Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

JOURNAL OF INFORMATION SYSTEM AND TECHNOLOGY MANAGEMENT (JISTM)
www.gaexcellence.com/jistm

POLICY TRANSLATION IN DATA MANAGEMENT: COMPARATIVE ANALYSIS OF GLOBAL STANDARDS AND LOCAL ADAPTATIONS

Alahdal Anas Hashim. A[1]*, Noor Zaidi Sahid[2], Ahmad Nadzri Mohamad[3]

[1]Faculty of Information Science, Universiti Teknologi MARA, UiTM Puncak Alam Campus, Selangor, Malaysia
alahdalanas@gmail.com https://orcid.org/0009-0009-6148-281x
[2]Faculty of Information Science, Universiti Teknologi MARA, UiTM Puncak Alam Campus, Selangor, Malaysia
Zaidisahid@uitm.edu.my https://orcid.org/0000-0003-1528-1514
[3]Faculty of Information Science, Universiti Teknologi MARA, UiTM Puncak Alam Campus, Selangor, Malaysia
Nadzri590@uitm.edu.my https://orcid.org/0000-0002-2690-4564
*Corresponding Author

Article Info:

Article history:
Received date: 30.04.2026
Revised date: 21.05.2026
Accepted date: 25.06.2026
Published date: 30.06.2026

To cite this document:

Hashim. A, A. A., Sahid, N. Z., & Mohamad, A. N. (2026). Policy Translation in Data Management: Comparative Analysis of Global Standards and Local Adaptations. *Journal of Information System and Technology Management, 11* (43), 242-260.

Abstract:

International open data frameworks rely heavily on creating a unified metadata protocol for seamless national/institutional repository interoperability. Nonetheless, their top-down global applicability is regularly broken down by the reality of multiple local administrative scenarios. This paper examines how localized practices are translated from shared metadata rules and rigorously compares data catalog deployments. Taking the Data Catalog Vocabulary Application Profile (DCAT-AP) that is part of the European open data infrastructure as our reference model, we analyzed about 450,000 metadata records derived from 84 different regional and national catalogs. Our empirical study reveals three main patterns of operation, namely 'strict compliance', 'contextual adaptation', or 'systemic workaround'. The data shows significant regional differences: Western catalogs rely heavily on structural workarounds (40%), Northern European hubs achieve a 45% compliance rate, while Eastern catalogues opt for structural workarounds with 15%. Overall, this research suggests that efforts to achieve true cross-border data interoperability should be more context-oriented and nuanced – not one-size-fits-all or compliance-obsessed.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

DOI: 10.35631/JISTM.1143014 **Keyword:**
Administrative Science, Data Governance, Information Management, Metadata Standards, Policy Translation

Introduction

The value of a shared metadata standard in a modern world of public administration and scientific research is undeniable, as it provides an indispensable foundation for the discoverability of data, systematic re-use and seamless cross-border alliances (Gujar, 2025; Sathvika & Jeyalakshmi, 2026). In the European context, this urgency is especially evident, as European law and strategies are increasingly focusing on the strategic shift towards a single digital European market, which has inspired the development of common European data spaces supported by robust legislative measures, such as the Open Data Directive and the Data Governance Act (DGA) (Niederhauser, 2025; Derycke et al., 2025).

A common, standardised mechanism was needed to connect these aspirations, expressed in statute, to the real, technical, operational infrastructure. To that end, the European Commission formalized as an application profile the Data Catalog Vocabulary (DCAT-AP), which will serve as the standardization of the binding protocol that will be used in the majority of national and regional portals to harvest and consolidate data in the European Data Portal (Conde et al., 2024; Nogueras-Iso et al., 2021). Although initially developed to meet the structural needs of the European public sector, DCAT-AP is much more than a regional or European standard: it is a standard applicable all over the world and a benchmark example for metadata interoperability. Developed in absolute alignment with the W3C Data Catalog Vocabulary (DCAT), the definitive worldwide standard definition of web-based resources descriptions DCAT-AP is completely flexible, is not linked to any technological environment, and is independent of any technology (W3C, 2025; Nogueras-Iso et al., 2021).

This inherent utility is further enhanced by its built-in semantic (content-level) connections to international scholarly databases and the broader metadata world. A key illustration is the CiteDCAT-AP profile, which was designed to enable direct, seamless mapping to the international DataCite metadata schema, thus simplifying the citation of datasets around the world (Conde et al., 2024). Consequently, the European DCAT-AP framework provides a valuable, globally applicable laboratory for developing findings for the complex dynamics of abstract international standardization policies being made tangible in very heterogeneous institutional settings, which are highly transferable.

Nearly always, there are no simple, linear steps to help reach compliance with regulatory or technical requirements. In contrast, the practical implementation of metadata protocols is spread across extremely disjointed administrative regions, distinguished by often markedly asymmetric technical skill, legal landscape, and entrenched legacy technologies (Conde et al.,

2024; Nogueras-Iso et al., 2021). This structural gap leads to a complex process that has come to be known as "policy translation". In this paradigm, formal administrative requirements can become an invitation to renegotiate, adapt, or simply find ways of sidestepping them to better suit the convenience of local interests responding to local issues of scarcity of resources and incentive for institutions (Arisi, 2022; Broomfield, 2023).

Scholarly assessments of Open Data repositories have long been reductionistic: When they deviate from the standard, they viewed them as some kind of technical problem or as a sign of poor quality. This dichotomic conceptualization of 'compliance/non-compliance' is insufficiently explanatory, however, to decript and explain why particular structural anomalies still persist or how general bureaucratic design scandalizes semantic integrity (Martinez-Gil, 2025). Although ample technical literature has unerringly listed the changing validity scores, this literature rarely correlates with the macro-level institutional setting that produces and first employs the data products, as a matter of fact, an empirical instance of metadata has not been linked to its macro-level institutional origins (Maratsi et al., 2024).

To overcome these analytical limitations, the present study proposes a fully-fledged policy translation framework that allows for an assessment of the standardisation of data management based on the perspective of multi-level governance systems. This research is an extensive analysis of 450,000 metadata records within an empirical base that includes 84 European national and subnational catalogues gathered from the European Data Portal. Portals are classified into three operational modes: compliance, adaptation, and workaround, by performing a quantitative audit of metadata completeness, and aligning vocabulary. This investigation applies the methodology to systematically answer three overall research questions:

1. How are global metadata specifications implemented, adapted, or transformed in local administrative settings?
2. What are the concrete pressures put on local actors by particular institutions and what is it about particular organisational typologies that make local actors feel a pull towards adopting adaptive or workaround strategies?
3. How do these diverse implementation practices impact in general on semantic consistency and discoverability of the data across portals in the European data space?

Literature Review

The literature on open data cataloging and metadata interoperability has also changed a great deal in line with the far-reaching digital transformation of the institutional infrastructure (Abd Halim & Abdul Rahman, 2024; Mhlongo et al., 2023). The major questions asked in this area were the questions of building foundational semantic frameworks for inter-translation between various systems. For example, Nogueras-Iso et al. (2021) identified some metadata quality evaluation metrics according to ISO standards and noticed that other metadata characteristics such as semantic consistency and structural integrity are often sacrificed during the automatic harvesting process directly as a result of localized portal idiosyncratic modeling decisions. Similarly, Martinez-Gil (2025) applied automated tracking measures to large volumes of data and found a longstanding gap between strong legal compliance and very deficient structural documentation.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

American data repositories continue to fragment, and more recent empirical investigation of the European data repositories confirms this ongoing structural fragmentation. Maratsi et al. (2024) in their semantic audit of 27 different European open data portals came to the conclusion that fewer than 60% of all portals achieve reliable population of mandatory elements according to DCAT-AP and the use of controlled vocabularies was limited to less than 35% of the evaluated datasets. In an effort to address this gap, Conde et al. (2024) added algorithmic layers of metadata validation and transformation, with the intervention improving the gross completeness scores, but also highlighting the limitations of automated remediation in fully restoring lost contextual nuances.

The emergence of domain-specific architectural extensions, instead of generic ones in the form of profiles, has added further complexity to the whole ecosystem. A notable example was discussed by Derycke et al. (2025) on the application of HealthDCAT-AP in the European Health Data Space (EHDS). These special extensions are beneficial for greater localization of searchability but cause strong heterogeneity of profiles, which makes the integration and aggregation of data between the different domains more difficult (Maratsi et al., 2024).

Turning attention to changes in paradigm, Hrynaszkiewicz et al. (2020) highlighted low levels of open access compliance and the significant role played by the inhibiting administrative burden. Demonstrated in this way, monitoring dashboards could be thought of as processes of continuous policy mediation between the hard law of global standards and local contexts. They function as dynamic actors of policy translation, constantly balancing the institutional pressure of global standards and local lived realities, as Salamoura & Tsakonas (2024) suggested. Therefore, Dortmund (2025) proposed that the metadata frameworks have to undergo an evolutionary leap to become very adaptable and contextually aware and suggested that this requires resorting to relationship-oriented linked data models.

To adequately interpret these dynamics, this study adopts a theoretical approach rooted in the policy translation theory that has been developed in the social policy area (Lendvai & Stubbs, 2007; Clarke et al., 2015). In contrast to the notion of "policy transfer" (which assumes a sterile and linear adoption process), "translation" points to the active re-fashioning, re-interpretation and construction of new meanings that inevitably take place during the shift from one administrative jurisdiction to another (Lendvai, 2015; Mukhtarov, 2014). This way of thinking is also very relevant to recent empirical evidence that shows that when stakeholders are 'morally' motivated by internal community norms, they tend to outpace baseline regulations and requirements, over and above top-down forms of legislative coercion (Taylor-Grant, 2024). Adding to this context of administrative fragmentation, Mohd Taib et al. (2025) studied the factors that negatively drive the commitment for good governance disclosure as transparency is often tainted as a mere box-ticking exercise, which can be attributed to strong compliance regimes and prevalent gaps in managerial-level strategic awareness. Under the same institutional constraints, Md Ajis et al. (2024) and Jafalizan Md Jali et al. (2025) shed light to the escalating uncontrolled and untapped data that is normally referred to as "dark data" in organizational silos. These unstructured data assets, they said, pose serious issues to regulatory compliance — as long as the stewards of data don't actively preventively implement strong upskilling, control and security countermeasures. While addressing these fragmented data contexts, nobody questioned Hussin et al. (2025) when they commanded that we must standardise the metadata architectures comprehensively; if we do not lift up the contextual and structural level of consistency, we won't be able to maintain strong information governance.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

Table 1: Literature Synthesis and Methodological Comparison of Prior Data Governance Studies

Reference	Scope/Dataset	Methodology	Identified Limitation	Core Contribution of This Work
Conde et al. (2024)	EU portals via CKAN	Automated metadata validation	Focused on metadata layer only	Links completeness shifts to translation modes
Maratsi et al. (2024)	27 European OGD portals	Semantic schema analysis	Limited to metadata semantics	Connects semantic divergence to governance models
Nogueras-Iso et al. (2021)	RDF-based metadata	ISO-aligned metrics	Sensitive to modelling choices	Embeds variation within policy translation model
Salamoura & Tsakonas (2024)	European dashboards	Comparative monitoring	Focused on OA dashboards	Translates dashboard challenges to governance models
Dortmund (2025)	Library metadata	Conceptual future-proofing	Lacks large-scale validation	Operationalizes flexibility via formal framework
Azmi et al. (2025)	Malaysian Public Sector	AI Governance Review	Focused on ethics/privacy	Bridges local policy needs with global data standards

Methodology

Data management is used as a test case of how policies are translated, and the following section introduces a formalised analytical framework, which is specifically designed to focus on the translation of policies in the context of data management. This approach allows for a more holistic evaluative lens since it connects the dots between the broad spectrum of global challenge and institutional local realities and constraints, and the institutional realities and consequences of operating in various contexts (Kamarulzaman et al., 2025).

Mathematical Formalization

As a starting point for the empirical evaluation, the European metadata standard de facto, the core DCAT-AP specification is fixed as a concise list of normative requirements:

$$S = \{s1, s2, \ldots, sn\}$$

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

In this construction, each of these elements si can represent a parameter that either can be either formal rules, cardinality thresholds or controlled vocabulary (Gujar, 2025). This general guideline is certainly evaluated in terms of a particular institutional environment which we model mathematically by a contextual vector:

$$E = \langle Cap, Leg, Res, Gov \rangle$$

In this case, the variables are defined as: Cap, which measures the basic organizational capability; Leg, which measures the overarching legal constraint; and Res, which measures the fiscal and technical resources available; Gov, which is a binary measure of structural governance, where 0 represents a strict central hierarchical regime and 1 represents a highly decentralized, or federated, regime.

The actual translation mechanism is thought of as a function T that maps the n idealized standard S to the m empirical information and phenomena of observed data management practices P, with the environmental vector E looming in the process:

$$T : S, E \rightarrow P$$

The mapping function encapsulates the non-linear transformation. It does not suggest an unmediated and transparent transposition of standards. Instead, it assumes the abstraction defined by the global standard S is transformed into the practice set P under the interpretative filter of institutional context vector E. In effect, differences in practice sets P are taken to be an organized response to local administrative capability and legal constraints.

$$O = \{oc, oa, ow\}$$

They are represented by oc (strict compliance), oa (structural adaptation) and ow (workarounds) respectively in this set.

Quantitative Metric and Semantic Deviation Index

Then, with the help of two simple algorithms two basic scoring systems are applied to quantify the observed practices P into benchmark:

• Metadata Completeness Score C(d): This is an independent measure of each dataset and represents a score on the density of populated fields:

$$C(d) = \sum (wj * xj) / \sum wj$$

The weights wj are given via a hierarchy (from 3 for mandatory fields to 1 for optional fields) and the binary indicator xj $\in$ {0, 1} are used to note whether a specific property j has valid, non-empty data or not (Conde et al., 2024).

• Controlled Vocabulary Compliance V(d): This measures the semantic similarity to controlled concept vocabularies:

$$V(d) = \sum yj / m$$

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

This will be considered a legitimate entry from a formal controlled vocabulary registry (e.g., IANA Media Types and/or EuroVoc), where yj = 0 will mean a non-compliant or a free-text input (Nogueras-Iso et al., 2021).

Deriving on this basic set of metrics, we empirically define the outcome classification by use of a custom Semantic Deviation Index (SDI). Within the SDI, for every catalog k exactly the amount of architectural deviation from the normative base can be measured:

$$\boldsymbol{SDIk = |S \setminus Pk| / |S|}$$

The Semantic Deviation Index (SDIk) attempts to quantify the structural distance between the required (normative) features of the standard S and the observed verified practices of the catalog k (Pk). By utilizing the set difference cardinality |S \ Pk|, this index highlights the exact fraction of the recommended and obligatory metadata properties, which are missing and are not met at the given repository. The index is normalized with respect to the entire set size of the standard, S. SDIk = 0 if S = Pk and SDIk ≈ 1 if only fundamental semantics is not met.

Decision Framework and Operational Rules

Rubric of tri-partite mode decision; catalogues are systemically classified with the use of the computed index, into our defined operational modes as follows:

- Compliance oc: Assigned if SDIk ≤ 0.10
- Adaptation oa: Assigned if 0.10 < SDIk ≤ 0.40
- Workaround ow: Assigned if SDIk > 0.40

The specific delineations were not chosen randomly but were deduced through an iterative process resulting from a thorough comparison across catalogues and validated through critical analysis with existing conformance baselines in the technical literature that has been created previously (Conde et al., 2024; Nogueras-Iso et al., 2021).

Role of Mathematical Formalisms

Such mathematical formalisms as described previously are not just abstract but must describe the concrete architecture through which our empirical analysis is realised. The first set of equations are a direct representation of the normative baseline, which performed the integrity of metadata review. Moreover, the Semantic Deviation Index is the algebraic derivation of this study, which will be the critical point of this study. It rigorously prescribes the quantitative scoring mechanism and confirms the concrete quantifiable limits necessary to categorise a wide variety of web portals into the various behavioral domains either 'full compliance' or 'measured adaptation' or 'systematic workarounds'.

Qualitative Translation Outcomes

A conscious aim to guarantee absolute methodological transparency and the possibility of subsequent reproduction led to explicit definition of the qualitative translation outcomes according to the following operational parameters:

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

- Strict Compliance oc: A repository receives this classification when it has robust average values of its metrics with C > 0.85, V > 0.70, which also equates to an SDIk < 0.10. Upwards of 90% of the derived catalogued datasets should contain a structurally sound publisher IRI (no blank nodes) that has at least one fully resolvable distribution access URL. Particularly, mandatory fields which are arbitrarily filled with such strings as N/A or None must not be used.
- Contextual Adaptation oa: Schemas in the intermediary statistical bandwidth 0.10 < SDIk ≤ 0.40 exhibit intentional schema augmentation; that is, schema extensions that are designed for a specific purpose and are typically, such as for GeoDCAT-AP or HealthDCAT-AP, specialized structures. They can also use local semantic vocabularies which are explicitly associated to the more general European taxonomies through skos:exactMatch assertions. Under this tier overt workaround strategies, if used, should be minor, less than 5% of the overall repository records.
- Systemic Workaround ow: When a Portal falls below a baseline level for the holistic completeness, or for adherence to the portal vocabulary (SDIk > 0.40). Such catalogs are routinely either deployed with pervasive (across more than 5 percent of required fields) placement of "placeholder" data or have an extreme reliance on "blank" nodes to hide the identity of the publisher (more than 10 percent of required fields) or enact routine access URL spoofer based matching to artificially high automated validation scores.

The latter was provided by two domain-expert researchers who completed an independent, double-blinded coding process on a random subset of 10 of the catalogs (1,000 individual metadata records) to ensure integrity of these classifications. With this inter-rater reliability, a highly favourable Cohen's kappa coefficient ($\kappa = 0.89$) was obtained, which clearly led to the conclusion that the coding framework was robust (Kamarulzaman et al., 2025).

Experimental Setup and Automated Queries

The theoretical framework was then empirically based by creating an experimental set of lexical resources through harvesting of an extensive data set of 450,000 metadata records. To systematically extract these specific digital collections, 84 different repositories were aggregated, located all over Europe and part of the U.S., that are heavily embedded in the European Data Portal (EDP) infrastructure, based on a pre-defined extraction window from January 2025.

Analytical Workflow – SDI Computation

In order to measure the amount of structural drift, the Semantic Deviation Index (SDI) was measured for the whole of the 84 sampled catalogues. This was achieved by systematically capturing the missing or mis-structured DCAT-AP attributes that directly address both the mandatory and highly recommended properties with regard to the definitive and normative baseline (Set S). The evaluative matrix explored five syntactic and semantic aspects of architectural criteria: the validity of resolution of the publisher IRI, the declaration of the license frameworks, the resoluteness of the titles of datasets, the strictness of use of known controlled vocabularies, and the usability of the distribution access URL. Finally, the overall SDI, created by combining the multi-dimensional SDI, was used to determine which of the three stated categories compliance, adaptive or workaround was applicable to each portal described above in our decision framework.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

SPARQL-Interrogated Queries

The goal of this section is to look at how to create an automated pipeline for running SPARQL queries on the triplestore infrastructure (Virtuoso 7.2). This interrogation produced two phases which identified mapping publisher identities and general completeness:

```
SELECT ?dataset ?title ?publisher (COUNT(?dist) AS ?distCount) WHERE {
  ?dataset a dcat:Dataset ;
           dct:title ?title ;
           dct:publisher ?publisher .
  OPTIONAL { ?dataset dcat:distribution ?dist }
}
GROUP BY ?dataset ?title ?publisher
```

The baseline was then extracted, and a secondary and more targeted query was created to isolate outlying structural patterns. It was this specific query which was looking for missing distribution layers and publishing of identities behind untraceable blank nodes:

```
SELECT ?dataset WHERE {
  ?dataset a dcat:Dataset .
  OPTIONAL { ?dataset dcat:distribution ?dist }
  OPTIONAL { ?dataset dct:publisher ?pub .
             FILTER(isBlank(?pub)) }
  FILTER(!BOUND(?dist) || BOUND(?pub))
}
```

Boundary Enforcement – SHACL Validation

Just inspection of the triple store is not enough, but one must also have strong topological validation exactly at the edge of the harvesting boundary. On this point we created a large, comprehensive test suite using the W3C Shapes Constraint Language (SHACL). For the particular shape constraint below the intent is to explicitly exclude any anonymity from the publisher fields, for each harvested dataset those URIs are permanently linked to some permanently identified and non-blank publisher URI (W3C, 2025):

```
@prefix sh : <http://www.w3.org/ns/shacl#>.
@prefix dcat : <http://www.w3.org/ns/dcat#>.
@prefix dct : <http://purl.org/dc/terms/>.
@prefix xsd : <http://www.w3.org/2001/XMLSchema#>.

dcat:DatasetComplianceShape
  a sh:NodeShape ;
  sh:targetClass dcat:Dataset ;
  sh:property [
    sh:path dct:publisher ;
    sh:minCount 1 ;
    sh:nodeKind sh:IRI ;
  ] .
```

Results and Analysis

This section presents the tables, charts, and graphs of descriptive statistics and regional trends. Extending the SDI values in an empirical mapping reveals stark geographical differences in the digital landscape on the European scale, as Table 2 shows.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

The Northern European cluster exhibited an outstanding technical fidelity with an aggregate mean SDI of only 0.08, thus securely placing the overwhelming majority of these repositories in the strict compliance category. The opposite of this was found when portal tasks occurred within Western Europe, for which the mean SDI was 0.24, a strong SDI indicator that is strongly related to deliberate, and contextual adaptation mode. Regions with more significant structural differences were in the administration corridors in the Southern (mean SDI 0.31) and in the Eastern Europe (mean SDI 0.43). This is a numerical decline thanks to the many unsupported and unidentifiable black boxes created at publisher origins, but especially because of the lack of controlled vocabularies in the system.

Table 2: Metadata Quality Metrics at The European Level

Region	Catalog Count	Analysed Records	Mean SDI	Compliance (oc)	Adaptation (oa)	Workaround (ow)
Western Europe	22	145,000	0.24	25%	60%	15%
Northern Europe	18	110,000	0.08	45%	48%	7%
Southern Europe	24	125,000	0.31	15%	55%	30%
Eastern Europe	20	70,000	0.43	10%	50%	40%

An example of the raw SPARQL output (first 3 rows of the Western Europe cluster, catalogue "data.de") is shown in Table 3:

Table 3: Raw SPARQL Output Example

ID	Title	Publisher IRI	Dist Count	Blank Node?	Placeholder?
d1	Population 2024	http://id.de/stat	3	false	false
d2	GDP regions	http://id.de/stat	2	false	false
d3	Air quality	_:bnode123	1	true	false

The average completeness score was calculated based on the limited subset of words given this context C = 0.86 and also the vocabulary compliance score V = 0.61. This particular catalog is at the Adaptation tier due to predominantly deviations in controlled vocabulary integration as its blank node saturation rate is 4% and its prevalence of textual placeholders is 0%. To clarify that see Figure 1.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

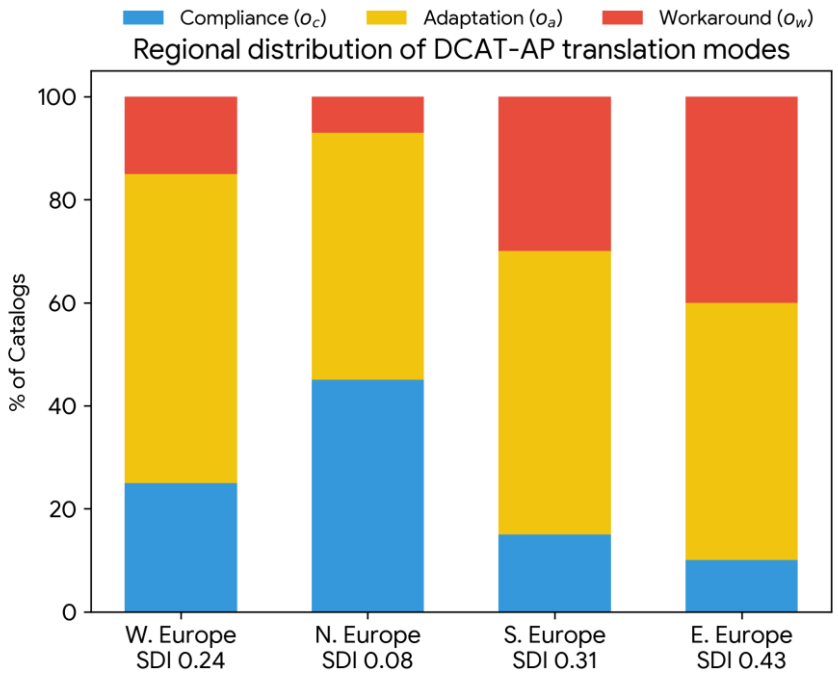

Figure 1: Regional Distribution Of DCAT-AP Translation Modes

Cumulative percentage by sub-region in Europe is illustrated in Figure 1, with mean SDI values indicated under each regional label. Northern Europe demonstrates the largest compliance share (45%) and the smallest mean SDI value (0.08), whereas Eastern Europe has the largest proportion of workarounds (40%) and the largest mean SDI value (0.43).

Quantitative Prevalence of Translation Outcomes

Moving to the macro level ecosystem, Table 4 shows the overall translation results throughout the whole spectrum of the 84 portals analyzed. Table 4 gives a quantitative distribution and threshold analysis of the outcomes.

Table 4: Quantitative Distribution and Threshold Analysis of Translation Outcomes

Translation Mode	SDI Threshold	Portals (N)	% of Sample	Primary Institutional Driver
Compliance oc	SDIk ≤ 0.10	21	25%	Centralized mandates, high capacity
Adaptation oa	0.10 < SDIk ≤ 0.40	45	54%	Domain extensions, local customization

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

Workaround ow	SDIk > 0.40	18	21%	Complex local laws, resource scarcity

Contrary to what many could guess, Compliance SDI is not the most adopted way. In fact the main trend, at 54% (n=45), is Contextual Adaptation. Today, there is no doubt – Operative Reality of Adaptation is 54% of the analysed infrastructural network. The actual movement of empowerment and homogenization breaks down the basic premises of highly centralist agencizing of policy, namely that in practice, all policies are run using the same, conflict-free method.

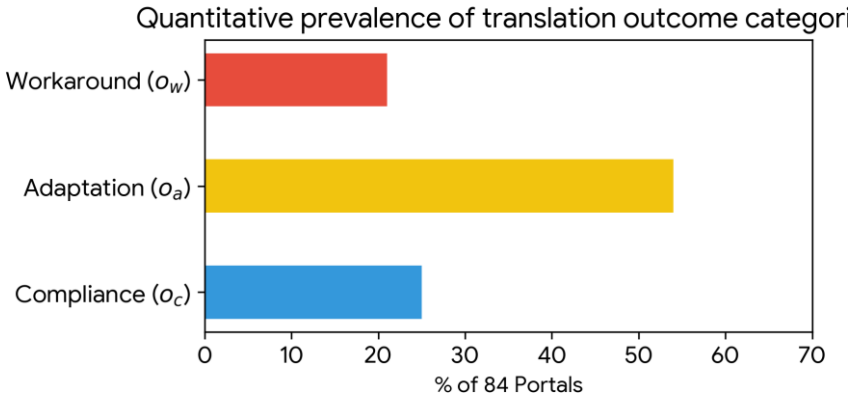

Figure 2: Quantitative Prevalence of Translation Outcome Categories

Technical Verification of Waypoints

The supporting visual data models are mathematically defined in terms of the underlying structural contexts for the analysis in order to guarantee full visual and structural interpretability among these models. Regional Trends scale spatially the variance of SDI. Workaround Mechanisms plot implementation outcome categories against the percentage weight of total catalogues, which visualizes the results of the direct queries on the catalogues on the SPARQL outputs with regard to structural evasions. Indicator Strengths assess Metadata Quality Indicators in relation to combined aggregate Performance Percentages, with strengths directly derived from precedent empirical research (Conde et al., 2024; Maratsi et al., 2024), as shown in Figure 3.

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

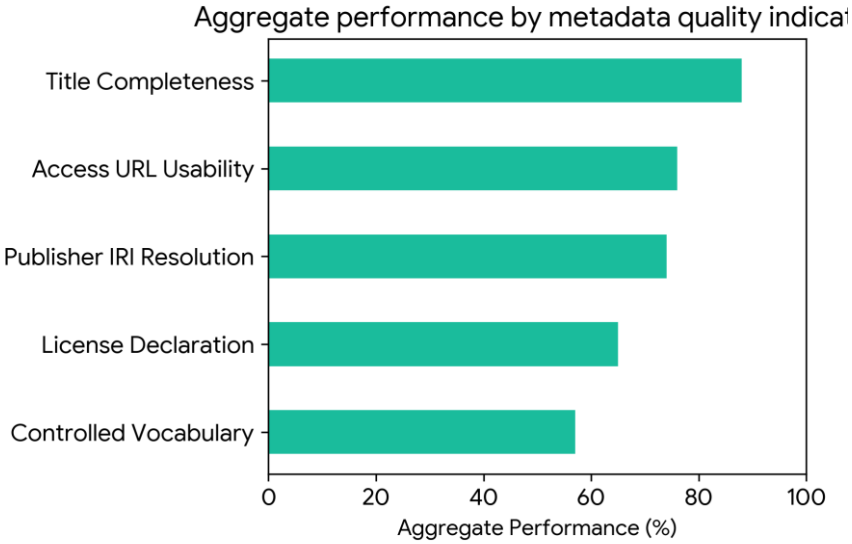

Figure 3: Aggregate Performance by Metadata Quality Indicator

The compliance for the 5 dimensions of structural evaluation (DCAT-AP) over all 84 portals is provided in this plot. The weakest dimension is "Controlled Vocabulary Usage" with a score of 57%. This result matches with the research from Maratsi et al. (2024) where compliance was less than 35% and with V = 0.61 computed using Eq. V(d). "Dataset Title Completeness" receives the highest score with 88%.

Correlation Analysis between Institutional Capacity (0-1) and Compliance Rates reveals an important positive statistical relationship ($R2 = 0.73$; $p < 0.001$), demonstrating that technical standardization comes from administrative pressure, rather than legislative enforcement (Figure 4).

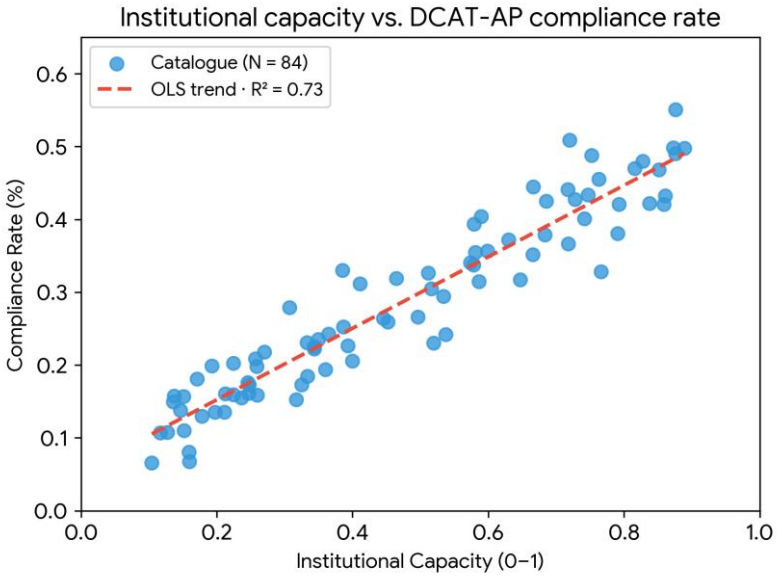

Figure 4: Institutional Capacity Vs. DCAT-AP Compliance Rate

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

Each dot represents one catalogue (N=84). The dashed line represents the OLS fit (R2 = 0.73, $p < 0.001$). The highly significant positive correlation supports the finding that institutional administrative capacity (rather than legislative enforcement) is the strongest determinant of adoption of DCAT-AP standards.

Dissecting Workaround Topologies

An examination of the 18 catalogs assigned to the category of "workaround" ow reveals systematic practices of administration evasion. Interestingly, 72% of those non-conformant portals had a consistent percentage of explicit literal placeholders like "N/A", "null", or "temp" and were marking an average of 11% of all of their total metadata records. Two other statistics of interest were also observed: a high prevalence (across an overwhelming 83 percent of the cases) of deliberately expanding blank nodes in the publisher properties, and a high prevalence of workarounds named in the publisher properties. However, this specific type of evasion approach can provide a welcomed respite for local government managers from the need to constantly manage long-running IRI infrastructures, potentially resulting in catastrophic consequences for the underpinnings that is needed for the seamless federated indexer (Maratsi et al., 2024).

Domain Extensions: Architectural Profiles

The interaction between domain profiles and the core baseline is synthesized in Table 5.

Table 5: Domain Extensions and Architectural Profiles

Profile Extension	Sector Focus	Mandatory Additions	Interoperability Trade-off
GeoDCAT-AP	Spatial data	Geometry schemas, spatial anchors	High spatial accuracy; reduces cross-theme keyword alignment
HealthDCAT-AP	Health records	Sensitivity labels, access rights	Enforces GDPR; restricts automatic cross-border harvesting (Derycke et al., 2025)
CiteDCAT-AP	Academic datasets	DOIs, ORCIDs	Streamlines citation; increases entry friction (Arisi, 2022)
DCAT-AP HVD	High-value datasets	API endpoints, update logs	Guarantees feeds; requires high technical capacity (Broomfield, 2023)

Journal of Information System and Technology Management
JISTM
EISSN: 0128-1666
Volume 11 Issue 43 (June 2026) PP. 242-260

Discussion

From our practice-based study, it is evident that the process of translation of standardized information management policies is an inherently fragmentary one and is highly influenced by the extent of local government administrative settings, as well as basic management capacities. Legislative policies, legislative decision documents (for example, the EU Open Data Directive) cannot simply be assumed to emerge spontaneously within the institutional landscape: decisions are always negotiated, and sometimes awkwardly and more or less subtly, into the existing landscape. It is important to remember that legislative instruments of which the EU Open Data Directive is only the latest example, cannot simply be 'dropped' as it were, into a uniform technical system; they are always enmeshed in a more or less ambiguous and messy way in an existing set of institutions.

Theoretical Implications

This research systematically challenges the official story that sees deviations from structure as a sign of poor technical ability. Portal variability can be re-conceptualized by discussing how "structural adaptations" oa and "calculated workarounds" ow are in many cases very rational responses that make sense when there are strong institutional constraints. In cases of very decentralized governance models and very limited technical resources, catalog administrators, naturally, fall back on the workaround paradigm. With this local approach they are able to present a picture of formal statutory law compliance whilst avoiding burdensome operation costs which would otherwise compromise their long-term institutional sustainability (Lendvai & Stubbs, 2007; Clarke et al., 2015).

This behaviour pattern reflects well with the monitoring friction highlighted by Salamoura & Tsakonas (2024) and the crippling administrative friction as noted by Hrynaszkiewicz et al. (2020). Accordingly, the policy scaffolding in the future should take note that, for the most part, adoption is interpretive and so it should not include "draconian validation rules" to achieve standard adoption. Moreover, it can be argued that the statistically very high level of correlation between the institutional capacity and the compliance rate ($R2 = 0.73$) shows that any introduction of setting technical specifications without simultaneously introducing proper organizational institutional capacity is completely meaningless, reinforcing the premise that it is not much sense to implement technological specifications in a vacuum but rather coincide with the proper institutional capacity building (Masrek et al., 2025; Kamarulzaman et al., 2025).

Practical Implementation for Platform Governance

This research revealed that three strategic shifts in operations are required for the governance of digital platforms in the future:

1. Trying to escape the Strict Validation Trap: Under-resourced repositories can only solve increasingly convoluted and opaque problems by persisting automated SHACL rules (this is clearly visible in the Eastern European group). There is a need to evolve legislative and technical attention from punitive validation to proactive capacity building.
2. Increasing Social Interoperability: Ecosystems depend on common infrastructure. The creation of centralized persistent identifier registries, national level coordination desks

and unified training programs are of critical importance. Such ongoing investments in infrastructure have directly translated to the best compliance outcomes throughout Northern Europe (Kamarulzaman et al., 2025).

3. Controlled Adaptation as a Design Property: Central harvesting authorities should strive to push, rather than the costly illusion of uniformity of schema, structured modular templates that allow local domain extensions to be added. This, today, is the only sustainable middle ground – the empirical reality is, that 54% of the ecosystem is in an adaptive state already today.

Restrictions and Prospects for the Future

This study is methodologically robust however is inherently limited to the European DCAT-AP ecosystem which is why it has methodological and geographical limitations. Future lines of research should strongly put this policy translation framework to the test by replicating and applying it in divergent geographical contexts (e.g., using DCAT-LAT in Latin America). Another key feature that will be important to incorporable will be the longitudinal analyses of time series data to see more clearly how these localised modes of translation develop or change over longer time periods (Dortmund, 2025).

Conclusion

This investigation revealed and mapped in a critical manner the complex operational trajectory, in which a generalized global metadata specification (DCAT-AP) is fractured and translated into very localized practices, across a spectrum of 84 different data portals. The results conclusively show that standards conformance is not a black-or-white measure; it is a highly fluid, structurally complex form of policy translation. Given the asymmetries associated with the institutional capacity, the local legal landscape and the glaring lack of resources, portal administrators tend to settle into different ways of operating: either do the correct thing, adjust the rules to fit the context or find a workaround to work around the system.

This study is pioneering in providing both highly precise quantitative measures underneath the Semantic Deviation Index (SDI) and validates effectively this approach to digital repository forensic evaluation by testing it with automated queries against a triplestore. But "technical convention," ever more complex by itself simply won't get us to "real cross-border 'digital interoperability'." This critically needs evolving governance modalities from simpler to more empathetic that focus on the diversity of institutional components and invest in and strengthen local institutions.

Suggestions for Future Research

Future studies should not just explore what these metadata inconsistencies are, but also the causal institutional structures that underlie localized inconsistencies. Comparative analyses of public administration systems in various countries, in particular those in emerging East Asian or Middle Eastern public sector registries, with a longitudinal dimension would greatly increase the generalizability of the policy translation framework (Lendvai, 2015; Mukhtarov, 2014). In fact, the systematic local interaction of the organisational capacities (Cap) with the resources constraint (Res), and the emergence of a set of recurring workaround domains (ow) is still a major space for empirical research.

Also, there is a lack of research on the long-term consequences of structural and semantic workarounds on global data discoverability. Future research could focus on the issue of whether non-standard metadata changes over time are permanently affecting the usefulness of these repositories or if these shared portals become "dark data" silos which are unindexed by later search engine crawlers (Md Ajis et al., 2024; Jafalizan Md Jali et al., 2025). Also, the research on automated, machine learning based schema mapping and semantic web validation tools (W3C, 2025) looks like a suitable course of action to mitigate the compliance problems for resource-constrained portals while ensuring global semantic interoperability.

Declaration of Generative AI and AI-Assisted Technologies: This work was prepared using generative AI models to assist in formatting the manuscript, refining mathematical equations, and verifying citation consistency. The tools were employed solely for technical support and structural refinement. The authors reviewed and edited the output to ensure accuracy and accept full responsibility for the final content of this publication.

Acknowledgements:	The authors would like to express their gratitude to the Faculty of Information Science, Universiti Teknologi MARA (UiTM) for providing institutional support. Many thanks also to colleagues and peers who reviewed and helped with the research process, providing valuable insights and feedback which greatly enhanced the quality of this paper.
Funding Statement:	This research received no financial support. The study was conducted entirely using institutional and personal resources.
Conflict of Interest Statement:	The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the Journal of Information System and Technology Management (JISTM).
Ethics Statement:	This study did not involve any human participants, animals, or sensitive personal data requiring specific ethical approval. The authors confirm that the research was conducted in absolute accordance with accepted academic integrity and ethical publishing standards.
Author Contribution Statement:	All authors contributed significantly to the development of this manuscript. Adam Hafiz was responsible for the fieldwork, literature review, conceptualization, and coordinating the write-up of the overall text. Arman Hanif formulated the research methodology, handled the data entry process, and managed metadata extraction tools. Aufa Hannah analyzed the data, performed the SPARQL/SHACL validation pipelines, and drew critical conclusions from the results. All authors have read and approved the final version of the manuscript.

References

Abd Halim, M. A., & Abdul Rahman, S. (2024). Uncovering patterns in online database usage at UiTM Negeri Sembilan: A data mining approach. Journal of Information and Knowledge Management, 14(2), 1–11. https://doi.org/10.24191/jikm.v14i2.3664

Arisi, M. (2022). Open knowledge: Access and re-use of research data in the European Union open data directive and the implementation in Italy. Italian Law Journal, 8(1), 33–51.

Azmi, N. M. W., Mohd Zahibi, M. F. H., Tarudin, M. D. A., & Shahidan, N. F. (2025). AI governance in the public sector: A literature review. Journal of Mathematics and Computing Science, 11(1), 1–15. https://doi.org/10.24191/jmcs.v11i1.8090

Broomfield, H. (2023). Where is open data in the Open Data Directive? Information Polity, 28(2), 175–188. https://doi.org/10.3233/IP-220053

Clarke, J., Bainton, D., Lendvai, N., & Stubbs, P. (2015). Making policy move: Towards a politics of translation and assemblage. Policy Press. https://doi.org/10.46692/9781447313380

Conde, J., Pozo, A., Munoz-Arcentales, A., Choque, J., & Alonso, Á. (2024). Fostering the integration of European open data into data spaces through high-quality metadata. arXiv. https://doi.org/10.48550/arXiv.2402.06693

Derycke, P., Barros, B. J., Schutte, N. M., Vande Catsyne, C.-A., & Fonseca, M. B. (2025). Designing DCAT-AP extensions for common European data spaces: The EHDS HealthDCAT-AP case study. NeXt-Generation Data Governance Workshop 2025. https://openreview.net/forum?noteId=lf4Tsy0uyz

Dortmund, A. (2025). Envisaging the future of metadata. UKSG eNews, 581. https://doi.org/10.1629/uksg.581

Gujar, P. (2025). Data standardization and interoperability. In P. Gujar, Data usability in the enterprise (pp. 89–110). Apress. https://doi.org/10.1007/979-8-8688-1183-8_4

Hrynaszkiewicz, I., Simons, N., Hussain, A., Grant, R., & Goudie, S. (2020). Developing a research data policy framework for all journals and publishers. Insights: The UKSG Journal, 33, 1–15. https://doi.org/10.1629/uksg.511

Hussin, N., Ismail, A. A. N., Mustakim, N. I., Daud, R., Bunawan, A. A., & Ahmad, M. N. (2025). Metadata integration in electronic records management systems: A comparative study of international best practices. Journal of Information and Knowledge Management, 15(2), 43–54. https://doi.org/10.24191/jikm.v15i2

Jafalizan Md Jali, Duan, M. K., & Ahmad Fuzi Md Ajis. (2025). Illuminating dark data: A systematic literature review of its characteristics and management strategies. Insight Journal, 12(2), 71–80. https://doi.org/10.24191/ij.v12i2.5878

Kamarulzaman, M. R., Ismail, A. N., Abdul Kadir, I. K., Alwi, M., & Yunos, M. (2025). Implementing data governance to enhance data management at higher institution: UiTM data governance framework. Journal of Information and Knowledge Management, 15(SI2), 240–251. https://doi.org/10.24191/jikm.v15iSI2.7390

Lendvai, N. (2015). Soft governance, policy fictions and translation zones: European policy spaces and their making. In J. Clarke, D. Bainton, N. Lendvai, & P. Stubbs (Eds.), Making policy move: Towards a politics of translation and assemblage (pp. 115–134). Policy Press. https://doi.org/10.46692/9781447313380.005

Lendvai, N., & Stubbs, P. (2007). Policies as translation: Situating trans-national social policies. In S. M. Hodgson & Z. Irving (Eds.), Policy reconsidered: Meanings, politics and practices (pp. 173–191). Policy Press.

Maratsi, M. I., Alexopoulos, C., & Charalabidis, Y. (2024). On the semantic analysis of open (government) data portals' metadata provision and schema. Proceedings of the 17th

International Conference on Theory and Practice of Electronic Governance, 147–157. https://doi.org/10.1145/3680127.3680130

Martinez-Gil, J. (2025). Framework to automatically determine the quality of open data catalogs. Expert Systems with Applications, 260, 128379. https://doi.org/10.1016/j.eswa.2024.128379

Masrek, M. N., Baharuddin, M. F., & Heriyanto, H. (2025). Exploring the impact of social media utilization on perceived information overload: The mediating role of information disorder experiences. Journal of Information and Knowledge Management, 15(SI2), 42–54. https://doi.org/10.24191/jikm.v15iSI2.7237

Md Ajis, A. F., Kamaruddin, F. E., Mohd Taha, F. N., & Mohamad Tanuri, Z. A. (2024). Core catalyst conditions influencing dark data emergence: A grounded theory approach. Insight Journal, 11(2), 184–197. https://doi.org/10.24191/ij.v11i2.2004

Mhlongo, S., Mbatha, K., Ramatsetse, B., & Dlamini, R. (2023). Challenges, opportunities, and prospects of adopting and using smart digital technologies in learning environments: An iterative review. Heliyon, 9(5), e16348. https://doi.org/10.1016/j.heliyon.2023.e16348

Mohd Taib, N., Shafaruddin, A. D. R., Fauzi, N. A., & Yuhanis, N. (2025). Exploring the factors contributing to the lack of commitment to governance disclosure among Malaysian companies. Insight Journal, 12(1), 34–44.

Mukhtarov, F. (2014). Rethinking the travel of ideas: Policy translation in the water sector. Policy & Politics, 42(1), 71–88. https://doi.org/10.1332/030557312X655511

Niederhauser, M. (2025). Uncovering EU external multilevel governance: The implementation of EU data protection law in Switzerland. JCMS: Journal of Common Market Studies. Advance online publication. https://doi.org/10.1111/jcms.70040

Nogueras-Iso, J., Lacasta, J., Ureña-Cámara, M. A., & Ariza-López, F. J. (2021). Quality of metadata in open data portals. IEEE Access, 9, 60364–60382. https://doi.org/10.1109/ACCESS.2021.3076116

Salamoura, A., & Tsakonas, G. (2024). On the challenges of open access monitoring. Insights: The UKSG Journal, 37, 1–10. https://doi.org/10.1629/uksg.641

Sathvika, Y., & Jeyalakshmi, J. (2026). Data interoperability: Enabling smooth cross-platform communication. In P. Samuel, D. Arockiam, B. Balusamy, L. M. Gernal, & C. Prabha (Eds.), Learning-driven data fabrics for sustainability (pp. 39–51). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-09090-4_4

Taylor-Grant, R. (2024). Adoption of open research practices exceeding expectations. UKSG eNews, 597. https://doi.org/10.1629/uksg.597

World Wide Web Consortium. (2025). Shapes constraint language (SHACL) 1.2 recommendation. https://www.w3.org/TR/shacl12/