

Article

From Stars to LETTERS: A Multi-Dimensional, FAIR-Aligned Framework for Geospatial Metadata Quality Evaluation

Claire Ponciano ¹, Falk Würriehausen ², Markus Schaffert ¹, Hartmut Müller ^{1,*} and Jean-Jacques Ponciano ¹

¹ i3mainz, Institute for Spatial Information and Surveying Technology, Department of Engineering, Mainz University of Applied Sciences, Lucy-Hillebrand-Straße 2, 55128 Mainz, Germany; claire.ponciano@hs-mainz.de (C.P.); markus.schaffert@hs-mainz.de (M.S.); jean-jacques.ponciano@hs-mainz.de (J.-J.P.)

² Federal Agency for Cartography and Geodesy, Richard-Strauss-Allee 11, 60598 Frankfurt, Germany; falk.wuerriehausen@bkg.bund.de

* Correspondence: hartmut.mueller@hs-mainz.de

Abstract

Star-based schemes, such as the 5-star Linked Open Data model and its geospatial extensions, are widely used to characterize openness and interoperability. However, in practice, higher star ratings are often assigned on the basis of technical properties such as RDF exposure or schema publication without requiring the satisfaction of foundational metadata quality conditions. This weakens the monotonic interpretation of star levels and can produce ambiguous signals for data users. To address this issue, we propose the LETTER framework, a multi-dimensional evaluation model in which seven independent binary dimensions describe metadata readiness for reuse: Provenance (P), Access (A), Structure (S), Connections (C), License (L), Identifiers (I), and Quality (Q). The framework is aligned with FAIR principles and mapped to ISO 19115 and ISO 19157 concepts. We evaluate it through an exploratory comparative case study of four purposively selected datasets from the German Spatial Data Infrastructure (GDI-DE): Municipal Points of Interest (Trier), Thuringia Digital Elevation Model (DEM), Administrative Units (VG250), and INSPIRE Digital Land Model (DLM). The results show that datasets receiving comparatively high star ratings may still lack machine-actionable provenance, quality evidence, stable identifiers, or robust access conditions. In particular, the analysis highlights a recurring ‘PDF Trap’, where relevant trust information exists only in narrative documentation and therefore remains inaccessible to automated reuse workflows. We conclude that LETTER provides clearer diagnostic power than scalar star ratings by exposing which metadata functions are actually satisfied and which remain missing.

Keywords: metadata; geospatial data; FAIR principles; data quality; interoperability; ISO 19115; Linked Open Data



Academic Editors: Wolfgang Kainz, Dev Raj Paudyal, Jingzhong Li and Binbin Lu

Received: 31 January 2026

Revised: 25 March 2026

Accepted: 1 April 2026

Published: 2 May 2026

Copyright: © 2026 by the authors.

Published by MDPI on behalf of the International Society for Photogrammetry and Remote Sensing. Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

1.1. Motivation: Reliable Metadata in Data-Intensive Geosciences

The volume of heterogeneous geospatial data has grown exponentially, driven by National Mapping Agencies (NMAs), open government portals, and environmental sensor networks [1,2]. While initiatives like INSPIRE [3] and national Spatial Data Infrastructures (SDI) have improved discoverability [4], a critical challenge remains: ensuring the “fitness for purpose” of discovered data [5,6].

The role of spatio-temporal information in governance has been highlighted by recent global crises, such as the COVID-19 pandemic, where the lack of standardized, reliable metadata hampered rapid decision-making across European borders [7]. Without transparent metadata describing the dataset itself (data metadata)—specifically its origin, quality, structure, and intended use—even openly accessible geospatial data remains difficult to interpret [8]. This distinction is crucial in infrastructures like INSPIRE or GDI-DE, where service metadata (e.g., for WMS) may be present, but the underlying data metadata required for fitness-for-use assessment is often missing or incomplete [8]. This is particularly true for Volunteered Geographic Information (VGI), where heterogeneity is inherent and quality assurance often relies on collaborative mechanisms rather than authoritative validation [9,10]. The FAIR principles (Findable, Accessible, Interoperable, Reusable) [11] extend the original Semantic Web vision [12] and provide a theoretical foundation for scientific data stewardship [13], but implementing these abstract principles in the geospatial domain requires operational models (e.g., maturity indicators, validation frameworks, or certification schemes) to translate principles into measurable technical requirements [14].

In this context, fitness for purpose should not be understood as a single universal property of a dataset but as the degree to which the metadata provides sufficient evidence for a given reuse decision. The framework proposed in this paper is therefore designed not to replace use-case-specific evaluation but to make explicit whether the metadata contains the machine-actionable evidence needed to judge suitability for particular analytical, legal, or interoperability requirements.

1.2. STAR-Based Evaluation and Its Ambiguities

To simplify assessment, the community has largely adopted “Star” thinking, originating from Berners-Lee’s 5-star Linked Open Data (LOD) scheme [15–17] illustrated in Figure 1.

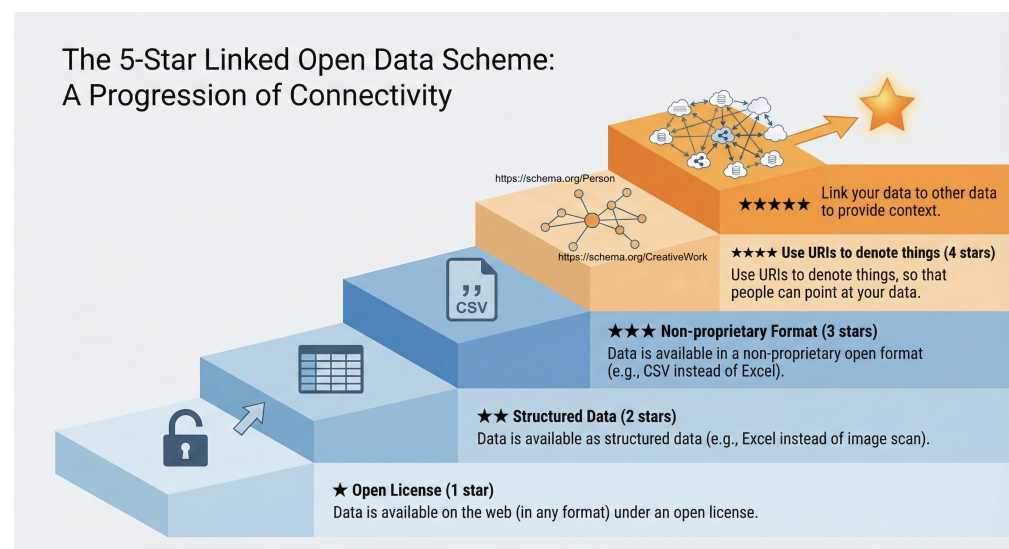


Figure 1. The classic 5-star Linked Open Data scheme [15], illustrating the monotonic progression from Open License (1 star) to Linked Data (5 stars).

Extensions to this model, such as the 8-star geospatial model or vocabulary-specific stars [18], attempt to add criteria for quality and trust. However, our analysis suggests these models suffer from a monotonicity failure. A dataset can reach “5 stars” (Linked Data) or even “8 stars” (Trusted Data) based on partial criteria, masking significant deficits [19]. For example, a dataset might be fully linked (RDF) but lack basic provenance [20], while a 3-star CSV file might be meticulously documented. This ambiguity leads to decision-makers

interpreting “stars” as a generic quality label rather than a specific technical interoperability score [21].

1.3. Research Questions

This paper addresses three questions:

- **RQ1:** How do linear “Star” schemes fail to capture the multi-dimensional nature of geospatial data quality described in standard frameworks [22,23]?
- **RQ2:** Can a mnemonic, non-hierarchical taxonomy (LETTER) improve diagnostic clarity regarding FAIR compliance?
- **RQ3:** How does the LETTER framework perform when applied to real-world datasets (Trier, Thuringia, VG250, INSPIRE) compared to the 8-star model?

2. Related Work

2.1. FAIR Principles and Machine Actionable Metadata

The FAIR Guiding Principles [11] have fundamentally reshaped data management by emphasizing that data must be Findable, Accessible, Interoperable, and Reusable. A distinguishing feature of FAIR is its specific focus on machine actionability, ensuring that computational systems can find, access, operate on, and reuse data with minimal human intervention [13].

This paradigm has been widely adopted in policy and research infrastructures, influencing research data policies and evaluation frameworks across multiple disciplines. González-Beltrán et al. [24] highlight how metadata is central to realizing these principles, transforming them from theoretical ideals into practical requirements for data repositories. However, in the context of geospatial metadata, FAIR implementation is still evolving. Recent research has begun to bridge the gap between theory and practice: Thompson et al. proposed a workflow for FAIR-compliant GIS map dissemination [25], and Degbelo established definitions and challenges for FAIR geovisualizations [26]. Despite these advances in defining workflows and challenges, FAIR often remains difficult to quantify in existing portals. While these initiatives provide robust frameworks, they do not yet function as a concrete scoring system or mnemonic metric that allows for rapid, comparative assessment of geospatial datasets by non-experts [14,27].

2.2. ISO 19115, ISO 19157 and Associated Geospatial Metadata Standards

The geospatial domain relies on the ISO 19100 series as its documentation backbone. ISO 19115-1:2014 serves as the core geospatial metadata standard, defining a comprehensive schema for describing the identification, quality, spatial and temporal extent, distribution, and constraints of geographic data and services [28]. This is complemented by ISO 19157:2013 [23], which establishes principles and components for describing and reporting data quality. It includes specific measures for completeness, positional accuracy, temporal quality, consistency, and thematic accuracy, providing a rigorous vocabulary for quality reporting. Crucially, the standard emphasizes that these measures are not absolute but must be evaluated against specific product specifications or use cases to determine fitness for purpose. Further extensions, such as ISO 19165-1:2018 [29], address long-term preservation metadata and provenance, ensuring the durability of digital geospatial assets.

Despite the richness of these standards, they face significant limitations in practice. and the prevalence of optional elements often leads to the ‘blank field’ syndrome—where metadata records are technically compliant but lack content in non-mandatory fields [30]. Furthermore, while they enable detailed reporting via free text fields, they do not directly provide a simple external “badge” or multi-dimensional code that summarizes quality in a way that is immediately interpretable by non-experts [31].

2.3. Metadata Quality Assessment and Usability Studies

Empirical studies highlight a gap between standard compliance and actual usability. Large-scale analyses of over 160,000 datasets in open data portals have demonstrated generally poor metadata quality and a weak correlation between usage metrics and metadata completeness scores [21,32]. This suggests that mere adherence to schema fields does not guarantee utility.

User studies corroborate this, reporting persistent difficulties in discovering and selecting spatial data. Even when ISO-compliant metadata is present, users struggle due to inconsistent content, extensive use of free text, and poor interface design [32,33]. While research has explored visual approaches for metadata selection—such as the multivariate visualizations proposed by Ahonen-Rainio [34]—these have not yet become standard features in mainstream SDIs. Consequently, practical operational guides still emphasize text-based audits, completeness checks, and the use of controlled vocabularies rather than relying on compact visual summary metrics [35].

2.4. State of the Art: Evolving Star-Based Schemes

To simplify openness assessment, the community widely cites Tim Berners-Lee’s 5-star Linked Open Data (LOD) scheme [15]. This model defines incremental steps from “available on the web” (1 star) to “linked data” (5 stars), serving as a powerful advocacy tool for technical interoperability.

However, researchers have long recognized that simple star ratings must be supported by explicit metadata to be effective. Early critiques by Martin et al. highlighted that star ratings often fail to communicate the actual openness level without accompanying metadata [36]. Consequently, significant efforts have been made to extend the model beyond five stars to address quality and trust. Hyvönen et al. introduced a **7-star model** for Linked Data Finland, adding requirements for validating schemas and documenting the definition of data elements [37]. This was recently expanded to an **8-star model**, which explicitly incorporates criteria for data quality, trust, and reuse [38]. Similarly, Al-khatib and Ali proposed a framework that adds a **validation layer** using SHACL shapes to the 5-star stack, ensuring that datasets are not just linked, but structurally valid [17].

Parallel efforts have aligned stars with domain-specific needs. In the geospatial domain, Wilson et al. proposed a **5-star guide for GIS**, specifically targeting replicability and reproducibility in spatial algorithms [39]. Furthermore, the relationship between these schemes and the FAIR principles has been formalized by Hasnain and Rebholz-Schuhmann, who mapped 5-star criteria against FAIR facets to highlight overlaps in interoperability [40].

Table 1 summarizes and compares the different definitions and extensions of star-based approaches: from the original technique [15], through the extension from 5 to 7 stars proposed for data quality and reusability [37], then the extension from 7 to 8 stars proposed for trust [38], and the application to the geographical domain aimed at scientific reproducibility [39].

Despite these advances, a “monotonicity gap” remains. While the 8-star model addresses quality, it remains a “ladder”: a dataset might fail a lower rung (e.g., proper licensing) but technically achieve a higher rung (e.g., linked data formats), creating ambiguity [21]. Our LETTER framework builds on these “validation” and “trust” extensions but re-organizes them into independent, non-hierarchical dimensions to prevent this ambiguity.

Table 1. Alignment of star-based progression scales.

Level	Berners-Lee (2006) [15] (The Web Standard)	Hyvönen et al. (2014, 2024) [37,38] (Extension: Quality & Trust)	Wilson et al. (2021) [39] (Extension: GIS Reproducibility)
★	Open License: data available on the Web under an open license (e.g., PDF, image).	Same as Berners-Lee	Data and Code: simple sharing of raw data and source code.
★★	Structured: machine-readable data but in a proprietary spreadsheet or office format.	Same as Berners-Lee	Basic Metadata: contextual information (ReadMe files, software environment description).
★★★	Open Format: non-proprietary format (e.g., CSV, JSON).	Same as Berners-Lee	Structured Metadata: complete and structured provenance.
★★★★	URI / RDF: use of unique identifiers to denote things.	Same as Berners-Lee	Standards & APIs: geospatial standards and links to appropriate APIs.
★★★★★	Linked Data (LOD): data linked to other data to provide context.	Same as Berners-Lee	Encapsulated Environment: full reproducibility via standard container technologies.
★★★★★★	(Not defined)	Explicit Schemas: publication of schemas/vocabularies used to describe data.	(Not defined)
★★★★★★★	(Not defined)	Validation: proof that data validates against published schemas.	(Not defined)
★★★★★★★★	(Not defined)	Veracity / Trust: documentation of correspondence to the real world (factually correct).	(Not defined)

2.5. Gap Analysis

Current approaches provide the necessary building blocks but lack a unified communication tool. We have solid conceptual frameworks (FAIR, ISO) and operational multi-star ladders (5-star LOD). We also have empirical analyses detailing the state of metadata quality.

The missing element is a compact, mnemonic, multi-dimensional notation that:

- Explicitly separates **Provenance, Access, Structure, Connections, License, Identifiers, and Quality**.
- Defines strong, binary criteria for each dimension to avoid ambiguity.
- Can be encoded directly inside ISO/DCAT metadata and used as an external label or “badge” for users.

The proposed LETTER framework aims to fill this gap.

3. Problem Analysis: Limitations of STAR Based Evaluation for Geodata

3.1. Conceptual Shortcomings of Linear Star Ladders

Linear “Star” schemes rely on a **monotonicity assumption**, where higher stars are intended to imply the full satisfaction of all lower-level requirements. In this theoretical model, a 5-star dataset is inherently “better” than a 3-star one. However, real-world workflows often assign stars independently based on specific technical features (e.g., the presence of RDF). Consequently, a “5-star” dataset may technically achieve Linked Data status while lacking basic metadata hygiene expected at lower levels, such as a clear license or schema documentation [21].

Furthermore, star schemes suffer from an **aggregation problem**. Distinct aspects of data usability—such as file format, linking capability, and content quality—are compressed into a single scalar number. This single-score approach hides critical trade-offs, making diagnosis difficult. A user cannot discern if a low score is due to a proprietary format (Structure) or a lack of documentation (Quality). Finally, there is a **scope mismatch**: star schemes originally targeted openness and linked data characteristics, not the full spectrum of geospatial quality, provenance, or identifier stability required for trusted spatial analysis [18].

3.2. Empirical Symptoms in Real Geospatial and Open Data Portals

The limitations of the linear model become evident when applying it to typical datasets found in Spatial Data Infrastructures (SDIs). Let us take the following two examples (vector and raster) observed in our study:

- **Municipal Points of Interest (POI):** Municipal POI dataset access and information are available on https://geoportal.trier.de/trier/mod_ogc/ (accessed on 31 March 2026). These datasets often appear as structured data in GML and JSON formats. While they typically achieve 3 stars for their open format in the original model of Berners-Lee [15], and potentially 6 stars in the extended model of Hyvönen et al. [37], if a schema is published (as seen in the Trier use case), they frequently lack stable dataset identifiers.
- **Digital Elevation Models (DEMs) of Thuringia:** Thuringia DEM datasets access and information are available on <https://geoportal.thueringen.de/gdi-th/download-offene-geodaten/download-hoehendaten> (accessed on 31 March 2026). These datasets typically achieve 5 stars in the original model of Berners-Lee [15] due to the use of non-proprietary formats, open license, open standards (OGC WMS) and link to INSPIRE. They also achieve **7 stars** in the model of Hyvönen et al. [37] thanks to available schema and the presence of quality information inside making them machine readable.

Municipal POI datasets achieve easily a baseline rating of 3 stars in Berners-Lee’s model [15] thanks to open license, formats and standards, but they achieve ratings of 6 stars in the model of Hyvönen et al. [37], even though they do not meet the basic criteria for connectivity and semantic links. These observations highlight the problem of ‘non-monotonic jumps’ in models extended to 7 and 8 stars. Thuringia DEM datasets achieve a good star ratings in both models, despite unstable access and a lack of stable identifiers. These observations highlight that star ratings obscure a lack of important aspects of data.

3.3. Impact on Scalability, Interoperability, and Comprehension

The reliance on star schemes has broader systemic impacts:

- **Scalability:** As new concerns such as provenance or quality arise, the community tends to invent new star variants or extended ladders (moving from 5 to 7, then 8 stars). This fragments the ecosystem, as different portals adopt different “ladders” [19].

- **Interoperability:** Different catalogues and projects interpret star levels differently. One portal's "3 stars" might require CSV, while another's requires a formalized API. This lack of standardization makes it difficult to translate assessments across borders or domains.
- **Comprehension:** Non-experts tend to read "5 or 8 stars" as a generic quality label—similar to a hotel rating—without insight into which specific aspects of the metadata are reliable.

Moreover, if we consider a linear star rating system where the highest level cannot be achieved unless the lower levels have been validated, as illustrated in Figure 2 on the left, in this case, the POI datasets would remain at a 3-star rating, losing information on schema publications and implying a priority on semantic connectivity over schema publication. However, the priority of certain aspects over others in determining data quality generally depends on the application case. Therefore, a multidimensional assessment, as illustrated in Figure 2 on the left, would be more appropriate for determining data quality in relation to the needs of the application case.

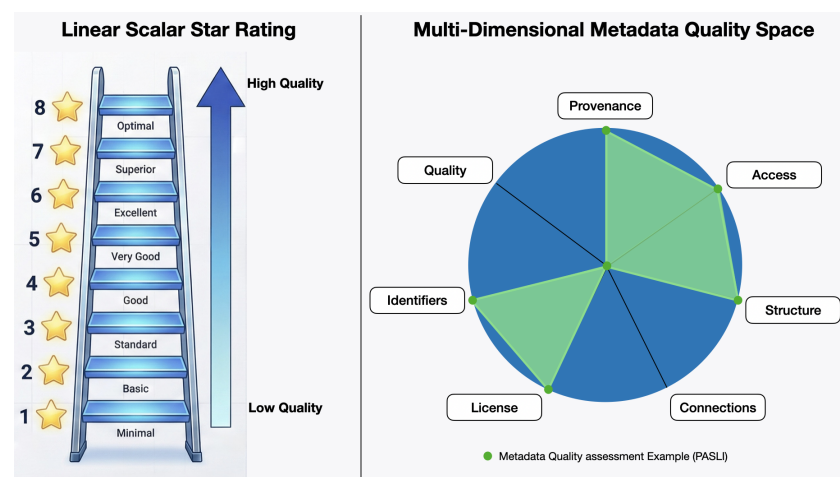


Figure 2. Conceptual contrast between scalar star ratings and multi-dimensional metadata assessment. Left: linear star progression as a scalar model. Right: binary dimension-wise assessment in the LETTER framework, where each axis is evaluated as present (1) or absent (0).

4. Proposed LETTER Taxonomy: Conceptual Analysis, Design, and Definition

4.1. Requirements Derived from Practice and Standards

To move beyond the limitations of scalar star schemes, the design of the LETTER framework was guided by four requirements derived from geospatial standards, FAIR-oriented data stewardship, and practical metadata use.

First, each LETTER had to remain conceptually separable. Each dimension should correspond to a distinct metadata function that answers a different user or machine question, rather than mixing technical openness, legal clarity, provenance, and trust into a single score.

Second, the framework had to be operationally assessable from existing metadata records. The selected dimensions therefore had to be evaluable through observable metadata evidence in standards-based catalogues, such as ISO 19115-1:2014 [28], ISO 19157:2013 [23], ISO/TS 19139:2007 [41], or DCAT [42]-aligned records.

Third, the framework had to preserve traceability to existing standards and principles. The dimensions were selected so that they could be mapped to FAIR requirements and to established geospatial metadata constructs such as identification, access, lineage, constraints, references, and quality reporting.

Fourth, the framework had to remain usable as a compact communication layer for non-specialist users while still being meaningful for automation. This excludes highly granular decompositions that may be standard-compliant but are difficult to communicate or compare at a glance.

The resulting seven dimensions were therefore not chosen merely for mnemonic convenience. They were selected because, taken together, they capture the minimum set of metadata functions required to determine whether a dataset can be parsed, retrieved, contextualized, legally reused, stably referenced, traced, and trusted through explicit quality evidence.

4.2. Solution: A Mnemonic, Multi-Dimensional Framework

4.2.1. Very Mnemonic LETTER Dimensions

We propose the LETTER taxonomy, defining seven independent dimensions. Each letter is chosen to be human-obvious and mnemonic:

- **P (Provenance):** Transparent lineage, source datasets, processing steps, and versioning.
- **A (Access):** Stability and documentation of technical access mechanisms.
- **S (Structure):** Structured, non-proprietary formats and documented schema.
- **C (Connections):** Links to other datasets, vocabularies, and registers.
- **L (License):** Clarity and machine-readability of reuse conditions.
- **I (Identifiers):** Stable dataset and feature identifiers plus coordinate reference systems.
- **Q (Quality):** Documented and accessible quality evaluation methods and results.

A dataset is described by a code containing the validated letters from among the seven, such as **PASCLIQ** for a fully validated dataset or **PASCLI** for a dataset that does not validate the quality dimension. This code is directly readable by humans and can be pronounced as a word.

4.2.2. Binary Baseline Criteria

To avoid the ambiguity of partial-credit ladder systems, the formal LETTER model uses binary baseline criteria for each dimension.

- **Binary baseline:** Each LETTER dimension is assigned a value of 1 only if all minimum machine-actionable conditions for that dimension are satisfied, and 0 otherwise.
- **No implicit compensation:** Strong performance in one dimension does not compensate for failure in another. For example, linked-data exposure does not substitute for missing provenance, and schema publication does not substitute for explicit licensing.
- **Diagnostic intention:** The purpose of this strict baseline is not to deny that metadata maturity can vary in degree but to ensure that the resulting code has a clear operational meaning. If a machine or user cannot reliably act on a metadata function, the corresponding letter is not assigned.

This means that the outcome of an evaluation is a seven-dimensional compliance vector rather than a single aggregated score. More fine-grained maturity levels may be introduced in future work as an extension for improvement planning, but they are not part of the formal baseline model proposed here.

4.2.3. Encoding and Visual Representation

The resulting code is compact (e.g., **PASCLI** means all letters are present except Quality). Visually, this can be represented using:

- **Radar Charts:** Showing the dimension-wise profile of a dataset with binary values (0/1) for each dimension, consistent with the formal baseline model (as shown in Figure 2).

- **Traffic Lights/Tick Marks:** A simple dashboard interface where each letter is indicated as present (lit/checked) or absent (unlit/unchecked) based on a binary criterion.

Because the formal model is binary, visualizations should preserve the interpretation of dimensions as present or absent. Binary dashboards, tick-based matrices, or radar plots constrained to 0 and 1 are therefore suitable representations. Continuous intermediate values may be useful for future maturity extensions, but they are not part of the present formal LETTER scoring model.

4.3. The LETTER Taxonomy Definition

Table 2 provides the formal definition of the taxonomy, linking each user-facing value statement to mandatory criteria and standard mappings.

Table 2. Formal definition of the LETTER taxonomy: dimensions, mandatory criteria, and standard mappings.

Code	User Value	Mandatory Criteria	ISO/FAIR Mapping
P	Traceability	Structured lineage, Processing steps defined	LI_Lineage; FAIR R1.2
A	Reliable retrieval	Working endpoint, Open protocol (HTTP/WFS)	CI_OnlineResource; FAIR A1
S	Parsing capability	Non-proprietary format, Schema available	MD_Format; FAIR I1, I2
C	Context	Resolvable links to external resources	MD_AssociatedResource; FAIR I3
L	Legal certainty for reuse	Explicit license URL, Machine-readable rights field	MD_LegalConstraints; FAIR R1.1
I	Referenceability	Persistent URIs (dataset/feature), Defined CRS	MD_Identifier, RS_Identifier; FAIR F1
Q	Trust	Quality report with methods/measures present	DQ_DataQuality; ISO 19157

4.4. Mapping with Other Standards

4.4.1. Mapping LETTER to FAIR and ISO

The framework aligns seamlessly with existing standards:

- **L and P** support **Reusability** (FAIR) and map to *MD_LegalConstraints* and *LI_Lineage*.
- **A and I** support **Accessibility and Findability** (FAIR), respectively, and map to identification and spatial reference classes.
- **C and S** support **Interoperability** (FAIR) by enforcing schemas and links.

4.4.2. Mapping with STAR Schemes

Table 3 shows the mapping of LETTER to STAR and illustrates the conceptual difference between the linear star model and the LETTER components.

Table 3. Comparison of Star Levels vs. Implied LETTERs.

STAR Level	Typical Conditions	Implied LETTERs
Star 1–3	Structured Format, Open License	A, S, L
Star 4	URIs for things	C, I (partial)
Star 5	Linked Data	C (strong)
Star 6–8 (Extended)	Provenance, Schema, Quality	P, S, Q (often missing in practice)

The comparison highlights that while stars 1–3 provide a solid foundation (ASL), higher stars become ambiguous. A “5-star” dataset often fails strictly defined I (identifiers at the feature level) or P (provenance), masking these gaps. The LETTER framework allows these dimensions to be evaluated independently, ensuring that “Links” is not assumed just because “Quality” is present.

5. Study Design and Empirical Evaluation

5.1. Case Selection and Analytical Scope

The empirical part of this paper is designed as an exploratory comparative case study. Rather than aiming at statistical representativeness, the objective is to test the diagnostic behavior of the LETTER framework across contrasting metadata publication situations.

The analytical sample comprises four purposively selected datasets from the GDI-DE ecosystem, which also constitute the complete set of detailed case studies reported in this paper: Municipal Points of Interest (Trier), Thuringia DEM, Administrative Units (VG250), and INSPIRE Digital Land Model.

The datasets were selected using a maximum-variation strategy. The aim was to include contrasting combinations of provider level, data type, access modality, licensing conditions, and metadata publication practices. More specifically, the selection sought variation in: (i) provider context, from municipal to national infrastructures; (ii) data structure, including vector and raster resources; (iii) access modality, including download resources and OGC service-based access; (iv) legal regime, including clearly open and more restricted reuse conditions; and (v) the extent to which provenance and quality information are published in structured, machine-actionable form rather than only in narrative documentation.

This design supports analytical comparison of how the same framework behaves under heterogeneous geospatial metadata conditions. The results should therefore be interpreted as comparative and diagnostic rather than statistically generalizable.

5.2. Assessment Protocol

Each dataset was evaluated through a structured three-step protocol.

First, metadata evidence was collected from the official catalogue records and the linked technical resources associated with each dataset. Priority was given to ISO 19115 or ISO 19139 records, DCAT descriptions, service capability documents, dataset landing pages, and linked documentation where relevant.

Second, a STAR rating was assigned using the interpretation of star-based schemes documented in this study. This step was included for comparative purposes only and focused on the type of signals that conventional star-based models communicate regarding openness, structured publication, linking, schema exposure, and related technical characteristics.

Third, a LETTER code was assigned by checking each of the seven dimensions against the formal binary criteria defined in Section 4. A dimension was assigned only when the minimum machine-actionable conditions were satisfied. For example, Q required explicit and accessible quality evidence beyond generic narrative claims, and P required structured provenance or lineage information sufficient to support traceability in a reuse workflow.

The purpose of this protocol was not to rank datasets globally but to compare how scalar star ratings and multi-dimensional LETTER codes differ in the type of diagnostic information they convey.

5.3. Consistency Control

To reduce interpretative drift, the evaluation relied on explicit dimension-wise checks derived from the mandatory criteria in Table 2. Each case was assessed using the same decision logic and the same types of publicly accessible metadata evidence.

This LETTER framework has been implemented with explicit coding rules to assess each dimension. These rules result from an inter-rater agreement analysis.

5.4. Evaluation of the Four Geospatial Case Studies

We analyzed four purposely selected datasets from the GDI-DE ecosystem in order to test the diagnostic behavior of the LETTER framework across contrasting metadata publication patterns: Municipal Points of Interest (Trier), Thuringia Digital Elevation Model (DEM), Administrative Units (VG250), and INSPIRE Digital Land Model.

- **Municipal POI (Trier):** These municipal POI datasets are released as Open Data, referenced via a DCAT License-URI, and are accessible through stable OGC Web Map (WMS) and Web Feature (WFS) services for both visualization and raw vector access. The data is provided in structured, non-proprietary formats like JSON and GML (supported by an XSD schema) and has available coordinate reference system. Features are identified by a generic integer gid and the dataset has a non-stable objective. The accompanying metadata adheres to INSPIRE standards, providing lineage and scale information; however, it is important to note that quality assurance indicators inside metadata apply strictly to the metadata validity, not the spatial data itself. These datasets achieved 6 stars in the extended model because it includes a schema. However, the LETTER code **A S L** reveals significant gaps. The data lacks persistent dataset identifier (missing **I**) and has no links to other resources (missing **C**). This “6-star” dataset is technically structured but isolated.
- **Reference Raster (Thuringia DEM):** The DEM Thuringia dataset is an Open Data resource referenced by a DCAT License-URI, available for download in the structured, non-proprietary ASCII xyz format with an accompanying XSD schema, though the download access is neither secure nor stable. The dataset utilizes standard coordinate reference systems and lacks stable identifiers, yet it maintains strong links to INSPIRE standards. Its metadata is rich in quality and provenance details, specifying height precision, compliance verification results, and tools used, alongside a complete lineage that documents the laser-based acquisition methods and dates. This dataset is rated at 5 stars in Berners-Lee’s original model [15]) due to its open license, non-proprietary format (ASCII xyz) and link to INSPIRE. It obtains a 7 stars rating in the extended model of Hyvönen et al. [37] thanks to information quality contained in metadata. Contrary to these good star ratings obtained, the LETTER code **P S C L Q** highlights that it lacks a stable and secure access (a download link may result in packet loss without notification) (missing **A**) and stable identifiers (missing **I**).
- **Administrative Units (VG250):** Administrative Units (VG250) dataset access and information are available on https://mis.bkg.bund.de/trefferanzeige?docuuiid=93a98c5c-cf03-4a95-bf0a-54001fbf3949#detail_overview (accessed on 31 March 2026). The Administrative Units (VG250) dataset is an Open Data resource referenced by a DCAT License-URI, accessible via OGC Web Feature Service (WFS) and available in Shapefile, Excel, and OGC-compliant GeoPackage formats. The WFS access is documented through standard OGC rather than direct documentation. The data is structured with a machine-readable XSD schema accessible via the DescribeFeatureType request. The 2025 version introduces a specific resource URI, unique feature identifiers, and defined coordinate systems. Although the metadata is INSPIRE-compliant and contains basic provenance details like lineage and lifetime, it excludes specific quality metrics, with detailed procedural information regarding version updates provided in separate PDF documentation. This dataset is rated at 3 stars in Berners-Lee’s original model [15]) due to its open license and non-proprietary format (GeoPackage). It obtains a 6 stars rating in the extended model of Hyvönen et al. [37] thanks to the available data schema. The LETTER code for this dataset is **A S L I**, highlighting stable identifiers (**I**) and the lack of connections (no **C**), which is hid by the 6-star rating.

- INSPIRE Digital Land Model (DLM):** INSPIRE DLM dataset access and information are available on <https://gdz.bkg.bund.de/index.php/default/inspire/inspire-verwaltungseinheiten/inspire-wfs-digital-landscape-model-1-250-000-wfs-dlm250-inspire.html> (accessed on 31 March 2026). The INSPIRE Digital Land Model is a copyrighted dataset accessible via INSPIRE Atom feeds and OGC Web Feature Services (WFS), with the latter’s capabilities documented through standard OGC protocols rather than direct user documentation. Delivered in the structured, non-proprietary GML format with an associated schema, the dataset maintains stable object identifiers and is explicitly linked to INSPIRE standards. However, the metadata provides only basic lineage information, lacking specific quality metrics or proof of conformance, while detailed provenance regarding modeling, format changes, and source files is contained exclusively in external PDF documentation. This dataset is rated at 5 stars in Berners-Lee’s original model [15]) thanks to its link to INSPIRE and a 6 stars rating in the extended model of Hyvönen et al. [37] thanks to the available data schema. Its LETTER code is **A S C I**, highlighting the lack of machine-readable license and limited usage (no **L**). It also highlights the lack of machine-readable quality and provenance (no **Q**, no **P**). The **P** dimension is missing because the transformation logic from national data is documented in a PDF format and not a machine-readable one.

Table 4 summarizes these findings, explicitly contrasting the star rating with the diagnostic nuances of the LETTER code.

Table 4. Comparison of STAR Rating vs. LETTER Code for Geospatial Datasets.

Dataset	STAR ([15]–[38])	LETTER	Key Strengths & Weaknesses
Municipal POI (Trier)	3–6	A S L	Strength: Schema present. Weakness: Isolated (No I, C).
Thuringia DEM	5–7	P S C L Q	Strength: provided Quality (Q) and Provenance (P). Weakness: Unstable access (No A), Unstable identifiers (No I).
Administrative Units (VG250)	3–6	A S L I	Strength: Stable identifiers (I). Weakness: “PDF Trap” hides Quality/Provenance (No Q, P).
INSPIRE DLM	5–6	A S C I	Strength: Persistent feature URIs (I) and connection to INSPIRE (C). Weakness: Copyright (No L), “PDF Trap” hides Quality/Provenance (No Q, P).

5.5. Cross-Case Synthesis

Across the four case studies, several patterns emerge. First, structural publication is comparatively mature: all four datasets satisfy **S**, indicating that non-proprietary structure and schema-related evidence are now common in the evaluated geospatial publication settings. Second, provenance and quality remain the least consistently satisfied dimensions. When these dimensions are missing, the reason is often not total absence of information, but its confinement to narrative or PDF-based documentation that is not directly actionable for automated workflows. Third, comparatively high star ratings do not reliably predict broader metadata readiness. Datasets that appear mature under scalar openness-oriented schemes may still lack stable identifiers, robust access conditions, explicit legal reuse conditions, or machine-actionable trust evidence. The comparative value of LETTER therefore lies less in ranking datasets than in revealing which metadata functions remain incomplete.

6. Analysis of Results

6.1. Structural Coverage and Diagnostic Power

The four-case comparison shows that the LETTER framework exposes patterns that remain difficult to see in scalar star ratings.

First, the dimensions are unevenly distributed across the cases. Structure (S) is present in all four datasets, while Provenance (P) and Quality (Q) are much less consistently satisfied. This indicates that formal publication formats have matured more rapidly than machine-actionable trust documentation.

Second, comparatively high star ratings can coexist with missing metadata functions that are critical for reliable reuse. In the present cases, a dataset may appear technically mature because it is structured, openly published, or linked to standard infrastructures, while still lacking stable identifiers, robust access conditions, explicit licensing, or structured lineage and quality evidence.

Third, the diagnostic advantage of LETTER is operational rather than merely descriptive. A code such as ASLI or ASCI immediately indicates which metadata functions remain absent and therefore where improvement efforts should be directed. In this sense, the framework provides a more actionable assessment grammar than a single scalar score.

Across the four cases reported here, the datasets satisfy on average 4.0 out of 7 dimensions. This value should not be interpreted statistically, given the purposeful case-study design, but it is useful in illustrating that partial metadata maturity is common even in datasets that are otherwise well published.

6.2. Interoperability with FAIR and ISO

The LETTER framework is designed to be computationally derived from existing standards. Table 5 presents the mapping between LETTER dimensions and specific ISO 19115/DCAT elements, demonstrating the feasibility of automated assessment.

Table 5. Mapping LETTER Dimensions to ISO and FAIR Elements for Automation.

Letter	ISO 19115 Element (XPath)	FAIR Maturity Indicator
P	<code>gmd:dataQualityInfo/*/gmd:lineage</code>	Gen2-MI-R1.2 (Provenance)
A	<code>gmd:distributionInfo/*/gmd:transferOptions</code>	Gen2-MI-A1.1 (Protocol open/free)
S	<code>gmd:distributionFormat/*/gmd:formatSpecificationCitation</code>	Gen2-MI-I2 (Knowledge representation language)
C	<code>gmd:aggregationInfo/*/gmd:aggregateDataSetName</code>	Gen2-MI-I3 (Qualified references)
L	<code>gmd:resourceConstraints/*/gmd:useLimitation</code>	Gen2-MI-R1.1 (License machine-readable)
I	<code>gmd:fileIdentifier OR gmd:dataSetURI</code>	Gen2-MI-F1 (Unique/Persistent Identifier)
Q	<code>gmd:dataQualityInfo/*/gmd:report</code>	Gen2-MI-R1.2 (Detailed provenance/quality)

This mapping indicates that computing LETTER codes via standard scripting approaches such as XSLT transformations or general-purpose XML/JSON parsers is technically feasible for ISO-compliant catalogues, without dependence on any specific commercial product or software release. Furthermore, the consistency with FAIR maturity indicators suggests that a LETTER assessment can support FAIR-oriented evaluation, while not replacing full FAIR assessment procedures.

7. Discussion

7.1. Advantages of LETTER Compared with STAR

The transition from stars to LETTERS offers four distinct advantages for the geospatial community:

1. **Clarity:** The meaning of each letter is explicit and mnemonic. “L” stands for License, directly tying the indicator to a user value (legal certainty), whereas “3 stars” requires knowledge of the specific star-scheme variant being used.
2. **Diagnostics:** By separating dimensions, the framework highlights exactly where improvements are needed. It avoids the aggregation problem where a high score in one area (e.g., Format) masks a deficit in another (e.g., Quality).
3. **Extensibility:** New dimensions can be added as additional letters (e.g., T for Timeline, where temporal metadata can be diverse such as creation date, update date and lifespan, which are among the most common. In the present baseline model, creation and update date are more seen as part of provenance information whereas lifespan as part of quality information. For dynamic products or time-critical infrastructures, a dedicated temporal dimension may be justified in future extensions. The present paper, however, focuses on a baseline set of dimensions that can be operationalized consistently across heterogeneous geospatial metadata records.) without redefining the existing hierarchy or “breaking” the ladder, solving the scalability issue inherent in 8-star models [19].
4. **Conceptual Transferability:** The present case studies demonstrate the framework on heterogeneous geospatial datasets. More broadly, the seven LETTER dimensions describe general metadata functions rather than geospatial file types alone. This suggests conceptual transferability to broader data spaces and other metadata ecosystems, provided that the operational criteria are adapted to domain-specific standards and publication practices.

The framework also contributes to fitness-for-purpose assessment, but in a bounded sense. LETTER does not claim to replace application-specific evaluation with a universal badge. Instead, it exposes whether the metadata provides the evidence needed to judge suitability for a particular task. For example, cross-dataset integration depends strongly on Connections and Identifiers, legal reuse depends on License and Provenance, and analytical confidence depends on Quality and Provenance. The framework therefore supports fitness-for-purpose screening by making explicit which metadata functions are available for decision-making.

7.2. Limitations and Threats to Validity

Despite its benefits, the framework faces adoption challenges:

- **Learning Curve:** While mnemonic, users must learn the meaning of seven distinct letters. This is cognitively more demanding than a simple “5-star” scale, effectively trading simplicity for accuracy.
- **Design Choices:** The initial thresholds for binary criteria (e.g., what constitutes a “machine-readable” quality report?) may require tuning per domain. Strict binary pass/fail criteria might discourage data providers who are “almost” compliant.
- **Evaluator Dependence:** Because several dimensions require interpretation of metadata evidence, manual assessments may be influenced by evaluator judgment, especially where provenance or quality statements are only partially structured. This risk can be reduced through explicit coding rules and inter-rater agreement analysis.

7.3. Integration into Existing Infrastructures

The LETTER code is designed to be embedded within existing infrastructure rather than requiring new systems.

For ISO-based SDIs, the code can be stored as a `gmd:keyword` or within the abstract. An automated service could crawl the CSW (Catalog Service for the Web), compute the code based on the mapping in Table 5, and inject it back into the metadata record.

For DCAT-based Open Data Portals, the code can be computed from standard DCAT fields (`dcat:license`, `dcat:accessURL`, `dct:provenance`).

Below is a concrete example of how the LETTER code ASCII would be encoded in standard metadata formats to ensure machine-readability.

The URI `https://data.example.org/dataset/1` is used here strictly as an illustrative placeholder (`example.org` is the IANA-reserved domain for documentation, RFC 6761) and is not intended to dereference to a live resource.

- ISO 19139 XML Snippet:

```
<gmd:descriptiveKeywords>
  <gmd:MD_Keywords>
    <gmd:keyword>
      <gco:CharacterString>LETTER_CODE:ASCII</gco:CharacterString>
    </gmd:keyword>
    <gmd:type>
      <gmd:MD_KeywordTypeCode codeList="http://..."
        codeListValue="theme"/>
    </gmd:type>
  </gmd:MD_Keywords>
</gmd:descriptiveKeywords>
```

- DCAT (Turtle) Snippet:

```
<https://data.example.org/dataset/1> a dcat:Dataset ;
    dct:title "Municipal Waste Points" ;
    dcat:keyword "LETTER_CODE:ASCII" ;
    # ... existing metadata ...
```

7.4. Diagnostic Power and the “PDF Trap”

One of the most important observations in the case studies is the recurrence of what we term the ‘PDF Trap’. In several datasets, relevant information on lineage, transformation logic, validation, or quality reporting exists, but only in external narrative documentation. While this may be sufficient for human interpretation, it is not directly reusable by automated FAIR-oriented workflows and therefore does not satisfy the requirement of machine-actionable trust evidence.

The LETTER framework strictly penalizes this lack of machine-readability. By stripping the Q and P letters from datasets where quality and lineage are locked in PDFs, the framework highlights a specific, actionable upgrade path: data providers must encode this information into structured *gmd:lineage* elements or PROV-O graphs [43]. This ensures that trust becomes machine-actionable, not just human-readable.

7.5. Future Work

To advance the LETTER framework from a proposal to a community standard, several steps are planned:

- Refinement: Community-driven refinement of the criteria is necessary, especially for domains in which metadata maturity is best represented through staged progression. A future extension could introduce graded maturity levels within each dimension, but such an extension should remain clearly distinct from the binary baseline proposed in this paper.
- Automation: We aim to develop open-source tools (e.g., a “LETTER-Validator” library) that automatically infer codes from ISO and DCAT endpoints.

- **Wider Evaluation:** Larger scale studies across multiple national infrastructures are needed to validate the framework on big data corpora.
- **Hackathon Experimentation:** To rigorously test the framework's usability and robustness, we propose organizing a dedicated Geospatial Metadata Hackathon. This event would invite developers, data stewards, and domain experts to apply the LETTER framework to their own datasets, develop visualization widgets (e.g., traffic lights), and stress-test the criteria against edge cases. This participatory approach will be crucial for broad adoption.

8. Conclusions and Standardization Proposal

8.1. Summary of Contributions

This study shows that star-based schemes, while useful as communication tools for openness and linked-data publication, are conceptually limited when used as general indicators of geospatial metadata quality. In practice, the monotonic interpretation of star levels can fail because technically advanced publication characteristics may coexist with missing provenance, quality, identifier, or access evidence.

In response, this paper introduced the LETTER framework as a multi-dimensional alternative based on seven independent binary dimensions: Provenance, Access, Structure, Connections, License, Identifiers, and Quality. The framework was designed as a compact but standards-traceable metadata assessment layer aligned with FAIR principles and geospatial metadata practice.

Its application to four real-world datasets from the GDI-DE ecosystem shows that LETTER provides clearer diagnostic power than scalar star ratings. Rather than collapsing heterogeneous metadata functions into a single score, it reveals which metadata capabilities are actually satisfied and which remain absent, especially where trust-related information is available only in non-machine-actionable documentation.

8.2. Proposal for Standardization and Adoption

To move from a theoretical proposal to community practice, we outline a roadmap for standardization and adoption:

1. **Publish Guideline:** The LETTER specification should be formalized and published as a community guideline, aligned with the best practices of ISO TC 211 and the Open Geospatial Consortium (OGC) [44].
2. **Tooling:** To lower the barrier to entry, reference implementations and open-source tools must be developed to automatically generate LETTER codes from existing repositories.
3. **Pilot Programs:** We encourage Spatial Data Infrastructures (SDIs), open data portals, and environmental data infrastructures to pilot these codes within their catalogues.

Crucially, the LETTER framework is positioned not as a replacement for existing standards, but as a compact quality summary layer that sits on top of rigorous ISO 19115/19157 and FAIR implementations, providing a user-friendly entry point to complex metadata.

8.3. Outlook: Toward Cross-Domain Metadata Assessment

The framework proposed here is grounded in geospatial metadata practice, but its underlying logic is not limited to geospatial resources. The seven dimensions correspond to general metadata functions that recur across many data ecosystems: structure, traceability, access, contextual linking, legal reuse conditions, stable referenceability, and explicit quality evidence.

This is why LETTER can also be applied to broader data space architectures. However, cross-domain reuse may require operational adaptation for each or certain criteria. To

transfer LETTER to another domain, it is necessary to verify that the implementation of each dimension complies with the standards, vocabularies, and publication conventions specific to that domain and to adapt it accordingly as needed.

Author Contributions: Conceptualization, Claire Ponciano and Hartmut Müller; formal analysis, Claire Ponciano and Falk Würriehausen; investigation, Markus Schaffert; writing—original draft preparation, Claire Ponciano; writing—review and editing, Jean-Jacques Ponciano. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Publicly available datasets were analyzed in this study. The data can be found here: Municipal POI dataset—https://geoportal.trier.de/trier/mod_ogc/ (accessed on 31 March 2026); Thuringia DEM dataset—<https://geoportal.thueringen.de/gdi-th/download-offene-geodaten/download-hoehendaten> (accessed on 31 March 2026); Administrative Units (VG250) dataset—https://mis.bkg.bund.de/trefferanzeige?docuuid=93a98c5c-cf03-4a95-bf0a-54001fbf3949#detail_overview (accessed on 31 March 2026); and INSPIRE DLM dataset—<https://gdz.bkg.bund.de/index.php/default/inspire/inspire-verwaltungseinheiten/inspire-wfs-digital-landscape-model-1-250-000-wfs-dlm250-inspire.html> (accessed on 31 March 2026).

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

References

1. Yang, C.; Huang, Q.; Li, Z.; Liu, K.; Hu, F. Big Data and cloud computing: Innovation opportunities and challenges. *Int. J. Digit. Earth* **2017**, *10*, 13–53. [CrossRef]
2. Li, S.; Dragicevic, S.; Castro, F.A.; Sester, M.; Winter, S.; Coltekin, A.; Pettit, C.; Jiang, B.; Haworth, J.; Stein, A.; et al. Geospatial big data handling theory and methods: A review and research challenges. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 119–133. [CrossRef]
3. European Commission. 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 (INSPIRE). *Off. J. Eur. Union* **2007**, *L 108*, 1–14.
4. Crompvoets, J.; Rajabifard, A.; van Loenen, B.; Delgado Fernández, T. *A Multi-View Framework to Assess Spatial Data Infrastructures*; Melbourne University Press: Carlton, VIC, Australia, 2008.
5. Chrisman, N.R. The role of quality information in the long-term functioning of a geographic information system. *Cartographica* **1984**, *21*, 79–87. [CrossRef]
6. Devillers, R.; Bédard, Y.; Jeansoulin, R.; Moulin, B. Towards spatial data quality information analysis tools for experts and end-users. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 261–281.
7. Müller, H.; Louwsma, M. The Role of Spatio-Temporal Information to Govern the COVID-19 Pandemic: A European Perspective. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 166. [CrossRef]
8. Nogueras-Iso, J.; Zarazaga-Soria, F.J.; Muro-Medrano, P.R. *Geographic Information Metadata for Spatial Data Infrastructures: Resources, Interoperability and Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2005. [CrossRef]
9. Goodchild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [CrossRef]
10. Haklay, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environ. Plan. B* **2010**, *37*, 682–703. [CrossRef]
11. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.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. [CrossRef] [PubMed]
12. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
13. Mons, B.; Neylon, C.; Velterop, J.; Dumontier, M.; da Silva Santos, L.O.B.; Wilkinson, M.D. Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles. *Inf. Serv. Use* **2017**, *37*, 49–56. [CrossRef]
14. Hodson, S.; Jones, S.; Collins, S.; Genova, F.; Harrower, N.; Laaksonen, L.; Mietchen, D.; Petrauskaitė, R.; Wittenburg, P. *Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data*; Technical Report; Publications Office of the European Union: Luxembourg, 2018. [CrossRef]
15. Berners-Lee, T. Linked Data—Design Issues. 2006. Available online: <https://www.w3.org/DesignIssues/LinkedData.html> (accessed on 31 March 2026).
16. Bizer, C.; Heath, T.; Berners-Lee, T. Linked Data—The Story So Far. *Int. J. Semant. Web Inf. Syst.* **2009**, *5*, 1–22. [CrossRef]

17. Al-khatib, B.; Ali, A.A. Linked data: A framework for publishing five-star open government data. *Int. J. Inf. Technol. Comput. Sci.* **2021**, *13*, 1–15.
18. Janowicz, K.; Hitzler, P.; Adams, B.; Kolas, D.; Vardeman, C. Five Stars of Linked Data Vocabulary Use. *Semant. Web* **2014**, *5*, 173–176. [[CrossRef](#)]
19. Debattista, J.; Auer, S.; Lange, C. Luzzu—A methodology and framework for linked data quality assessment. *ACM J. Data Inf. Qual.* **2016**, *8*, 4. [[CrossRef](#)]
20. Moreau, L.; Clifford, B.; Freire, J.; Futrelle, J.; Gil, Y.; Groth, P.; Kwasnikowska, N.; Miles, S.; Missier, P.; Myers, J.; et al. The Open Provenance Model core specification (v1.1). *Future Gener. Comput. Syst.* **2011**, *27*, 743–756. [[CrossRef](#)]
21. Kubler, S.; Robert, J.; Neumaier, S.; Umbrich, J.; Le Traon, Y. Comparison of open government data quality frameworks. *Gov. Inf. Q.* **2018**, *35*, 13–29. [[CrossRef](#)]
22. Wang, R.Y.; Strong, D.M. Beyond Accuracy: What Data Quality Means to Data Consumers. *J. Manag. Inf. Syst.* **1996**, *12*, 5–33. [[CrossRef](#)]
23. ISO 19157:2013; Geographic Information—Data Quality. Technical Report. International Organization for Standardization: Geneva, Switzerland, 2013.
24. González-Beltrán, A.; Sansone, S.A. The role of metadata in the FAIR principles. *Data Intell.* **2019**, *1*, 1–16. [[CrossRef](#)]
25. Thompson, P.T.; Ojha, S.; Powell, C.D.; Pennell, K.G.; Moseley, H.N. A proposed FAIR approach for disseminating geospatial information system maps. *Sci. Data* **2023**, *10*, 389. [[CrossRef](#)] [[PubMed](#)]
26. Degbelo, A. FAIR geovisualizations: Definitions, challenges, and the road ahead. *Int. J. Geogr. Inf. Sci.* **2022**, *36*, 1059–1099. [[CrossRef](#)]
27. Jacobsen, A.; de Miranda Azevedo, R.; Juty, N.; Batista, D.; Coles, S.; Cornet, R.; Courtot, M.; Crosas, M.; Dumontier, M.; Evelo, C.T.; et al. FAIR Principles: Interpretations and Implementation Considerations. *Data Intell.* **2020**, *2*, 10–29. [[CrossRef](#)]
28. ISO 19115-1:2014; Geographic Information—Metadata—Part 1: Fundamentals. Technical Report. International Organization for Standardization: Geneva, Switzerland, 2014.
29. ISO 19165-1:2018; Geographic Information—Preservation of Digital Data and Metadata—Part 1: Fundamentals. Technical Report. International Organization for Standardization: Geneva, Switzerland, 2018.
30. Olfat, H.; Kalantari, M.; Rajabifard, A.; Williamson, I.P. Spatial Metadata Automation: A Review of the State of the Art. In Proceedings of the GSDI 13 World Conference, Quebec City, QC, Canada, 14–17 May 2012; pp. 1–14.
31. Lush, V.; Bastin, L.; Lumsden, J. Geospatial data quality indicators. In Proceedings of the 10th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (Accuracy 2012), Florianopolis, Brazil, 10–13 July 2012; pp. 367–372. Available online: <https://www.spatial-accuracy.org/system/files/Lush2012accuracy.pdf> (accessed on 31 March 2026).
32. Benitez-Paez, F.; Degbelo, A.; Trilles, S.; Huerta, J. Contribution to the analysis of the usability of Open Government Data. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 474.
33. Degbelo, A.; Granell, C.; Trilles, S.; Bhattacharya, D.; Casteleyn, S.; Kray, C. Opening up smart cities: Citizen-centric challenges and opportunities for GIScience. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 16. [[CrossRef](#)]
34. Ahonen-Rainio, P. Visualization of Geospatial Metadata for Selecting Geographic Datasets. Ph.D. Thesis, Helsinki University of Technology, Espoo, Finland, 2005.
35. Map Library Standards Committee. *Best Practices for Geospatial Data Management*; Map Library: London, UK, 2019.
36. Martin, S.; Foulonneau, M.; Turki, S. 1–5 stars: Metadata on the openness level of open data sets in Europe. In *Proceedings of the Research Conference on Metadata and Semantic Research*; Springer: Cham, Switzerland, 2013; pp. 234–245.
37. Hyvönen, E.; Tuominen, J.; Alonen, M.; Mäkelä, E. Linked Data Finland: A 7-star model and platform for publishing and re-using linked datasets. In *Proceedings of the The Semantic Web: ESWC 2014 Satellite Events*; Springer: Cham, Switzerland, 2014; pp. 226–230.
38. Hyvönen, E.; Tuominen, J. 8-star linked open data model: Extending the 5-star model for better reuse, quality, and trust of data. In Proceedings of the Posters, Demos, Workshops, and Tutorials of the 20th International Conference on Semantic Systems (SEMANTiCS 2024), Amsterdam, The Netherlands, 17–19 September 2024; CEUR Workshop Proceedings; Volume 3759. Available online: <https://ceur-ws.org/Vol-3759/paper4.pdf> (accessed on 31 March 2026).
39. Wilson, J.P.; Butler, K.; Gao, S.; Hu, Y.; Li, W.; Wright, D.J. A five-star guide for achieving replicability and reproducibility when working with GIS software and algorithms. *Ann. Am. Assoc. Geogr.* **2021**, *111*, 1311–1317. [[CrossRef](#)]
40. Hasnain, A.; Rebholz-Schuhmann, D. Assessing FAIR data principles against the 5-star open data principles. In *Proceedings of the Semantic Web: ESWC 2018 Satellite Events*; Springer: Cham, Switzerland, 2018; pp. 469–477.
41. ISO/TS 19139:2007; Geographic Information—XML Schema Implementation. Technical Report. International Organization for Standardization: Geneva, Switzerland, 2007.
42. World Wide Web Consortium (W3C). Data Catalog Vocabulary (DCAT)—Version 2. W3C Recommendation. 2020. Available online: <https://www.w3.org/TR/vocab-dcat-2/> (accessed on 31 March 2026).

43. World Wide Web Consortium (W3C). PROV-O: The PROV Ontology. W3C Recommendation. 2013. Available online: <https://www.w3.org/TR/prov-o/> (accessed on 31 March 2026).
44. Federal Geographic Data Committee. *National Spatial Data Infrastructure Strategic Plan 2021–2024*; Technical Report; FGDC: Reston, VA, USA, 2020.

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