# UVAJANJE TEHNOLOGIJE LINKED DATA V GEODETSKIH UPRAVAH: SISTEMATIČNI PREGLED LITERATURE

# GEOSPATIAL LINKED DATA PROLIFERATION IN NMCAS: SYSTEMATIC LITERATURE REVIEW

*Marjan Čeh, Jernej Tekavec*

UDK: 528:028.4:004.005.55
Klasifikacija prispevka po COBISS.SI: 1.02
Prispelo: 10. 4. 2023
Sprejeto: 31. 5. 2023

DOI: 10.15292/geodetski-vestnik.2023.02.213-234
REVIEW ARTICLE
Received: 10. 4. 2023
Accepted: 31. 5. 2023



## IZVLEČEK

*V raziskavi smo se osredotočili na sistematični pregled literature o objavljanju geoprostorskih podatkov v tehnologiji povezanih podatkov (linked data – LD) pri geodetskih upravah evropskih držav, odgovornih za zemljiško administracijo. Analizirali smo znanstvene študije, objavljene v letih od 2014 do 2023. Odkrivali smo vzroke za razmeroma počasno uveljavljanje LD-tehnologij za objave geoprostorskih podatkov. Metodologija temelji na analizi problemov, posegov, primerjav, izidov in okoliščin (PICOC) ter iskanju s tako imenovanim »vzvratnim valjenjem snežne kepe«. Postopek presoje upravičenosti za vključitev člankov v sistematični pregled literature je predstavljen v Smernicah za prednostne postavke poročanja za sistematične preglede in metaanalize (PRISMA). V naboru devetnajstih izbranih člankov smo analizirali vsebino in identificirali dvanajst ključnih področij. S frekvenčno analizo smo pridobili pregled LD objavljenih geoprostorskih tem, ki so najpogosteje objavljene pri geodetskih upravah sedmih evropskih držav. Na podlagi petih raziskovalnih vprašanj smo ugotovili, da lahko poleg sistematične podpore razvijalcem besednjakov ter ontologij na uspeh objave LD vplivajo še zaupanje med deležniki, ozaveščenost razvijalcev in poznavanje najboljših praks LD izkušenih geodetskih uprav.*

## KLJUČNE BESEDE

povezani podatki, geodetske uprave, geoprostorska semantika, interoperabilnost, sistematični pregled literature

## ABSTRACT

*This research focuses on a systematic literature review of the geospatial Linked Data (LD) publication at National Mapping and Cadastral Agencies in Europe (NMCA) responsible for land administration. We analysed scientific studies published from 2014 to 2023. Our research aimed to find out the reasons for the relatively slow adoption of geospatial data publication as LD. Therefore, we searched for the most common problems, solutions, and challenges for publishing prioritised data themes in NMCAs. Applied research methodology relies on well-established approaches such as analysing Population Problems, Intervention, Comparison, Outcomes, and Circumstances (PICOC) and backward snowballing. The eligibility process is presented by Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines. We analysed nineteen extracted papers and identified twelve key issues from full-text reading. Frequency analyses resulted in the overview of the most frequently published geospatial themes as LD in several European countries. We addressed five research questions and concluded that in addition to systematic support for developing vocabularies and ontologies, trust, awareness, and knowledge about best practices could influence the success of LD publication at NMCAs.*

## KEY WORDS

Linked Data, National Mapping and Cadastral Agency, NMCA, geospatial semantics, interoperability, Systematic Literature Review, SLR

## 1 INTRODUCTION

National Mapping and Cadastral Agencies (NMCA) store large amounts of data (Big Data) and update official geodatabases in a data warehouse architecture as data silos (Crompvoets et al., 2016). They must be referenced and linked to datasets from other national administrations and agencies or commercial sources. The Directive on open data and the reuse of public sector data (2019) of the European Commission mandate the creation of conditions for multiple reuses of public sector data, which requires online technologies for geospatial services - location-based services and data publication distribution with enabled connectivity.

Spatial Data Infrastructures (SDIs) have enhanced data discovery and accessibility based on OGC (Open Geospatial Consortium) web services. However, finding, accessing, and using data disseminated via SDIs are still tricky for non-expert users (Vilches-Blázquez and Saavedra, 2019). Combining spatial data from SDIs using Semantic Web technologies is an important step forward (Wiemann and Bernard, 2016).

As observed by Crompvoets and Vancauwenberghe (2018), since 1990, most European countries have established strategic advisory boards to define the SDI strategy and systems to control SDI implementation.



Linked data are vital in spatial Data Infrastructure (SDI) (Huang et al., 2019) as interoperability enables software connection of data and information from different sources that can be used to make better decisions and improve the efficiency of various operations. Current SDIs face the problem of data integration and semantic heterogeneity due to their partially siloed data organisation. The production of official geospatial data is often in the hands of highly specialised public agencies that have traditionally followed their paths and established production frameworks (Ariza-López et al., 2021). These siloed data are isolated and cannot be shared and integrated across organisations and communities (Vilches-Blázquez & Saavedra, 2019).

The integration process may involve various data preparation techniques, including data selection, cleansing, transformation, and delivery (extract, transform, load - ETL) (Folmer et al., 2019; Mallek et al., 2020). Geoportals are the visible parts of SDIs focusing on interoperability by implementing standards such as (the Open Geospatial Consortium) OGC web services for finding and using geographic data and services (Owusu-Banahene, 2018).

Open standardisation Ma (2017) for web services in the domain of geospatial data linkage plays an important application role in the implementation of interoperability using linked data principles and online publication of data. Geospatial data interoperability refers to the ability of different spatial information systems, software, and spatial data formats to harmonise (Lanucara et al., 2019).

At different events, such as the ENDORSE conference on reference data and semantics and the Conference on Formal Ontologies in Information Systems, to name just two, experts also discuss barriers to the uptake of LD.

### 1.1 Motivation for literature review

This study was motivated by a desire to present the structure and scope of research at the intersection between the semantic open-linked data and geospatial data in NMCAs. The research challenge was also

to determine the current state of the art and to identify scholars considered pioneers in the geospatial linked data web publication who focus their research on specific NMCAs problems. Such authors can be crucial for the dissemination of knowledge and the advancement of the research field. Finally, the main driver for conducting the literature review described in this article is the slow proliferation of LD and the non-existence of such studies in the domain of NMCAs. We are interested in understanding the barriers to the uptake of LD in general (Attard et al., 2017; Polleres et al., 2020). The paper is intended to provide us and professionals involved in the strategic and operational development of NMCA with insight into the current state of geospatial data publishing as Linked Data. We are interested in recent expansion, issues, and challenges of LD, published in journal articles and scientific reports available as open literature in the last decade.

New technologies are not easy to implement, as shown by assessments in the form called Gartner Group's hype cycle (O'Leary, 2008). The Smart world is one of four recent Gartner's (Nguyen, Jump and Casey, 2022) most influential themes. The emerging technologies with the most significant impact are Smart spaces, Metaverse, and Digital twins, all of which are also in the NMCA domain of interest. The theme of Critical enablers includes Knowledge Graphs where Linked data participate as an enabling technology. This article provides a systematic review and comments on the literature in the field of geospatial data linkage for LD to identify the current status of the use of LD in geospatial administrations (NMCAs). We raise the question of the challenges faced by those administrations that have already started using the LD semantic technology as open-linked data (OLD). We are also interested in the reasons for the relatively slow uptake of interlinking geospatial technologies. We used the systematic literature review method described in more detail in Section 2 to analyse the available literature. Several previous studies provide a review of the literature in the field of Linked Data. With this review, we attempt to close the gap in reviewing the literature, specifically the NMCA's publishing geospatial data.

### 1.2 Consideration of research questions

Over the past decade, the Linked Data concept has shown a strong proliferation tendency into various geospatial domains, including NMCAs. Several authors have reported early successful implementation of geospatial data in the form of the Semantic Web, such as Goodwin, Dolbear, and Hart (2008), Hyvönen et al. (2014), Saavedra, Vilches-Blázquez, and Boada (2014), Debruyne et al. (2016) Beek and Folmer (2017).

The current understanding of LD does not yet offer evidence of a robust positive influence or improvement in economic outcomes from potential alternated business models of NMCAs. From the initial analyses of the available literature, it appears that there are some limitations to the proliferation of existing knowledge about the linked data after the invention of the World Wide Web (Berners-Lee, 1996). Our research aims to identify the main challenges and multidisciplinary issues reported in the literature concerning LD implementation in NMCAs and discover what needs to be done to increase its application. The review is based on the following research questions:

- RQ1: What are the main problems that studies address for publishing NMCA geospatial data as LD?
- RQ2: What solutions and requirements are identified by NMCAs with LD publishing experience?
- RQ3: What are priority data themes published by NMCAs from LD active counties?

– RQ4. What challenges/issues have NMCAs identified for the future development of linked data?
– RQ5: What are the reasons for the relatively slow adoption/dissemination of geospatial data LD by NMCAs?

The present work is organised as follows. In the methodology section, we explain the planning phase of the systematic literature review (draft protocol) and then describe how the study was conducted. Finally, we report the results, discuss the synthesis of findings, and answer the research questions. In conclusion, we summarise the main findings.

## 2 METHODOLOGY

The article follows the methodology of a systematic literature review. A systematic review (SR) is a method used to find and summarise all relevant existing evidence related to a particular research question of interest. One of the activities associated with the process of SR is the selection of primary studies (Octaviano et al., 2015). In the preliminary investigation, we conducted a literature review on semantic web technologies for spatial data. We systematically examined documented research articles describing already implemented semantic online linked data applications (LD). The target domain was limited to the context of spatial data and National Mapping and Cadastral Agencies (NMCA).



### 2.1 Planning the Review and Protocol

Systematic reviews must be conducted according to a predetermined search strategy. The general process of the SR is divided into three phases: planning, conducting, and reporting the review. The phases associated with planning the review are: 1. Identify the need for a review, 2. Prepare a review protocol, and 3. Conduct an SR. The steps associated with conducting the review are: 1. Identify the research, 2. Select primary studies, 3. Assess the study quality, 4. Extract and monitor data, 5. Synthesise data. Reporting the review is a single step (Kitchenham, 2004). In software engineering, the research questions can be adapted to software engineering issues such as assessing the effect of a software engineering technology, assessing the frequency or rate of the adoption of technology, identifying cost and risk factors associated with a technology, recognising the impact of technologies on reliability, reusability, performance and cost models, cost-benefit analysis of software technologies (Person et al., 2021).

Research synthesis is a collective term for a family of methods for summarising, integrating, combining, and comparing the results of different studies on a topic or research question. In software engineering, comparing and contrasting such evidence is necessary to infer empirical support for a phenomenon. Such synthesis may also highlight important areas and questions not adequately addressed in previous empirical research. Research synthesis, then, is a way of understanding the claims of several studies (Cruzes and Dybå, 2010).

### 2.2 Protocol for literature review

The selection of this literature review is based on evidence (studies). We followed the PICOC research approach (Richardson et al., 1995) and formulated research criteria about the Population Problem, Intervention, Comparison, Outcomes, and Circumstances (context). The first question should address

the problem in a particular population, and the question about intervention facilitates the search for a precise solution. The comparison question provides the criteria for comparing the results of the intervention with the status of the population before the intervention. Outcomes indicate the effects of an intervention on the population's problems. Circumstances describe the context of the population with a problem (Booth et al., 2019).

For the selection of geospatial LD research studies, we provided criteria items that corresponded to the aspects of the PICOC approach:

1. Problem: Siloed geospatial data
2. Intervention: Publishing geospatial data as Linked Data using semantic web technologies
3. Comparison: web searchability, linkability
4. Outcomes: machine-readable, published geodata as open-linked data
5. Circumstances: National Mapping and Cadastral Agency (NMCA)

Keywords and synonyms: geospatial, "linked data", publication, "semantic web", "National Mapping and Cadastral Agency", NMCA

Databases: SCOPUS, IEEE, Google Scholar

Criteria for the selection of studies:

- Inclusion of publications: Journal articles, conference papers, scientific reports;
- Exclusion of publications: duplicates, documents that do not fall within the scope of the land management system;
- Language limitations: English language;
- Time limitation: publication years from 2014 to 2023.

Search query:

TITLE-ABS-KEY (linked data AND geospatial) AND ( LIMIT-TO ( PUBYEAR, 2023 ) OR LIMIT-TO ( PUBYEAR, 2022 ) OR LIMIT-TO ( PUBYEAR, 2021 ) OR LIMIT-TO ( PUBYEAR, 2020 ) OR LIMIT-TO ( PUBYEAR, 2019 ) OR LIMIT-TO ( PUBYEAR, 2018 ) OR LIMIT-TO ( PUBYEAR, 2017 ) OR LIMIT-TO ( PUBYEAR, 2016 ) OR LIMIT-TO ( PUBYEAR, 2015 ) OR LIMIT-TO ( PUBYEAR, 2014 ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )

In exhaustive data selection for eligible studies (43), we scored each study with six measures on the short scale (No = 0, Partially = 0.5, Yes = 1).

Table 1: Assessing studies before detailed data extraction with criteria/questions

| Criteria | Eligibility question before extraction |
|---|---|
| C1: | Is the study relevant and essential to our research question? |
| C2: | Does the study focus on geospatial data? |
| C3: | Does the study clearly define the problems of LD publication? |
| C4: | Does the study explain the challenges of LD publication? |
| C5: | Does the study explain the best practices of LD publication? |
| C6: | Does the study cover the NMCAs domains? |

The results of selecting and evaluating the studies for eligibility by criteria (Table 1) for detailed information extraction are presented in Table 2.

Data extraction

To summarise the information about the studies selected for extraction, we created Table 6 to collect and synthesise information defined by the PICOC research criteria. We collected data on the characteristics of the studies under investigation by prioritising the geospatial data domain (topics), geographic coverage and country experiences, issues addressed in the studies, and future challenges of NMCA development LD.

Research Questions

An essential component of the protocol is the selection of research questions related to the motivation for the research. Research questions are presented in Section 1.2

Limitations of this literature review: In this research, we do not deal with legal issues of LD such as liability - privacy - intellectual property - competition law, technical matters (formats, tools) and standards of open linked data.



## 3 RESULTS

A systematic review attempts to collate empirical evidence from a relatively small number of studies about a focused research question (Higgins et al., 2022). To present the process for the selection of a small number of studies to be included in the review, we used the flowchart template, adopted from the guideline Preferred Reporting Items for Systematic Reviews and meta-analyses PRISMA (Page et al., 2021), Figure 1.

We initially identified 133 records from three source databases: Scopus (61 records, 46%), IEEE (2 records, 2%), and Google Scholar (70 records, 53%). In addition, 13 studies were added manually (a total of 146 records). Before screening titles and keywords, 14 duplicates were removed (9%). When screening titles and keywords, 26 (18%) records were removed because they were not related to our research problem.

One hundred six records (papers) were screened based on abstract (73 % of broadly identified 146 records). Sixty-three records (43 %) were excluded because they were irrelevant in the context of land administration systems (LAS). We screened (by reading the introduction and conclusions of the papers) 43 studies (29 %) for eligibility. Three eligibility criteria excluded twenty-four (16 %) records: not describing the NMCA domain 12 (8 %), not dealing with critical geospatial datasets 7 (5%), and not relevant from the EU location context 5 (3 %). Nineteen studies (13 % of the total 146 initially identified records) were selected for full-text reading, information extraction, analysis, and synthesis.

In the process of careful selection, concerning the 43 selected studies, we assessed each for eligibility with six criteria on the short quality scale (No = 0, Partially = 0.5, Yes = 1) (Table 1). We selected the studies that achieved the interval score [3 – 6] for the extraction.

The selection of studies and reports for full-text reading and key information extraction (19 articles) sorted by year of publication are presented in Table 2. Other studies were excluded from the review.

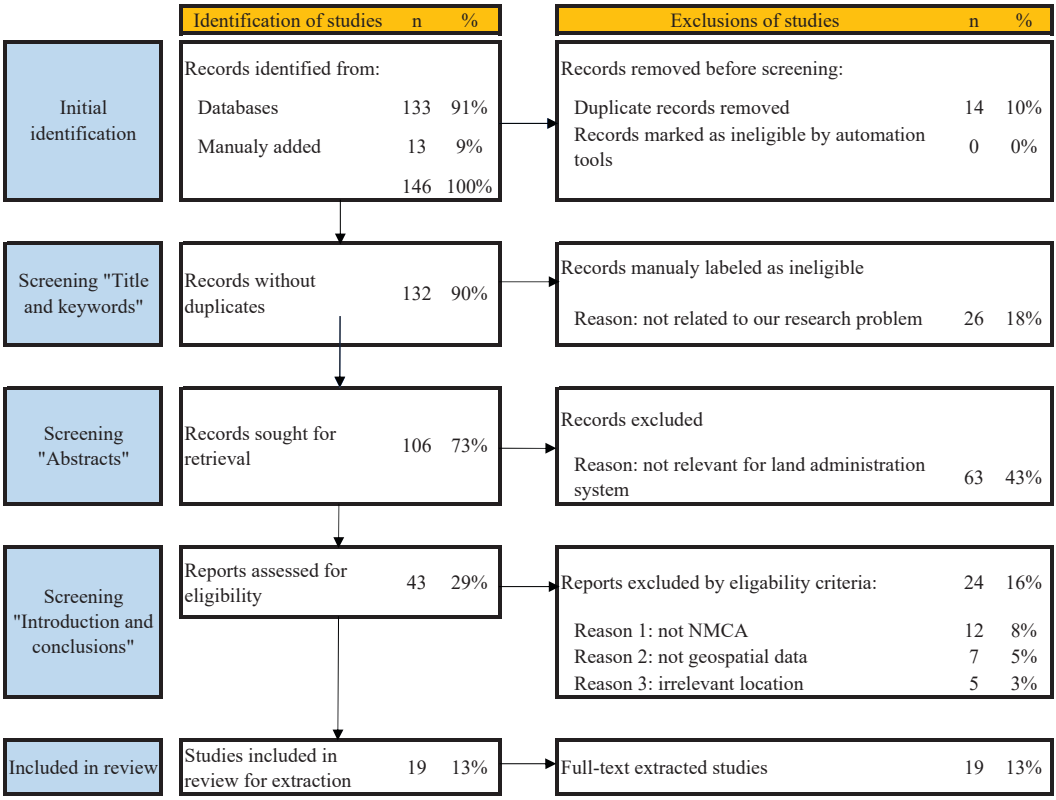


Figure 1: Results of studies selection in the form adopted from the PRISMA 2020 flow chart (Page et al., 2021)

Table 2: Detailed selection of studies and quality assessment criteria

| ID | Author | C1 | C2 | C3 | C4 | C5 | C6 | Total score | % |
|---|---|---|---|---|---|---|---|---|---|
| | | Relevance | Geospatial | Problems defined | Challenges defined | Best practices | NMCA domain | | |
| 1 | Saavedra et al., 2014 | 0.5 | 1 | 1 | 1 | 0.5 | 1 | 4.0 | 67% |
| 2 | Çağdaş and Stubkjær, 2015a | 1 | 1 | 0.5 | 0,5 | 0 | 1 | 4.0 | 67% |
| 3 | Çağdaş and Stubkjær, 2015b | 1 | 1 | 0.5 | 1 | 0 | 1 | 4.5 | 75% |
| 4 | Hietanen et al., 2016 | 0.5 | 1 | 0.5 | 0,5 | 0 | 1 | 3.0 | 50% |
| 5 | Beek and Folmer, 2017 | 1 | 1 | 0.5 | 0,5 | 1 | 1 | 5.0 | 83% |
| 6 | Bucher et. al., 2020 | 1 | 1 | 0.5 | 1 | 0.5 | 1 | 5.0 | 83% |
| 7 | Debruyne et.al., 2017 | 1 | 1 | 1 | 1 | 1 | 0.5 | 5.5 | 92% |
| 8 | Kyzirakos et al., 2018 | 1 | 1 | 0.5 | 1 | 0 | 0 | 3.0 | 50% |
| 9 | van den Brink et al., 2018 | 1 | 0 | 1 | 1 | 1 | 0 | 4.0 | 67% |
| 10 | Folmer et al., 2018 | 1 | 1 | 0 | 0 | 1 | 1 | 4.0 | 67% |
| 11 | Ronzhin, 2019 | 1 | 0.5 | 1 | 1 | 1 | 1 | 5.5 | 92% |
| 12 | Forlmer et al., 2019 | 1 | 1 | 0.5 | 1 | 1 | 1 | 5.5 | 92% |
| 13 | Rowland et al., 2020 | 1 | 1 | 0.5 | 0 | 0.5 | 1 | 3.5 | 58% |

| ID | Author | C1 | C2 | C3 | C4 | C5 | C6 | Total score | % |
|---|---|---|---|---|---|---|---|---|---|
| | | Relevance | Geospatial | Problems defined | Challenges defined | Best practices | NMCA domain | | |
| 14 | Folmer et al., 2020, | 1 | 1 | 1 | 1 | 0.5 | 1 | 5.5 | 92% |
| 15 | Yaman et al., 2021 | 1 | 0 | 1 | 1 | 0.5 | 0.5 | 4.0 | 67% |
| 16 | Issa et al., 2021 | 1 | 1 | 1 | 1 | 1 | 0 | 5.0 | 83% |
| 17 | Rowland et al., 2021 | 1 | 0 | 1 | 1 | 1 | 1 | 5.0 | 83% |
| 18 | Rowland et al., 2022 | 1 | 1 | 1 | 1 | 1 | 1 | 6.0 | 100% |
| 19 | Yaman et al., 2022 | 1 | 0.5 | 1 | 1 | 1 | 1 | 5.5 | 92% |
| | | **1.00** | **0.79** | **0.72** | **0.82** | **0.65** | **0.79** | **4.6** | **76%** |

From Table 2, it can be concluded that the most recent publications of selected studies come from Rowland (3), Folmer (3), and Yaman (2). The results from Table 2, sorted by year, confirm that recent publications lead in the score concerning the assessed quality and detailed selection criteria. This may indicate a slight bias in our selection criteria or a recognised need to cover the problematic issues of LD publication by authors who make more effort to provide updated and reliable information about their solutions to solve the problems in the domain of LD publication at NMCAs.



### 3.1 Snowballing of search results of systematic literature review

In software engineering and information systems, it is recommended to perform a systematic literature review, upgraded with the snowballing technique described in (Wohlin et al., 2022) as a hybrid search strategy. The secondary studies (references) may pinpoint gaps or highlight areas that need more attention from researchers or practitioners. It is possible to perform forward snowballing, i.e., identifying articles that have cited the articles found in the basic search or backward snowballing (records from the reference lists) (Jalali and Wohlin, 2012).

In this literature review, we applied backwards snowballing in the reference lists of the initially identified articles for SR of Linked Data, and it returned 1,706 reference studies in total. We were interested in the top ten journals that most frequently publish about geospatial LD (Table 3), and the results of the authors' backward snowballing are summarised in Table 4.

Table 3: Number of publications per journal from snowballing in present SLR

| Journal | Publications |
|---|---|
| International Journal of Geographical Information Science | 32 |
| Computers & Geosciences | 26 |
| Semantic Web | 18 |
| ISPRS International Journal of geo-information | 16 |
| Environmental Modelling & Software | 16 |
| Transactions in GIS | 15 |
| Communications of the ACM | 12 |
| Computers | 11 |
| Semantic Web | 9 |
| GeoInformatica | 9 |

The International Journal of Geographical Information Science published about 20% of the papers dealing with or at least mentioning geospatial linked data. Still, the first three journals in the above list (Table 3) published almost half of the articles in this shortened list of ten journals.

We also identified the most fruitful authors in the domain of Linked Data, and the results of backward snowballing are collected from the initial search, as presented below in Table 4. Some authors are included in our final list of nineteen studies (Folmer, E. and Kyzirakos, K.). Rowland, A., as the first author of the most qualitative studies from Table 2, is not in the list of most fruitful authors (Table 4) because she has published most recently, and her work has not yet been included in references records enough.

Table 4: Backward snowballing searches counted the references per author from the initial examination.

| Author | Publications |
|---|---|
| Janowicz, K. | 26 |
| Lehmann, J. | 23 |
| Koubarakis, M. | 23 |
| Auer, S. | 20 |
| Kraft, M. | 19 |
| Bizer, C. | 17 |
| Corcho, O. | 14 |
| Hitzler, P. | 14 |
| Bereta, K. | 14 |
| Folmer, E. | 13 |
| Kyzirakos, K. | 13 |

Before the extraction phase, we also performed frequency analyses of the words in the sub-selection of articles (43). Of the 25.511 concepts detected automatically from the relevant articles, the twenty most frequent words are listed in Table 5.

Table 5: Frequency analyses of the words in eligible articles

| Word | Count | % |
|---|---|---|
| data | 7116 | 27.9% |
| spatial | 2401 | 9.4% |
| linked | 1491 | 5.8% |
| geospatial | 1468 | 5.8% |
| information | 1403 | 5.5% |
| web | 1397 | 5.5% |
| semantic | 1169 | 4.6% |
| RDF | 946 | 3.7% |
| org | 834 | 3.3% |
| level | 778 | 3.0% |

| Word | Count | % |
|---|---|---|
| geo | 756 | 3.0% |
| quality | 674 | 2.6% |
| query | 672 | 2.6% |
| datasets | 659 | 2.6% |
| knowledge | 658 | 2.6% |
| geometry | 647 | 2.5% |
| GEOSPARQL | 643 | 2.5% |
| ontology | 624 | 2.4% |
| geographic | 596 | 2.3% |
| model | 579 | 2.3% |

Following the sequence of expected general terms used in the initial search, on top of the sorted collection (Table 5), technological LD concepts are high in frequency, such as Web (5.5 %), RDF (3.7 %), GEOSPARQL (2.5 %) and ontology (2.4 %). The concepts of RDF and GEOSPARQL are demonstrated here as necessary for LD publication but are not explicitly considered in this review. We extracted information from the selected studies for further analyses and syntheses presented in Section 4.



Extraction of data

In a systematic review, data extraction captures critical characteristics of studies in a structured form based on information from journal articles and reports. The most common fields for extraction in literature reviews are defined in the PICOC framework (problem, intervention, comparison, outcome, circumstances) (Schmidt et al., 2021). In this step of the systematic review, we developed evidence tables, in which we collected vital information from each study (using the PICOC framework as a guide) and a summary table that provided a high-level overview of the critical findings from our review evidence. Summary Table 6 helped to perform a quantitative synthesis.

## 4 ANALYSES AND SYNTHESIS OF SMAS LINKED DATA PUBLICATION

Research synthesis is a collective term for a family of methods for summarising, integrating, combining, and comparing the findings of different studies on a topic or research question. Research synthesis is a way of making sense of a collection of studies, e.g., Cruzes and Dybå (2010). Limited attention is paid to research synthesis in software engineering, but comparing and contrasting evidence is necessary to conclude empirical support for LD. This systematic literature review has exposed several problems to be solved and challenges for the future development of publishing and interlinking geospatial data maintained by NMCSs.

To classify the given LD key characteristics discovered in the selected studies, we introduced five research and development concepts to guide us through the final phase of the literature review: a synthesis. The proposed concepts adopted from the PICOC framework are “problem”, “solution” (intervention), “to be improved” (comparison), “requirement” (goal, outcome) and concept “circumstances” (context).

Table 6: Summary table for the synthesis of data collected from nineteen selected studies on issues of geospatial LD publishing

| item | LD issues | description | study ID | count | % |
|---|---|---|---|---|---|
| | | | | | |
| 1 | interoperability | shared, open data accessibility, reuse, integration - combining different sources | 3, 4, 8, 9, 12,15, 17 | 7 | 14,3% |
| 2 | vocabulary | common thesaurus for the domain terminology, multilingual terminology, cross-border terminology | 2, 3, 4, 6, 7, 9, 18 | 6 | 12,2% |
| 3 | support | users support browsing, searching, querying, analyses, viewing | 5, 9, 10, 12, 13, 17 | 6 | 12,2% |
| 4 | heterogeneity | disjoint semantically, structurally, and syntactically; diversity of contexts, terms definitions | 1, 6, 9, 11, 14, 19 | 5 | 10,2% |
| 5 | ontology | knowledge, semantics, relations, explaining instances meaning to the outside world | 4, 7, 9, 11, 18 | 5 | 10,2% |
| 6 | quality | LD publication quality metrics as completeness, updating, accuracy | 1, 11, 16, 19 | 4 | 8,2% |
| 7 | readability | readability by humans and by machines | 4, 6, 9, 15, 18 | 4 | 8,2% |
| 8 | big data | publishing large datasets, siloed data | 5, 8, 9, 14 | 4 | 8,2% |
| 9 | efficiency | business issues of OLD, publication automation, ETL, time- and cost-efficient delivery, restrictions | 8, 14, 17, 18 | 4 | 8,2% |
| 10 | data context | metadata - information about data | 2, 4 | 2 | 4,1% |
| 11 | serialisation | provide data in different LD formats | 4 | 1 | 2,0% |
| 12 | scalability | a scalable approach to linked data publication | 18 | 1 | 2,0% |

Table 6 addresses the issues of publishing LD in nineteen studies analysed. We identified, manually extracted, and analysed twelve key knowledge issues related to research at LD, listed in the "LD issues" column. The frequency analyses of the critical issues are presented in the last two columns of Table 6. The table is sorted in descending order by the frequency of studies dealing with specified LD issues. The frequency of crucial matters from nineteen studies ranges from seven studies (14%) for the knowledge issue "interoperability" to one study addressing "scalability" (2%).

In the following analyses, the concepts of the PICOC framework are written in brackets. The LD items are considered in the order of frequency, as calculated in Table 6.

Item 1: "Interoperability"

> The general term "interoperability" occurs most frequently in the studies, which also define the knowledge term with related concepts such as sharing, open data access, reuse, and integration - combining different sources.
>
> We classified the item interoperability as a *requirement* (*goal, outcome*) for the publishing process of LD because interoperability comprises at least three types (syntactic, structural, semantic), all of which must be met for systems to function seamlessly and interactively.

Item 2: "Vocabulary."

> The knowledge element "vocabulary" is also defined as a "common thesaurus for specialised terminology", "multilingual terminology", and "cross-border terminology" (by authors of studies number 2, 7, and 18).

It is the highest-ranked item among the technical terms on LD, indicating its essential importance for successful publication. We consider building vocabulary a *solution* (*intervention*) component that must be established to enrich the data from the siloed database effectively.

Item 3: "Support "

The item "support" is defined in the studies as helping users browse, search, query, analyse, and view LD. The item has a high score in the presented frequency classification, comparable to the "vocabulary" item, even though it is not technical. User support may be essential in the relatively low prevalence of LD in NMCAs. We note this with the label "*to be improved*" (*comparison*) because LD can work well in the research environment despite insufficient support. However, to incorporate LD into more NMCA processes and user applications, supporting activities and tools need to be provided and improved tremendously.

Item 4: "Heterogeneity"

The knowledge element "heterogeneity" is also described in the studies as "semantically, structurally, and syntactically disjunctive", "diversity of contexts"," and as "heterogeneous definitions of terms".



We see "heterogeneity" as the most critical *problem* that needs to be addressed to achieve good interoperability with LD. Regarding the score in the classification, the term "heterogeneity" corresponds to the frequency of the word "ontology ", which is the next term to be considered.

Item 5: "Ontology"

The knowledge term "ontology" is described in the analysed studies as "knowledge", "semantics", "relations", and "explanation of the meaning of instances for the external world".

We refer to "ontology" as a *solution* (*intervention*) to the heterogeneity problem in the LD publication projects, combined with the previously considered phrase "vocabulary". Both terms together form the basis for LD publication projects.

Item 6: "Quality"

The knowledge element "quality" is described in the studies with quality metrics such as "completeness," "update frequency," and "accuracy." It appears in 8% of the analysed studies, which is the same level as the knowledge items "readability", "Big Data", and "efficiency".

Quality is undoubtedly the general concept that needs *to be improved* (*comparison)* in most areas of information management and LD.

Item 7: "Readability"

The knowledge element "readability" is the readability of data by humans and machines.

Machine readability is a *requirement* (*outcome*) and prerequisite in LD publication projects. However, the readability of linked data is not so obvious or practical for non-specialists. This reason could be another solid obstacle to the inhibited diffusion of LD publication in NMCA.

Therefore, it could be concluded that "readability" "combined with the knowledge aspect of "support" could have the most significant impact on the dissemination of LD in the geospatial field.

Item 8: "Big data" «

Large, siloed datasets are the context of data management in NMCs. The knowledge item "Big Data" is used in the analysed studies relatively as frequently as the previously considered item "Readability". However, if we placed the word "readability" in the *requirement (outcome)* class, we would define Big Data as the *circumstances* class of research.

Item 9: "Efficiency"

The knowledge item "efficiency" is associated in the analysed studies with terms such as "business issues of open LD", "publication automation", "ETL", "time- and cost-efficient delivery", and "constraints". We consider this an attribute of LD that needs *to be improved* (*comparison*) to achieve broader adoption in NMCA geospatial data. Improved efficiency will likely increase the dissemination potential of LD with improved quality and support.

Item 10: "Data context"

The knowledge element "data context" is often described in studies as the availability of metadata (information about data).

We consider the availability of "data context" as one of the *problems* for enforcing the LD paradigm in the practice of NMCAs, since metadata are often too poorly prepared, even with the help of guidelines such as INSPIRE.

Item 11: "Serialisation"

The knowledge term "serialisation" is equated in the studies examined with "the publication of the same data in different formats LD ". The occurrence of the item "serialisation" is meagre, and serialisation in different formats supports different analysis platforms. Therefore, we consider "serialisation" a *prerequisite* (*outcome*) for disseminating LD to NMCAs.

Item 12: "Scalability"

The "scalability" aspect of knowledge refers to the ability of the infrastructure and computer system of LD to increase capacity in response to demand. Since this item is more related to the hardware components of the SDI system, the frequency of this item is low. Nevertheless, we should seriously consider this point as a *requirement* (*outcome*) because the scalability of the LD publishing approach can also significantly improve the efficiency of publishing LD.

From the above synthesis of twelve knowledge points extracted from the nineteen studies analysed, it appears that for a broader application of LD in NMCAs, solutions must be found to overcome these identified problems. Problems of heterogeneity and data contexts (metadata) hinder interoperability requirements such as readability and serialisation. Most importantly, NMCAs must improve the quality, efficiency, and support for publishing and reusability of LD. Most of the answers to these questions can be found in papers published in the last five years, and they will be presented shortly in the discussion section.

## 5 DISCUSSION OF THE RESEARCH QUESTIONS

In this section, we discuss the results of the synthesis phase of the literature review and answer the research questions established in the protocol of this literature review.

### 5.1 RQ1: What problems should be solved when publishing NMCA geospatial data as LD?

The most frequently mentioned problem is the heterogeneity of data, i.e., semantic, structural, and syntactic inconsistency. The issue of heterogeneity, disjointed and siloed data has been pointed out by several authors (Çağdaş and Stubkjær, 2015a; Debruyne et al., 2016; Huang et al., 2019; Huerta, Schade, Granell, 2014; Mai et al., 2019; Ronzhin et al., 2019; van den Brink et al., 2018; Vilches-Blázquez & Saavedra, 2019).

Because of the semantic heterogeneity (Bucher et al., 2020) of legal definitions and terminology among independent land administration government agencies, the top-down approach is hindered (or even impossible) (Ronzhin et al., 2019). Land registries and cadasters were specified according to different data models and standards, which led to interoperability problems (Çağdaş & Stubkjær, 2015a). Even cadastral producers and data are characterised by heterogeneity issues (Huerta, Schade and Granell, 2014). There is a growing gap between the classical ways Geographic Information Systems (GIS) are still used today and the open-ended, exploratory approaches used to retrieve and consume data from knowledge graphs such as Linked Data (Mai et al., 2019). The heterogeneity of the available spatial data makes it difficult for data users, web applications, and services to discover, interpret and use the information in large and distributed web systems (van den Brink et al., 2018). Spatial data tend to be very large (big data), posing difficulties when sharing or consuming open spatial data over the Web (van den Brink et al., 2018).



Our research also revealed the definition of "data context" as one of the *problems* in enforcing the LD paradigm in the practice of NMCAs, as metadata are often too poorly prepared. This means that even metadata standardised and regulated by directives such as "Infrastructure for Spatial Information in the European Community (INSPIRE)," 2007, are not semantically defined within the context-rich enough reference system.

### 5.2 RQ2: What are the solutions and requirements for LD publication at NMCAs?

We identified six essential solutions and their combination for the successful publication of LD in the NMCAs listed in Table 6: Vocabularies, ontologies, metadata, and solutions using standardised technologies that meet requirements such as readability and serialisation. Designing vocabularies to create and publish LD requires understanding the knowledge domain and intended use (Çağdaş and Stubkjær, 2015b; Ronzhin et al., 2019). A standard set of agreed *ontologies,* vocabularies, and keyword sets of geographic *metadata* for describing the relationships between spatial features should be established (Wiemann & Bernard, 2016). Regarding *standards*, OGC has developed an architecture supporting its vision of geospatial technology and data interoperability, called the OGC Abstract, which provides the conceptual foundation for most OGC specifications (Ariza-López et al., 2021). One real-world object can have various data representations and serialisations, sometimes stemming from different organisations or exposing different facets of the available data. (Schleidt et al., 2020). The serialisation of the LD domain is applied in RDF/XML, Turtle, JSON-LD, and HTML (Beek & Folmer, 2017; Hietanen et al., 2016; Schleidt et al., 2020).

### 5.3 RQ3: Which are the priority data themes at NMCAs LD?

NMCAs provide varying amounts of LD data and geospatial themes in the LD publication area. We are interested in the most frequently openly published geospatial topics. Therefore, we summarised the LD data available on the NMCA portals (Table 8) of the selected countries in Table 7.

Table 7: Geospatial themes of selected countries most often published as LD at NMCA.

| Geospatial Theme | NMCA | | | | | | | Count |
|---|---|---|---|---|---|---|---|---|
| | NL | UK | IE | FI | FR | CH | NO | |
| Administrative Units | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| Postcode Units | | 1 | | | | | | 1 |
| Geographical Names | | | | 1 | | | | 1 |
| Place Names | | 1 | | 1 | | | 1 | 3 |
| Streets | | | | | | 1 | | 1 |
| Addresses | 1 | | | | | 1 | | 2 |
| Building Identifiers | 1 | | 1 | 1 | | | | 3 |
| Topography (small scale) | 1 | | | | | | | 1 |
| Topography (large-scale) | 1 | | | | | | | 1 |
| National Roads | | | | | | | 1 | 1 |
| Cadastre | 1 | | | | | | | 1 |
| Coordinates Reference System (CRS) | | | | | 1 | | | 1 |
| Public Transport Stops | | | | | | 1 | | 1 |
| **Summary** | **6** | **3** | **2** | **4** | **2** | **4** | **3** | **24** |

According to Table 7, the most frequently published themes are administrative units (7), followed by place names (3), building identifiers (3), and addresses (2). We conclude that NMCAs prioritise publishing topics that allow indirect referencing (query, navigation) for external users. We can suggest IGN France, which has recognised the need for and created a vocabulary to describe coordinate reference systems (CRS) in machine-readable form (van den Brink et al., 2018).

Table 8: Overview of addresses of NMCAs in the linked data portal

| state | NMCAs LD portals | Web address |
|---|---|---|
| NL | PLDN linked data environment | https://data.labs.kadaster.nl/kadaster/-/overview |
| UK | Ordnance Survey Linked Data Platform | https://data.ordnancesurvey.co.uk/ |
| IE | OSI's National Map (Prime2) data | https://osi.ie/blog/linked-data/ |
| FI | Linked Data Finland | https://www.ldf.fi/datasets.html |
| FR | Ontology of IGN administrative units | http://data.ign.fr/def/geofla/20210218.htm |
| CH | The Linked Data Service | https://ld.geo.admin.ch |
| NO | Kartverkets LOD | https://rdf.kartverket.no/api/1.0#/adminstrativeunit |

The most productive countries in terms of publishing LD are the Netherlands (administrative units, topography, addresses and buildings, large-scale topography, cadastre), Switzerland (administrative units, public transport stops, street directory, address directory, and others), and Finland (geographic names, place names, administrative units, buildings). On average, the NMCAs of seven selected countries have already published about three geospatial themes as LD.

### 5.4 RQ4: What challenges/issues have NMCAs identified for the future development of linked data?

The challenges NMCAs face in implementing geospatial LD publishing are multi-faceted, ranging from technological to policy-related issues. These challenges, stated in the extracted papers, include deciding on priorities for development, keeping pace with technological trends, adhering to EU and state policies on open data implementation, and compensating for the financial costs of data sharing. Additionally, the skills of staff working on local state projects or experts in NMCA and the availability of best practice guidance may pose challenges in implementing LD publishing.

Our research question about the challenges NMCAs concerning geospatial LD publication could be supported by the concepts expressed in the studies of the selected authors and others (Table 9).

Table 9: Challenges for NMCAs concerning geospatial LD publication

| Challenges for NMCAs concerning geospatial LD publication | Authors |
|---|---|
| LD endpoints connectivity | Mai et al., 2019 |
| Use of ontologies to derive appropriate GIS representations | Mai et al., 2019 |
| A multilingual environment of the EU and broader | Çağdaş and Stubkjær, 2015b |
| Generating ontologies from existing UML models | Ronzhin et al., 2019 |
| A scalable approach to reduce the time and cost associated with publishing | Rowland et al., 2022 |
| Automation is the process of transforming input geospatial data into linked data | Kyzirakos et al., 2018 |
| Creating thematic maps by combining attributes from the linked open data cloud with geometry stored in a spatial database server (SDS) | Owusu-Banahene, 2018 |
| Choosing between two approaches in dealing with geometries: GML, GeoJSON files or structured RDF-based linked data | van den Brink et al., 2018 |
| Provision of adequate data quality for metadata models | Beyza et al., 2020 |

The issues stated by the authors mentioned above are partially resolved in the selected studies, as most of them are initial and original investigations that attempt to address the challenges. Fewer studies offer stable production solutions that may still be limited to their local use case (Ma, 2017). They pose a resilient challenge to other NMCAs working in a different context.

### 5.5 RQ5: What are the reasons for the relatively slow adoption of geospatial LD in NMCAs?

The last research question, RQ 5, is the most general and therefore requires a closer look at the different aspects and suggestions for improving the dissemination of LD in NMCAs. Corresponding elements from reviewed studies are collected to support the argument for the slow proliferation of LD.

National Mapping and Cadastral Agencies are 'official' government bodies for which the state has determined the national interest to mandate a public authority to produce maps supporting a general purpose. Switching to an open data policy challenges the business model of National Mapping and Cadastral Agencies (NMCAs) (Folmer et al., 2020). Therefore, LD should be part of business discussions, including strategic vision and ambitions statements. The introduction of the Linked Data approach to the land administration domain may make land administration datasets more accessible via the Web (Çağdaş & Stubkjær, 2015a). NMCAs collect reliable, trustworthy, and well-maintained geospatial information,

but they are not the only public entities that must do so. They cooperate with other government bodies, such as statistics, state agencies, and municipalities.

Based on the well-established cooperation of public entities, we may expect further quantifiable cost savings and efficiency gains for the administration from using a national standards-based LD geospatial platform (Coumans, 2017). A collaborative framework as a platform will contribute to the interoperability and integration of different types of public sector and other geospatial information (Crompvoets et al., 2016). Understanding NMCA's data is complex and requires knowledge from at least two domains: geoinformation and land administration (Ronzhin et al., 2019).

Unfortunately, many data providers have abandoned support and maintenance of public SPARQL endpoints in recent years, thus harming the trust between consumers and open Linked Data providers. Due to the bureaucracy and complex legislation characterising public services, public bodies tend to be much slower in adopting new technologies, such as Linked Data, than private organisations and enterprises (Mouzakitis et al., 2017).

While linked data and semantic web technologies are not new, their uptake in various industries has been slow, perhaps due to limited tooling to support these endeavours (Rowland, Folmer, and Beek 2020). Contemporary Linked Open Data tooling is often unable to properly advertise datasets' usefulness to potential users, thereby hampering reuse (Folmer et al., 2019; A. Rowland et al., 2021; Rowland et al., 2022). The Data on the Web Best Practices give general guidelines for publishing data on the Web but no specific guidelines about spatial data (van den Brink et al., 2018). Over the last few years, Linked Open Data have seen a relatively slow adoption speed, partly due to the lack of sufficiently distinctive and functional browsing and viewing facilities. (Folmer et al., 2018).

The lack of knowledge on managing knowledge graphs and the link datasets is a limiting factor for the "five-star" deployment scheme (Tim Berners- Lee, 2010), also applicable to open geospatial data publication. And these are:

1. Make data available on the Web in whatever format.
2. Make data available as structured data (e.g., Excel).
3. Use non-proprietary formats (e.g., CSV instead of Excel format).
4. Use URIs to denote things so people can point at your data.
5. Link your data to other data to provide context.

Semantic Computing Research Group (SeCo), authors of the Linked Data Finland platform (Hyvönen et al., 2014), proposed to data publishers two extra stars:

6. Provide the dataset with explicit metadata schema that explains the dataset.
7. Validate the data against the schema for better quality.

This comprehensive approach enables the reuse of data to understand better the dataset's characteristics and evaluate the data quality for the intended purpose (Hyvönen et al., 2014). Therefore, bridging professionals' awareness about the extended reusability of data in a linked form for Web publication is critical for future research direction. In this context, it would be interesting to know how to share ownership and responsibility for maintaining a knowledge graph (Ronzhin et al., 2019). The GeoSPARQL benchmark can help implementers improve RDF storage solutions (Jovanovik et al., 2021).

Velitchkov (2021) raised the survey question "What are the barriers to the uptake of Linked Data?" to a small group of LD professionals and presented the findings at the ENDORSE (2021) Conference. The most interesting answers from LD professionals are that LD requires people to think in new ways, that LD is too academic, and that publishers do not use LD themselves. Additionally, LD fails to attract IT vendors due to the lack of professional demand.

In addition to technical solutions, several non-technical issues must be solved at the institutional, policy, legal, and social levels before operational deployment (Wiemann and Bernard, 2016). Therefore, the Framework for Effective Land Administration (FELA) was developed as an overarching policy guide (United Nations Expert Group on Land Administration and Management, 2019). There are six sets of governance instruments of national SDIs in Europe: collective decision-making structures, strategic management, allocation of tasks and responsibilities, creation of markets, inter-organisational culture, knowledge management, and regulation and formalisation of the infrastructure (Crompvoets and Vancauwenberghe, 2018), which are appropriate to be adopted directly in LAS, including NMCA's publishing LD. By answering all the research questions, we have provided a more structured overview of the issues, solutions, and challenges NMCAs face concerning geospatial LD publishing.

In summary, the adoption of LD by NMCAs requires a holistic approach that addresses both technical and non-technical challenges. By implementing appropriate guidelines, tooling, and expertise, NMCAs can ensure that their LD publishing efforts are successful and meet the needs of their stakeholders. Similarly, addressing institutional, policy, legal, and social issues can ensure that LD publishing efforts are aligned with national priorities and can be sustained over the long term.



## 6 CONCLUSIONS

In the present literature review, we searched for the current status of LD use in NMCAs and the challenges faced by NMCAs. The systematic review resulted in the final nineteen selected studies for full-text information extraction. If all the initially identified 132 studies were included in detailed data extraction, it would significantly affect the quality of the presented work. The reasons for the non-inclusion of specific documents lie in the systematic literature review methodology, which aims at the exhaustive study of the limited number of eligible articles due to the otherwise high workload. With the support of the review protocol, we assessed the quality of the studies using six criteria for study selection as eligible for answering research questions.

Answers to the five research questions of the present research are summarised as follows.

Answers to RQ1: Problems

Several authors have pointed out the issue of heterogeneity, context, metadata and siloed data as the main problems in publishing NMCA geospatial data as LD (Çağdaş and Stubkjær, 2015a; Debruyne et al., 2016; Huang et al., 2019; Huerta, Schade, Granell, 2014; Mai et al., 2019; Ronzhin et al., 2019; van den Brink et al., 2018; Vilches-Blázquez & Saavedra, 2019).

Answers to RQ2: Solutions and requirements

Based on the analyses of the nineteen papers, we identified six essential solutions and their combination for the successful publication of LD in the NMCAs listed: vocabularies, ontologies, metadata, and solu-

tions using standardised technologies that meet requirements such as readability and serialisation. In the LD development projects, it is essential to provide systematic support to creators and users developing vocabularies and ontologies and the scalability of approaches to deal with large datasets.

Answers to RQ3: Priority data themes

On average, the NMCAs of seven considered countries have already published about three geospatial themes as LD. The most frequently published themes are administrative units (7), followed by place names (3), building identifiers (3), and addresses (2).

Answers to RQ4: Challenges

As stated in the analysed papers, challenges include setting development priorities, keeping pace with technological trends, adhering to EU and state policies on open data implementation, and compensating for the financial costs of data sharing. Additionally, the skills of staff working on local state projects or experts in NMCA and the availability of best practice guidance may pose challenges in implementing LD publishing.

Answers to RQ5: Reasons for the slow proliferation of LD

Public bodies tend to adopt new technologies, such as Linked Data, much slower than private organisations and enterprises (Mouzakitis et al., 2017). While linked data and semantic web technologies are not new, their uptake in various industries has been slow, perhaps due to limited tooling to support these endeavours (Rowland, Folmer, and Beek 2020). The lack of knowledge on managing knowledge graphs and the link datasets limits the "five-star" deployment scheme (Tim Berners- Lee, 2010), which also applies to open geospatial data publication. Contemporary Linked Open Data tooling is often unable to properly advertise datasets' usefulness to potential users, thereby hampering reuse (Folmer et al., 2019; A. Rowland et al., 2021; Rowland et al., 2022). Over the last few years, Linked Open Data have seen a relatively slow adoption speed, partly due to the lack of sufficiently distinctive and functional browsing and viewing facilities. (Folmer et al., 2018). Velitchkov (2021) assumes that publishers do not use LD themselves, and LD fails to attract IT vendors due to the lack of professional demand.

While the review did not address legal and complex technical issues, the authors plan to explore LD more detail to understand which combinations of approaches, protocols, and standard semantic software technologies are most effective for publishing geospatial data in NMCAs in online ecosystems.

It is important to note that the conclusions drawn from the systematic literature review may be limited by the outdated information in some of the analysed articles. Furthermore, the study did not address technologies, standards, processes, or best practice guidance for LD publication at NMCAs, which will be the subject of our future investigations.

**Acknowledgements**

The Slovenian Research Agency and the National Mapping and Cadastral Agency (NMCA) of the Republic of Slovenia (RS) fund this research under project CRP V2-2295.

## Literature and references:


Ariza-López, F. J., Rodríguez-Pascual, A., Lopez-Pellicer, F. J., Vilches-Blázquez, L. M., Villar-Iglesias, A., Masó, J., Díaz-Díaz, E., Ureña-Cámara, M. A., González-Yanes, A. (2021). An Analysis of Existing Production Frameworks for Statistical and Geographic Information: Synergies, Gaps and Integration. ISPRS International Journal of Geo-Information, 10(6), 374. DOI: https://doi.org/10.3390/ijgi10060374

Attard, J., Orlandi, F., Auer, S. (2017). ExConQuer: Lowering barriers to RDF and Linked Data reuse. https://www.researchgate.net/publication/312889722_ExConQuer_Lowering_barriers_to_RDF_and_Linked_Data_re-use, accessed 20-May-23.

Beek, W., Folmer, E. (2017). AN INTEGRATED APPROACH FOR LINKED DATA BROWSING. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W2, 35–38. DOI: https://doi.org/10.5194/isprs-archives-XLII-4-W2-35-2017

Berners-Lee, T. (1996). The world wide web - Past, present and future. Journal of Digital Information, 1(1).

Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., Shakibazadeh, E. (2019). Formulating questions to explore complex interventions within qualitative evidence synthesis. BMJ Global Health, 4(Suppl 1), e001107. DOI: https://doi.org/10.1136/bmjgh-2018-001107

Bucher, B., Folmer, Erwin, Brennan, R., Beek, Wouter, Hbeich, E., Würriehausen, F., Rowland, L., Maturana, R. A., Alvarado, E., Buyle, R., Di Donato, P. (2020). Spatial Linked Data in Europe: Report from Spatial Linked Data Sessions at Knowledge Graph in Action. https://www.eurosdr.net/sites/default/files/uploaded_files/eurosdr_publication_ndeg_73.pdf, accessed 16-May-23.

Çağdaş, V., Stubkjær, E. (2015a). Core immovable property vocabulary for European linked land administration. Survey Review, 47(340), 49–60. DOI: https://doi.org/10.1179/1752270614Y.0000000093

Çağdaş, V., Stubkjær, E. (2015b). A SKOS vocabulary for Linked Land Administration: Cadastre and Land Administration Thesaurus. Land Use Policy, 49, 668–679. DOI: https://doi.org/10.1016/j.landusepol.2014.12.017

Coumans, F. (2017). National Mapping as a Service: Ordnance Survey Ireland Shows New Perspectives. https://www.gim-international.com/content/article/national-mapping-as-a-service, accessed 16.3.2023.

Crompvoets, J., Streilein, A., Masser, I. (2016). How should NMCAs adapt to alternative sources for NMCA data? (European Spatial Data Research No. Official Publication No 65).

Crompvoets, J., Vancauwenberghe, G. (2018). Governance of national spatial data infrastructures in Europe. International Journal of Spatial Data Infrastructure Research, 253–285. DOI: https://doi.org/10.2902/1725-0463.2018.13.art16

Cruzes, D. S., Dybå, T. (2010). Synthesising evidence in software engineering research. In International Symposium on Empirical Software Engineering and Measurement.

Debruyne, C., Clinton, É., Mcnerney, L., Nautiyal, A., O'Sullivan, D. (2016). Serving Ireland's Geospatial Information as Linked Data.

The Directive on open data and the reuse of public sector data (European Commission 2019).

Folmer, E., Beek, W., Rietveld, L., Namangaya A.H., B. M. (2018). Linked data viewing as part of the spatial data platform of the future. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42(4W8). DOI: https://doi.org/10.5194/isprs-archives-XLII-4-W8-49-2018

Folmer, E., Beek, W., Rietveld, L., Ronzhin, S., Geerling, R., Haan, D. den, Bui T.X. (2019). Enhancing the usefulness of open governmental data with linked data viewing techniques. Proceedings of the Annual Hawaii International Conference on System Sciences, 2019-January. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85083956109&partnerID=40&md5=de9a48645a167b7040a65824bc343194.

Folmer, E., Ronzhin, S., van Hillegersberg, J., Beek, W., Lemmens, R. (2020). Business rationale for linked data at governments: A case study at the Netherlands kadaster data platform. IEEE Access, 8, 70822–70835. DOI: https://doi.org/10.1109/ACCESS.2020.2984691

Goodwin, J., Dolbear, C., Hart, G. (2008). Geographical Linked Data: The Administrative Geography of Great Britain on the Semantic Web. Transactions in GIS, 12, 19–30. DOI: https://doi.org/10.1111/j.1467-9671.2008.01133.x

Hietanen, E., Lehto, L., Latvala, P. (2016). PROVIDING GEOGRAPHIC DATASETS AS LINKED DATA IN SDI. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B2, 583–586. DOI: https://doi.org/10.5194/isprs-archives-XLI-B2-583-2016

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., Welch, V. A. (2022). Cochrane Handbook for Systematic Reviews of Interventions. https://training.cochrane.org/handbook, accessed 20-May-23.

Huang, W., Raza, S. A., Mirzov, O., Harrie, L. (2019). Assessment and Benchmarking of Spatially Enabled RDF Stores for the Next Generation of Spatial Data Infrastructure. ISPRS International Journal of Geo-Information, 8(7), 310. DOI: https://doi.org/10.3390/ijgi8070310

Huerta, Schade, Granell (Ed.) 2014. Cadastral data integration through Linked Data agile2014_153, accessed 17.3.2023.

Hyvönen, E., Tuominen, J., Alonen, M., Mäkelä, E. (2014). Linked Data Finland: A 7-star Model and Platform for Publishing and Reusing Linked Datasets. In V. Presutti, E. Blomqvist, R. Troncy, H. Sack, I. Papadakis, & A. Tordai (Eds.), Lecture Notes in Computer Science. Semantic Web: ESWC 2014 Satellite Events (Vol. 8798, pp. 226–230). Cham: Springer Nature. DOI: https://doi.org/10.1007/978-3-319-11955-7_24

Infrastructure for Spatial Information in the European Community (INSPIRE). Directive 2007/2/EC of the European Parliament and the Council. (2007). https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32007L0002, accessed 23-May-23.

Jalali, S., Wohlin, C. (2012). Systematic literature studies: Database searches vs. backward snowballing. In Proceedings of the 2012 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 29–38). DOI: https://doi.org/10.1145/2372251.2372257

Jovanovik, M., Homburg, T., Spasić, M. (2021). A GeoSPARQL Compliance Benchmark. ISPRS International Journal of Geo-Information, 10(7), 487. DOI: https://doi.org/10.3390/ijgi10070487

Kitchenham, B. (2004). Procedures for Performing Systematic Reviews. https://www.semanticscholar.org/paper/Procedures-for-Performing-Systematic-Reviews-Kitchenham/29890a936639862f45cb9a987dd599dce9759bf5.

Lanucara, S., Praticò, S., Modica, G. (2019). Harmonisation and Interoperable Sharing of Multi-temporal Geospatial Data of Rural Landscapes. In F. Calabrò, L. Della Spina, & C. Bevilacqua (Eds.), Smart Innovation, Systems and Technologies: Vol. 100. New Metropolitan Perspectives: Local Knowledge and Innovation Dynamics Towards Territory Attractiveness Through the Implementation of Horizon/E2020/Agenda2030 - Volume 1 (Vol. 100, pp. 51–59). Cham: Springer International Publishing; Imprint: Springer. DOI: https://doi.org/10.1007/978-3-319-92099-3_7

Ma, X. (2017). Linked Geoscience Data in practice: where W3C standards meet domain knowledge, data visualisation and OGC standards. Earth Science Informatics, 10(4), 429–441. DOI: https://doi.org/10.1007/s12145-017-0304-8

Mai, G., Janowicz, K., Yan, B., Scheider, S. (2019). Deeply integrating Linked Data with Geographic Information Systems. Transactions in GIS, 23(3), 579–600. DOI: https://doi.org/10.1111/tgis.12538

Mouzakitis, S., Papaspyros, D., Petychakis, M., Koussouris, S., Zafeiropoulos, A., Fotopoulou, E., Farid, L., Orlandi, Fabrizio, Attard, Judie, Psarras, J. (2017). Challenges and opportunities in renovating public sector information by enabling linked data and analytics. Information Systems Frontiers, 19(2), 321–336. DOI: https://doi.org/10.1007/s10796-016-9687-1

Nguyen, T., Jump, A., Casey, D. (2022). Emerging Tech Impact Radar: 2023: Gartner Research Excerpt. https://emtemp.gcom.cloud/ngw/globalassets/en/doc/documents/767259-emerging-tech-impact-radar-2023-excerpt.pdf, accessed 10-May-23.

O'Leary, D. E. (2008). Gartner's hype cycle and information system research issues. International Journal of Accounting Information Systems, 9(4), 240–252. DOI: https://doi.org/10.1016/j.accinf.2008.09.001

Octaviano, F., Felizardo, K., Maldonado, J., Fabbri, S. (2015). Semi-automatic selection of primary studies in systematic literature reviews: is it reasonable? Empirical Software Engineering, 20, 1898–1917.

Owusu-Banahene, W. (2018). Making attributes from the Linked Open Data (LOD) cloud a part of Spatial Data Infrastructures (SDIS) for thematic mapping. AFRREV STECH: an International Journal of Science and Technology, 7(1), 20–33. DOI: https://doi.org/10.4314/stech.v7i1.3

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, Tianjing, Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, James, Tricco, A. C., Welch, V. A., Whiting, P., Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical Research Ed.), 372, n71. DOI: https://doi.org/10.1136/bmj.n71

Person, T., Ruiz-Rube, I., Mota, J. M., Cobo, M. J., Tselykh, A., Dodero, J. M. (2021). A Comprehensive Framework to Reinforce Evidence Synthesis Features in Cloud-Based Systematic Review Tools. Applied Sciences, 11(12), 5527. DOI: https://doi.org/10.3390/app11125527

Polleres, A., Kamdar, M. R., Fernández, J. D., Tudorache, T., Musen, M. A. (2020). A more decentralised vision for Linked Data. Semantic Web, 11(1), 101–113. DOI: https://doi.org/10.3233/SW-190380

Richardson, W. S., Wilson, M. C., Nishikawa, J., Hayward, R. S. (1995). The well-built clinical question: a key to evidence-based decisions. ACP Journal Club, 123(3), A12. DOI: https://doi.org/10.7326/ACPJC-1995-123-3-A12

Ronzhin, S., Folmer, E., Maria, P., Brattinga, M., Beek, W., Lemmens, R., van't Veer, R. (2019). Kadaster knowledge graph: Beyond the fifth star of open data. Information (Switzerland), 10(10). DOI: https://doi.org/10.3390/info10100310

Rowland, A., Folmer, E., Beek, W. (2020). Towards Self-Service GIS—Combining the Best of the Semantic Web and Web GIS. ISPRS International Journal of Geo-Information, 9(12), 753. DOI: https://doi.org/10.3390/ijgi9120753

Rowland, A., Overeem, J., Folmer, E., & Yaman B., Sherif M.A., Ngomo A.-C.N., Haller A. (Eds.) 2021. Demo: GeoDataWizard for linked spatial data creation. : Vol. 2977: CEUR-WS. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117757990&partnerID=40&md5=7acbf3e48131d15b8c0bb13785e6ded9.

Rowland, Alexandra, Folmer, Erwin, Beek, Wouter, Wenneker, R. (2022). Interoperability and Integration: An Updated Approach to Linked Data Publication at the Dutch Land Registry. ISPRS International Journal of Geo-Information, 11(1), 51. DOI: https://doi.org/10.3390/ijgi11010051

Saavedra, J., Vilches-Blázquez, L. M., & Boada, A. (Eds.) 2014. Cadastral data integration through Linked Data.

Schleidt, K., O'Grady, M., Grellet, S., Feliachi, A., van der Schaaf, H. (2020). ELFIE - The OGC Environmental Linked Features Interoperability Experiment. In Birukou & Athanasiadis (Eds.), IFIP Advances in Information and Communication Technology. Environmental Software Systems. Data Science in Action (Vol. 554, pp. 188–193). Cham: Springer International Publishing. DOI: https://doi.org/10.1007/978-3-030-39815-6_18

Schmidt, L., Olorisade, B. K., McGuinness, L. A., Thomas, James, Higgins, Julian. (2021). Data extraction methods for systematic review (semi)automation: A living systematic review. F1000Research, 10, 401. DOI: https://doi.org/10.12688/f1000research.51117.1

Tim Berners- Lee. (2010). 5-star Open Data. https://5stardata.info/en/, accessed 15-May-23.

United Nations Expert Group on Land Administration, Management. (2019). Framework for Effective Land Administration: A reference for developing, reforming, renewing, strengthening or modernising land administration and management systems.

Van den Brink, L., Barnaghi, P., Tandy, J., Atemezing, G., Atkinson, R., Cochrane, B., Fathy, Y., García Castro, R., Haller, A., Harth, A., Janowicz, K., Kolozali, Ş., van Leeuwen, B., Lefrançois, M., Lieberman, J., Perego, A., Le-Phuoc, D., Roberts, B., Taylor, K., Troncy, Raphäel. (2018). Best practices for publishing, retrieving, and using spatial data on the Web. Semantic Web, 10(1), 95–114. DOI: https://doi.org/10.3233/SW-180305

Velitchkov, I. (2021). Endorse 5 star deployment scheme for Linked Data: Barriers to the uptake of Linked Data. Endorse Conference https://op.europa.eu/en/web/endorse-2021/programme. https://op.europa.eu/documents/10120270/10133951/VELITCHKOV_presentation_Endorse+5+stars+deployment+scheme+for+Linked+Data.pdf/74ecc573-d874-37f3-f133-13968b1cc4eb?t=1616618192854, accessed 22-May-23.

Vilches-Blázquez, L. M., Saavedra, J. (2019). A framework for connecting two interoperability universes: OGC Web Feature Services and Linked Data. Transactions in GIS, 23(1), 22–47. DOI: https://doi.org/10.1111/tgis.12496

Wiemann, S., Bernard, L. (2016). Spatial data fusion in Spatial Data Infrastructures using Linked Data. International Journal of Geographical Information Science, 30(4), 613–636. DOI: https://doi.org/10.1080/13658816.2015.1084420

Wohlin, C., Kalinowski, M., Romero Felizardo, K., Mendes, E. (2022). Successful combination of database search and snowballing for identification of primary studies in systematic literature studies. Information and Software Technology, 147, 106908. DOI: https://doi.org/10.1016/j.infsof.2022.106908





Čeh M., Tekavec J. (2023). Geospatial Linked Data Proliferation in NMCAs: Systematic Literature Review. Geodetski vestnik, 67 (2), 212-234. DOI: https://doi.org/10.15292/geodetski-vestnik.2023.02.213-234

***doc. dr. Marjan Čeh***
*UL FGG, Jamova cesta 2, Ljubljana, Slovenija*
*e-naslov: marjan.ceh@fgg.uni-lj.si*

***asist. dr. Jernej Tekavec***
*UL FGG, Jamova cesta 2, Ljubljana, Slovenija*
*e-naslov: jernej.tekavec@fgg.uni-lj.si*