.
67. Podobnikar, T.; Vrečko, A. Digital Elevation Model from the Best Results of Different Filtering of a LiDAR Point Cloud. *Trans. GIS* **2012**, *16*, 603–617. https://doi.org/10.1111/j.1467-9671.2012.01335.x.
68. Podobnikar, T. Production of Integrated Digital Terrain Model from Multiple Datasets of Different Quality. *Int. J. Geogr. Inf. Sci.* **2005**, *19*, 69–89. https://doi.org/10.1080/13658810412331280130.
69. Podobnikar, T. Combination of Error Characterization and Spatial Error Model to Improve Quality of Digital Elevation Models. In Proceedings of the Proceedings of Spatial Accuracy; 2016; pp. 241–248.
70. Joshi, S. Review of Data Pipelines and Streaming for Generative AI Integration: Challenges, Solutions, and Future Directions. *Int. J. Res. Publ. Rev.* **2025**, *6*, 2348–2357. https://doi.org/10.55248/gengpi.6.0225.0919.
71. Podobnikar, T. Multidirectional Visibility Index for Analytical Shading Enhancement. *Cartogr. J.* **2012**, *49*, 195–207. https://doi.org/10.1179/1743277412Y.0000000012.
72. Podobnikar, T. Senčenje Reliefa Za Večstransko Uporabo (Relief Shading for Versatile Use). *Proteus* **2013**, *75*, 355–360.
73. Podobnikar, T. Nadgradnja Modela Reliefa Slovenije z Visokokakovostnimi Podatki : High-Quality Data for Enhancement of the Terrain Model of Slovenia. *Geod. Vestn.* **2008**, *52*.

74. Podobnikar, T. Simulation and Representation of the Positional Errors of Boundary and Interior Regions in Maps. In *Geospatial Vision: New Dimensions in Cartography*; Lecture Notes in Geoinformation and Cartography; Springer, 2008; pp. 141–169.
75. Podobnikar, T. Methods for Visual Quality Assessment of a Digital Terrain Model. *SAPIENS Surv. Perspect. Integrating Environ. Soc.* **2009**, *2*, 15–24.