DCPAPERS

Geographical Provenance of Open Government Datasets: Evaluating Geospatial Metadata in Municipal Open Data Portals

Florian Hahn[1,*]

[1] *Chemnitz University of Technology, Straße der Nation 62, 09111 Chemnitz, Germany*

Abstract

The study investigates the hypothesis that the adoption of metadata standards such as DCAT in German municipal Open Data portals improves the assignability of datasets to specific geographical regions. It focuses on municipal Open Data portals and compares those that utilise DCAT-compliant metadata structures to portals with non-standard formats, such as ArcGIS JSON. The harvesting and analysis of metadata from each portal is conducted via public endpoints. The classification routine is based on metadata fields such as geocoding URI, dct:spatial and keywords. Each dataset is evaluated for its regional geographical capability. Preliminary results suggest that DCAT-based portals significantly outperform non-standard implementations in terms of regional assignability of datasets, error rate, and metadata completeness.

Keywords

Geospatial Metadata, DCAT-AP, Open Government Data Dataset Metadata, Open Data Portal Metadata, Open Government Data

1. Introduction

The Open Data landscape is characterised by fragmentation, with municipalities implementing a variety of metadata practices. Although DCAT-AP is promoted as a unifying metadata profile in Europe[1], its practical impact on analytical usability remains underexplored, specifically in terms of the assignability of datasets to regions. The present study investigates whether portals utilising DCAT/DCAT-AP/DCAT-RDF/CKAN/SPARQL demonstrate superior geographical provenance for regional datasets in comparison to non-standard systems such as ArcGIS Hub via JSON/XML. Chosen portal categories are selected according to Wenige et al.[2].

2. Research Question and Hypothesis

Research Question: Does the utilisation of DCAT-compliant metadata in municipal German Open Data portals enhance the assignability of datasets to correct regions in comparison with non-standardised alternatives?

DCMI-2025 International Conference on Dublin Core and Metadata Applications

*Corresponding author.

florian.hahn@informatik.tu-chemnitz.de (F. Hahn)

https://www.tu-chemnitz.de/informatik/dm/team/fh.php.en (F. Hahn)

0009-0008-1126-9319 (F. Hahn)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

This article is licensed under a **Creative Commons Attribution 4.0 International License**, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and cite the source.

https://doi.org/10.23106/dcmi.952512517

Hypothesis: Municipal portals utilising DCAT-compliant metadata demonstrate considerably higher proportions of regionally assignable datasets in comparison to portals employing non-standardised systems.

3. Methodology

We selected german municipal portals (ADM1/ADM2) on prior technical experience and reliable API access representing both standardised (DCAT-AP/DCAT-RDF/CKAN/SPARQL) and non-standardised metadata implementations. Using a script[1], metadata was systematically harvested via CKAN APIs, SPARQL endpoints, REST interfaces or scraping. Each dataset was classified as assignable (information is found and accurate), unassignable (information is not found), or incorrect (information is found but refers to the wrong provenance).

4. Expected Results

Preliminary results show that DCAT-compliant portals achieve ∼90% regionally assignable datasets vs. ∼55% for non-standard portals.

Table 1
Expected Portal Comparison Summary (in %)

Portal Type	Assignable	Unassignable	Incorrect
DCAT (DCAT-AP/DCAT-RDF/CKAN/SPARQL)	90	4	6
Non-standard (JSON/XML/ArcGIS/OpenDataSoft)	55	17	28

5. Recommendations

We propose the following actions for municipalities:

1. Adopt DCAT-AP 3.0 and GeoDCAT-AP [3] for structured metadata exchange
2. Provide spatial metadata via `dct:spatial`[4], using standard URIs or bounding boxes
3. Use CKAN or other systems with RDF export functionality[5]
4. Use Political Geocoding URI[6] based on regional keys[7] from DCAT-AP or Spatial URI.

6. Conclusion

DCAT-compliant portals provide more regionally assignable metadata, confirming the hypothesis. The limitations of the work are the portals that were not selected, as these do not provide easy access to the required information. Future work will include them and a mapping portals by administrative level. This supports the strategic adoption of metadata standards in Open Data governance for analytical interoperability across jurisdictions.

[1]The full metadata extraction and classification script is implemented and available at: https://github.com/SODIC-research/SODRAM.

OPEN ACCESS
This article is licensed under a **Creative Commons Attribution 4.0 International License**, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and cite the source.
https://doi.org/10.23106/dcmi.952512517

References

[1] SEMIC Support Centre / European Commission, Dcat application profile for data portals in europe (dcat-ap) — version 3.0.0, https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe/release/300, 2024. Published 07 August 2024; Accessed: July 17, 2025.

[2] L. Wenige, C. Stadler, M. Martin, R. Figura, R. Sauter, C. W. Frank, Open data and the status quo – a fine-grained evaluation framework for open data quality and an analysis of open data portals in germany, arXiv preprint arXiv:2106.09590 (2021). URL: https://arxiv.org/abs/2106.09590. doi:10.48550/arXiv.2106.09590.

[3] GeoDCAT-AP Working Group / European Commission, Geodcat-ap — geospatial extension for dcat-ap, version 3.0.0, https://semiceu.github.io/GeoDCAT-AP/releases/3.0.0/, 2024. Released 04 October 2024; Accessed: July 17, 2025.

[4] I. Magnisalis, N. Comet, S. Schade, Towards a standard-based open data ecosystem: analysis of dcat-ap use at national and european level, Electronic Government, an International Journal 18 (2022) 137–180. doi:10.1504/EG.2022.10036962.

[5] Open Knowledge Foundation, Ckan dcat extension — dcat and rdf export for ckan, https://docs.ckan.org/projects/ckanext-dcat/en/latest/, 2025. Accessed: July 17, 2025.

[6] GovData / DCAT-AP.de Coordination, Political geocoding level vocabulary, https://www.dcat-ap.de/def/politicalGeocoding/Level/, 2023. Controlled vocabulary for administrative level classification; Accessed: July 17, 2025.

[7] GovData / DCAT-AP.de Coordination, Political geocoding municipality key vocabulary, https://www.dcat-ap.de/def/politicalGeocoding/municipalityKey/, 2025. Controlled vocabulary for municipality identification; Accessed: July 17, 2025.

OPEN ACCESS

This article is licensed under a **Creative Commons Attribution 4.0 International License**, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and cite the source.

https://doi.org/10.23106/dcmi.952512517