A graph-based representation of knowledge for managing land administration data from distributed agencies – A case study of Colombia

Luis M. Vilches-Blázquez and Jhonny Saavedra

ABSTRACT
Multiple efforts have been performed worldwide around diverse aspects of land administration. However, land administration data and systems’ notorious heterogeneity remains a longstanding challenge to develop a harmonized vision. In this sense, the traditional Spatial Data Infrastructures adoption is not enough to overcome this challenge since data sources’ heterogeneity implies needs related to harmonization interoperability, sharing, and integration in land administration development. This paper proposes a graph-based representation of knowledge for integrating multiple and heterogeneous data sources (tables, shapefiles, geodatabases, and WFS services) belonging to two Colombian agencies within a decentralized land administration scenario. These knowledge graphs are developed on an ontology-based knowledge representation using national and international standards for land administration. Our approach aims to prevent data isolation, enable cross-datasets integration, accomplish machine-processable data, and facilitate the reuse and exploitation of multi-jurisdictional datasets in a single approach. A real case study demonstrates the applicability of the land administration data cycle deployed.

1. Introduction

Land administration is described as the process of recording and disseminating information about the ownership, use, and value of land and its related resources and incorporates restrictions and responsibilities associated with land rights, use and value, and impact of development processes (UNECE 1996). In this sense, it is undoubted that land administration is the cornerstone for securing tenure, taxation, valuation, land/resources management, and spatial planning (van Oosterom and Lemmen 2015).

Traditionally, land administration information was registered in hard-copy formats. Geographic Information Systems (GIS) advancement initiated a digitalization process and made this information available in digital form. Later on, Williamson et al. (2010) argued that effective and efficient Land Administration Systems required Spatial Data Infrastructures (SDI) to progress. Supporting this assertion, we have witnessed SDIs presence in almost every country around the world over the last 20 years. These initiatives have combined systems, data sources, standards, network linkages, and institutional issues to provide geospatial information from different sources to the broadest audience (Giuliani et al. 2017). In this scenario, SDIs have been built for land and property context as one of their principal domains and some of the most extensively used (Borzacchiello and Craglia 2013).

Additionally, several international bodies have worked to create a standardized model for land administration. The International Federation of Surveyors (FIG) set the basis for a future cadastral system in (FIG 1998). Other efforts to create standardized models and facilitate cadastral data exchange were the Cadastral Data Content Standard for the National Spatial Data Infrastructure (FGDC 2008a), the Geographic Information Framework Data Standard Part 1, Cadastral (FGDC 2008b), the INSPIRE Data Specification on Cadastral Parcels (INSPIRE 2014), or the ISO standard 19152 Land Administration Domain Model (LADM) (ISO 2012).

Despite the aforementioned efforts, a longstanding challenge for developing a harmonized vision of land administration remains. It is associated with the notorious heterogeneity of land administration data and systems (Psyllidis et al. 2015). This heterogeneity causes data to differ dramatically, mainly when obtained from various providers or custodians, although it may appear even within the same domain’s datasets. These issues are associated with many causes, such as changing definitions, historical legacies, standards across jurisdictions, etc. (Chen et al. 2018). Therefore, the traditional SDI adoption is not enough to overcome this challenge.
since heterogeneity of multiple data sources implies needs related to harmonization, interoperability, sharing, and integration in land administration development (Rajabifard, Ho, and Soheil 2016; Lemmen et al. 2018). Current SDI approaches have limitations to accomplish these requirements (Tandy, van den Brink, and Barnaghi 2017).

Over the last decade, Semantic Web technologies have been increasingly utilized to deal with some longstanding issues in the geospatial domain (Huang and Harrie 2019). More recently, knowledge graphs have appeared as an extension of Semantic Web practices and are embraced by diverse companies such as Google, IBM, Facebook, or Microsoft (Noy et al. 2019). Knowledge graphs promote the creation, reuse, and recovery of human and machine-readable structured data about real-world objects using a graph-based representation (Paulheim 2017). In this way, knowledge graphs have become one of the principal ways to integrate diverse data (Cudré-Mauroux 2020) so that multiple heterogeneous datasets can be handled and interlinked in a single system (Krötzsch and Thost 2016; Bellomarini, Sallinger, and Vahdati 2020). In order to achieve integration, ontologies have been applied for the explication of hidden and implicit knowledge to overcome semantic heterogeneity problems (Wache et al. 2001) since they allow modeling semantic relationships between distinct structures and forming an integrated and coherent view of multiple and heterogeneous datasets (Krötzsch and Thost 2016).

Several works have addressed land or cadastral information from a semantic perspective (Hess and de Vries 2006; Li et al. 2012; Çağdaş and Stubkjær 2015a; Sladić et al. 2015; Çağdaş and Stubkjær 2015b; Shi and Roman 2018). Moreover, some works have transformed land or cadastral information to Linked Data (Saavedra, Vilches-Blázquez, and Boada 2014; Díaz and Vilches-Blázquez 2014; Shi et al. 2017; Folmer, Beek, and Rietveld 2018) and even The Netherlands’ Kadaster built an initial experience of knowledge graphs (Ronzhin et al. 2019). However, these previous works have not considered integrating land information associated with decentralized and multi-jurisdictional land administration agencies in the same country. Therefore, the mentioned approaches have not handled a datasets scenario with notorious differences due to different viewpoints, systems, historical legacies, etc. Furthermore, even though there exists an earlier experience of knowledge graphs in this area, there are no previous works where land administration knowledge graphs have been developed using data collected from tables, shapefiles, geodatabases, and Web Feature Services (WFS).

The main challenge of this work is to deal with heterogeneous data sources to prevent data isolation, enable cross-datasets integration, and accomplish (semantically) machine-processable data. Thereby, this work proposes a graph-based representation of knowledge for integrating multiple and heterogeneous data sources of land administration. These knowledge graphs are constructed on an ontology-based knowledge representation, which improves the existing LADM ontology model (Soon 2013). Furthermore, this article provides a real case study where land administration information from two Colombian agencies is semantically integrated by knowledge graphs and enriched with other data according to Linked Data principles. In this way, our work allows integrating land administration information, preserving the semantics of diverse data sources, and facilitating the reuse and management of this information.

This paper is structured as follows. The following section provides a description of the Colombian land administration’s context as our application scenario. Section 3 outlines the background and related work. The proposed ontology-based knowledge representation for our land administration scenario is described in Section 4. Section 5 presents the deployed process to generate the knowledge graphs of our case study. Finally, discussion and conclusions are provided in Section 6.

2. The Colombian land administration’s context

In many countries, land administration’s responsibilities and tasks are scattered among different branches of governmental agencies. Sometimes those organizations deal with other administrative territories, all of which may have subdivisions again: central, regional, or local responsibilities, with public or private functions. It entails that the dataset’s quality and governance aspects vary (van Oosterom and Lemmen 2015).

This distributed scenario is given in Colombia, whose cadasters have been most developed in Latin America. In this context, there are different governmental organizations, such as the Colombian National Geographic Institute (Instituto Geographic Agustín Codazzi – IGAC), which manages the National cadastral information, and five additional cadastral agencies, which regulate land administration in a decentralized and multi-jurisdictional way in the cities of Medellin, Bogota, Cali, Barranquilla, and the Antioquia Department.

This decentralized management of land administration systems has entailed diverse views, interpretations, uses, and applications of the gathered information. These facts involve the appearance of different and heterogeneous models, vocabularies, and production and management systems, appearing even within different sections in the same agency. IGAC built a Cadastral National System to consolidate a unique cadastral model for the National data to
overcome these drawbacks. This work was used to strengthen different existing models within the National organization and structure associated data in a single national database. Additionally, this proposal introduced a cadastral code for sharing data with the decentralized cadasters. Despite this effort, different agencies carried out land administration management in different ways in Colombia. Therefore, relevant obstacles exist for an integrated view of Colombian land administration since governmental agencies and their datasets remain isolated silos.

In order to foster data interoperability, some of these agencies have developed SDI initiatives, where several datasets are available as a diverse type of OGC (Open Geospatial Consortium) web services, highlighting the publication of Web Map Services (WMS) and WFS. Among these cases, we find different initiatives at the local level, such as Bogota SDI (IDECA – Infraestructura de Datos Espaciales para el Distrito Capital), Cali SDI (IDESC – Infraestructura de Datos Espaciales de Santiago de Cali) or GeoMedellin, and the Colombian National SDI (ICDE – Infraestructura Colombiana de Datos Espaciales) at the National level.

More recently, a new effort has been launched in the country supported by The World Bank (World Bank 2019). This project aims to build a multipurpose cadaster in chosen municipalities to strengthen tenure and security and access cadaster information. The project incorporates several components, standing out those associated with the development of land administration systems able to offer nationwide land administration services effectively and permanently, land administration decentralization by delegating the generation and updating processes of these data, and strengthening the National Colombian SDI. Further details associated with this project can be found in (World Bank 2019).

In this context, one of the main achievements has been the National profile of LADM, from which some works (Guarin et al. 2017; Jenni et al. 2017; Rodriguez, Páez, and Rajabifard 2017; Palicot and Daniel 2019) have been developed. Nevertheless, although these proposals promise inspiring results, they suggest relevant changes in supported models and processes for each decentralized agency.

Even although some of the mentioned initiatives have allowed increasing the (syntactic) interoperability level of the available information, we identified that the Colombian land administration continues supporting layers overlapping as the data integration method, without reaching a (semantic) interoperability at the data level (Vilches-Blázquez et al. 2014). Hence, these initiatives consolidate data silos, despite changes performed with the SDI initiatives. Thus, we present a proposal for a graph-based representation of knowledge for managing land administration in the Colombian context. This work allows resolving heterogeneity problems between different governmental agencies by means of an ontology network and performing (semantic) data integration without modifying the current process, models, and vocabularies used by agencies. In doing so, we deal with heterogeneous datasets belonging to the national (IGAC) and local (Bogota cadastral) agencies to present different issues related to our case study, keeping the provenance of the various sources and producers. Finally, we show the results can be managed, queried, and exploited in an integrated way providing several examples.

3. Background and literature review

This section presents some background information about ontologies and knowledge graphs and provides a description of the main proposals related to our work.

3.1. Background

3.1.1. Ontologies

A widespread definition of ontology is described as a formal, explicit specification of a shared conceptualization (Gruber 1993). From this definition, “formal” indicates that ontology is machine-understandable. “Explicit” implies that concepts and their constraints are explicitly defined. “Shared” means that consensual knowledge is specified in the ontology, so a community accepts it, whereas “Conceptualization” reveals an abstract model of relevant concepts.

According to (Suárez-Figueroa et al. 2012), there exist three different possibilities when ontologies are created:

- A single ontology is an ontology that does not present any relationship (domain-independent or dependent) with other ontologies.
- A set of interconnected single ontologies. It comprises a set of ontologies with some sort of domain-dependent relation.
- An ontology network is a set of ontologies related together via a variety of relationships, such as modularization, alignment, dependency, and version.

Ontologies can specify the information sources’ semantics and make the contents explicit in the information integration context (Wache et al. 2001). According to Wache et al. (2001), it is often performed in one of the three following manners:

- Single ontology approaches. These approaches use one global ontology providing a shared vocabulary to specify the semantics associated with the information source.
- Multiple ontologies. Each information source is represented by its ontology.
3.1.2. Knowledge graphs

The concept Knowledge Graph was created by Google when they presented their vision about a new Web search strategy in 2012. It entails changing from pure text processing to a more symbolic description of knowledge, described in the following way:

"The Knowledge Graph enables you to search for things, people or places that Google knows about—landmarks, celebrities, cities, sports teams, buildings, geographical features, movies, celestial objects, works of art, and more—and instantly get information that’s relevant to your query. This is a critical first step towards building the next generation of search, which taps into the collective intelligence of the web and understands the world a bit more like people do."

According to Paulheim (2017), from a broad perspective, any graph-based representation of some knowledge could be viewed as a knowledge graph since there exists no standard definition about what a knowledge graph is and what it is not (Ehrleringer and Wöß 2016). However, we adopt the vision pushed by Semantic Web in this work, where the graph-based representation is associated with RDF, the standard knowledge representation language for this Web. RDF provides a reliable infrastructure to share, publish, and query structured data on the Semantic Web (McDonald and Levine-Clark 2017). In this sense, a simple triple of a graph could be understood as two nodes describing real-world entities connected by a relation. These triples (nodes and relations) allow diverse datasets to interlink, leading to data in graph-based representations, so-called knowledge graphs (Bellomarini, Sallinger, and Vahdati 2020).

There exist several examples of knowledge graphs in the Semantic Web that are frequently available using the Linked Data principles (Berners-Lee 2006): (1) use of URIs as names for things; (2) use of HTTP URIs so that people can look up those names; (3) when someone looks up a URI, provide useful information through standards RDF and SPARQL Protocol and RDF Query Language (SPARQL); and (4) include links to other URIs so that they can discover more things. Some of these knowledge graphs are Cyc, OpenCyc, Wikidata, DBpedia, etc. A detailed description of these knowledge graphs and other ones are collected in (Paulheim 2017).

3.2. Literature review

One of the most critical issues in the land administration domain is related to (cadastral) models. In this sense, various international efforts have provided standardized proposals for land modeling. Thereby, the International Federation of Surveyors (FIG) developed Cadastre 2014 (FIG 1998), a proposal that became the stimulus of the modern cadasters by means of the technology deployment in cadastral systems services. The Federal Geographic Data Committee (FGDC) made another relevant effort with the Cadastral Data Content Standard for the National Spatial Data Infrastructure (FGDC 2008a). This initiative provided semantic definitions of objects related to land ownership, land records, and land surveying information to contribute to data sharing. The FGDC also built the Geographic Information Framework Data Standard Part 1 – Cadastral (FGDC 2008b) in association with the previous standard. This one provided information to identify the presence and source of parcel-level cadastral information.

More recent efforts to facilitate cadaster data exchange are INSPIRE Data Specification on Cadastral Parcels (INSPIRE 2014) and the ISO standard 19152 Land Administration Domain Model (LADM) (ISO 2012). The latter provides terminology and models for land administration and enables combining land administration information from different sources coherently. A comprehensive explanation of the existing modeling approaches for land administration systems is collected in (Ponsard and Touzani 2019).

With respect to data integration, it has been treated in the context of databases (Doan, Halevy, and Ives 2012). Thereby, a vast amount of integration frameworks have been developed (Golshan et al. 2017), which implement data integration systems following three-fold: the local-as-view (LAV), global-as-view paradigms (GAV), and global and local as view (GLAV) (Lenzerini 2002). Different data integration surveys have been described with a general viewpoint (Wache et al. 2001; Noy 2004), and diverse works have focused on the geospatial domain to deal with semantic heterogeneity challenges (Stuckenschmidt et al. 1999; Hakimpour and Timpf 2001; Fonseca et al. 2002; Lutz et al. 2009; Jung, Sun, and Yuan 2013; Vilches-Blázquez et al. 2014; van den Brink et al. 2017; Chen et al. 2018; Huang and Harrie 2019).

(Geospatial) ontologies play an essential function in representing domain knowledge and enabling knowledge sharing and reusing as well as promoting
interoperability (Siricharoen and Pakdeetrakulwong 2014). As a result, some approaches deal with land (cadastre) information from a semantic perspective. In (Hess and de Vries 2006), the authors presented a proposal to query translation based on a core cadastral model. The prototype performed query translation based on semantic relations between the core model and Greek and Dutch national cadastral systems. Li et al. (2012) described a cadastral domain model oriented at unified real estate registration of China from legal and spatial perspectives. This ontology-based model considered the legal and core cadastral domain model proposed by (van Oosterom et al. 2006) and was validated by a prototype system. Soon (2013) defined a domain ontology for Land Administration from definitions included in the ISO 19152 standard. In (Çağdaş and Stubkjær 2015a), the authors showed a thesaurus for the cadaster and land administration domain in SKOS format. This proposal was developed with the guidelines of the ANSI/NISO and collected terms from ISO 19152, reports of the Property Formation in the Nordic countries project, GEMET, STW, and AGROVOC thesaurus, and INSPIRE Spatial Data themes. Sladic et al. (2015) described an ontological model divided into four levels (upper-level, geospatial feature ontology, core ontology for cadaster based on ISO 19152, and domain ontology for the Serbian Cadaster). This ontology was tested using data from the Serbian real estate cadaster. Çağdaş and Stubkjær (2015b) developed a core immovable property vocabulary as an extension to the e-Government Core Vocabularies of the European Commission. This approach captured some characteristics of data specifications and standards related to immovable property units and their core attributes for land administration processes. Besides, the authors performed a publication of land administration data sets using RDF. Shi and Roman (2018) surveyed ontologies for real property, including geospatial standards for representing concepts and attributes relevant to real properties, LADM, and aspects related to real property transactions and data integration.

Regarding Linked Data initiatives, there are a few works that have addressed cadastral information from this perspective. Saavedra, Vilches-Blázquez, and Boada (2014) showed an implementation of Linked Data principles in the cadastral domain using the LADM standard (ISO 19152) and GeoSPARQL. In (Díaz and Vilches-Blázquez 2014), the authors built an ontology network using INSPIRE Cadastral Specifications and LADM. This ontology was used to publish Linked Data employing Spanish cadastral data in a use case. Shi et al. (2017) explained details associated with the Linked Data publication of state-owned real estates in Norway. In (Folmer, Beek, and Rietveld 2018), the authors described two viewing components (Data Stories and FacetCheck) for Linked Open Data provided by the Land Registry and Mapping Agency of the Netherlands. In addition, an initiative related to land property, called “The Land Register of UK”13 published transactions and prices as Linked Data. Finally, Ronzhin et al. (2019) presented an initial experience constructing the first open government knowledge graph in The Netherlands. The authors proved the value of using such a graph in data browsing, the development of location-aware chatbots, and multicriteria analysis for urban planning.

As identified in most of the literature mentioned above, several proposals have developed ontologies to model semantically land or cadastral information. However, these approaches have not developed an ontology network with a hybrid approach to integrate several decentralized land administration management viewpoints within the same country. Moreover, although there exist some works publishing Linked Data, to the best of our knowledge, there are no previous works where the knowledge graphs generation of land information is driven by data sources associated with tables, shapefiles, geodatabases, and WFS services from various agencies and at different jurisdictional (national and local) levels.

4. The proposed ontology-based knowledge representation for land administration

Ontologies are considered as formal models of how a domain is perceived and present a logical and accurate description of the intended meaning of terms (including concepts, relations, and attributes in a knowledge domain), data structures, and other elements abstractly describing the real world (Hakimi et al. 2020), and offer a universal and stable schema for data harmonization (Chen et al. 2018). Therefore, considering that each Colombian land administration agency’s data sources present different jurisdictions, interests, disciplines, and needs, one of the main elements of this work focuses on developing an ontology network to connect these different land administration visions. This ontology network aims to provide the semantics enabling data mediation and to build land administration knowledge graphs.

To achieve this goal, we built an ontology network through a hybrid approach (Wache et al. 2001). Thereby, the semantics of each agency is described by its own ontology, and mappings among ontologies are set manually. In addition, one global shared ontology is built using national and international standards.

We applied the NeOn methodology (Suárez-Figueroa et al. 2012), which specifies nine scenarios that incorporate commonly occurring circumstances in the ontological development process, e.g. when available ontologies require to be re-engineered, aligned, modularized, or integrated with non-ontological resources, putting a singular emphasis on
re-engineering and reusing knowledge resources (ontological and non-ontological). Further details about this methodology and its related scenarios are described in (Suárez-Figueroa et al. 2012). Next, we show details associated with different modules that define our ontology network.

4.1. Modeling land administration in the Colombian’s context

Taking into account that the interesting point in hybrid approaches is how the local ontologies are described (Wache et al. 2001), we begin the description of our ontology network on the bottom layer, that is, with the ontologies developed for each agency (IGAC and Bogota cadaster). Later on, we present two modules related to LADM (the National LADM profile and the LADM standard) and geospatial aspects. Finally, we provide details about these different modules’ connections to achieve our land administration ontology network.

4.1.1. National land administration’s module (IGAC model)

This module presents the National Cadastral System model, an attempt to consolidate a unique cadastral model developed by the IGAC. This model proposes a distinction between rural and urban to distinguish different administrative boundaries and keep consistency with previous information management systems. In the National Cadastral System model, Land is one of its main land administration features. Land refers to the geographical space with a defined extension occupied by a (rural or urban) property. Each of these different Land is part of a Vereda (administrative division) or Block (for RuralLand or UrbanLand, respectively) and can have zero or many (0:N) Building(s).

The module was developed using Scenario 1 (from specification to implementation) and Scenario 2 (reusing and re-engineering non-ontological resources) of the NeOn methodology since we reused the Entity-Relationship model of the mentioned Cadastral National System and its associated feature catalog14 to build this module. Figure 1 depicts more details about the discussed classes and their relations.

4.1.2. Local land administration’s module (Bogota Cadastre model)

The development of this module was based on the geographical feature catalog15 from the Bogota Spatial Data Infrastructure, developed according to the methodology for feature cataloging (ISO 2005). Considering this resource type, we also used Scenario 1 (from specification to implementation) and Scenario 2 (reusing and re-engineering non-ontological resources) of the NeOn methodology for building this module.

Within this module, Plot is a reference feature at this level of land administration. It refers to the minimum (urban or rural) geographic unit where one or more Properties are located. A Plot has associated a Use, may also contain Properties and Buildings, and is part of a Vereda or Block. At the same time, a Property has associated with a Stratum (a socio-economic classification of properties). Figure 2 depicts the aforementioned classes and their relationships.

4.1.3. Multipurpose cadaster module (The National LADM profile)

As we mentioned above, the multipurpose cadaster is a Colombian National project sponsored by The World Bank that aims to support increased tenure security and provide access to cadaster information through a land administration system. In this project, the Colombian National profile of the LADM has been developed. This is the common exchange model enabling cadastral information registration and management founded on the National Resolution (499/202016). We used the Unified Modeling Language (UML) models of this profile for developing this module, applying Scenario 1 (from specification to implementation) and Scenario 2 (reusing and re-engineering non-ontological resources) of the NeOn methodology. This module includes classes such as PieceOfLand (a land portion with a defined geographical extension), Plot (An specialized class that describes the Basic Administrative Unit (BAUnit)), or Easement (A type of spatial unit that allows the representation of a right-of-way associated with a BAUnit), among others. Figure 3 shows these classes and their relationships.

It is important to note that we focused on some packages of the National LADM profile to develop this ontological module, concretely on the Spatial Units package and parts of the Basic Administrative Units and Spatial sources packages. Likewise, we did not consider information related to legal/administrative aspects because this information was not available since it concerns personal data privacy.

4.1.4. LADM module

The LADM module development was associated with reusing the LADM standard described in ISO 19132 (ISO 2012) and reusing the ontology described in (Soon 2013). With these resources, we adopted Scenario 2, Scenario 3 (reusing ontological resources), Scenario 4 (reusing and re-engineering ontological resources), and Scenario 8 (restructuring ontological resources) of the NeOn methodology for developing this module.
Figure 1. The core of the National land administration model.

Figure 2. The core of the Bogota land administration model.
Regarding the ontology described in (Soon 2013), we discovered several anomalies or pitfalls that demonstrated this ontological resource did not conform to ontology modeling best practices after an initial analysis. We found these pitfalls using OOPS! (Ontology Pitfall Scanner!17), a tool to detect pitfalls in ontologies (Poveda-Villalón, Gómez-Pérez, and Suárez-Figueroa 2014). A list of the most common ontologies problems is gathered in (Rector et al. 2004; Poveda, Suárez-Figueroa and Gómez-Pérez, 2010). Next, we present some pitfalls that we addressed in the aforementioned ontological resource applying Scenario 4 of the mentioned methodology:

- **Missing annotations**: ontology terms lack annotations properties. These kinds of properties are valuable to enhance ontology usability and understanding.
- **The effect of range and domain constraints as axioms.** This pitfall is also identified as a recursive definition, that is, an ontology element (a class, an object property, or a datatype property) is utilized in its own definition.
- **Defining inappropriate inverse relationships.** Two relationships are characterized as inverse relations when they are not necessarily inverse.
- **Missing domain or range in properties.** Relationships and/or attributes without range or domain (or none of them) are incorporated in the ontology.
- **Defining multiple domains or ranges in properties.** The range or domain (or both) of a property (relations and attributes) is described by stating more than one rdfs:domain or rdfs:range statements.

In addition to the mentioned pitfalls, we discovered that the ontological resource provided by Soon (2013) did not completely conform to the LADM standard. Therefore, we performed reusing and re-engineering processes (Scenario 4) and an extension (Scenario 8) of the mentioned ontology using this ISO standard as a non-ontological resource (Scenario 2). In this line, we decided to adopt the general transformation rules described in (ISO 2015) to transform the UML models to OWL ontologies. Some examples of changes carried out in the ontology are presented in Tables 1 and 2, where axiom
definitions are formally depicted associated with a class that did not exist (Table 1). Several objects and data properties are added to the new ontology version (Table 2).

Figure 3 presents a partial view of the connections between classes of the LADM module and the multipurpose cadaster module, differentiated by using different prefixes (ladm and ladm_co, respectively) and color plots.

4.1.5. Geospatial module

The geospatial module was developed following Scenario 3 of the NeOn methodology, reusing GeoSPARQL (Perry and Herring 2012), a vocabulary to describe geospatial data in RDF and a SPARQL extension for processing geospatial data. This proposal allows distinct geometries (e.g., lines, points, polygons, multipoints, etc.), handling multiple coordinate reference systems, and incorporating spatial relations (e.g., touches, intersects, overlaps, etc.) for querying geographic datasets. Then, geometries are defined by the class Geometry and the property defaultGeometry, and the coordinates can be encoded employing Well-Known-Text (WKT) or Geography Markup Language (GML).

4.1.6. Ontology network development

Some examples of how these different ontology modules relate to them in order to create our ontology network for the Colombian land administration are shown in Figure 4. For instance, igac:Neighborhood is equivalent to idea:Neighborhood, igac:Block is equivalent to idea:Block, and igac:Vereda is equivalent to idea:Vereda. These classes are subclasses of ladm: SpatialUnitGroup, which is associated with ladm: SpatialUnit class using ladm:describesWhole relation. Furthermore, ladm:SpatialUnitGroup is defined by a geo:Geometry using the relation geo:defaultGeometry. This latter relation allows the aforementioned subclasses to be associated with a spatial representation through its geometric component. More formally:

![Figure 4. An overview of our integrated proposal for the Colombian land administration.](image-url)
This network was developed defining relations between the mentioned modules to provide an integrated model of the Colombian land administration. It was supported by Protégé and expressed in OWL2.

5. Land administration knowledge graphs. A case study

Colombian land administration’s knowledge graphs consist of several datasets containing information about land use, ownership, value, and associated resources provided by two agencies at different levels. It is important to note that we uptake this information to create knowledge graphs, but we omit some details in this article to preserve owners’ privacy. The deployed data cycle allows generating knowledge graphs, connecting them to the Linked Data cloud, and exploiting them in an integrated manner (see Figure 5). Next, we present details of the steps of this data cycle.

5.1. Data sources

We considered data sources generated, maintained, and updated by two different land administration agencies, such as the National (IGAC) and a local agency (Bogota cadaster). As we mentioned before, this entails that data sources have different issues related to heterogeneity (formats, models, vocabularies, or resolutions). These data sources were initially described in tables, shapefiles, geodatabases, and OGC WFS services.

- Tables. The Bogota cadaster provides Open Data about land properties (predios) and related uses. This dataset contains 22 attributes (e.g., area, address, cadastral identifier, type of use, updating date, among others) and offers more than 5 million records in version 12.19 (June 2020). It is available in CSV format using a Creative Commons Attribution 4.0 license.
- Shapefiles and Geodatabases. We collected a set of these types of files from both agencies. Concerning IGAC, diverse shapefiles and a geodatabase were gathered from its Open Data portal. These sources contain information for urban and rural sectors related to buildings, blocks, addresses, plots, etc. On the other hand, the Bogota cadaster files were collected from its Open Data portal and contain information associated with blocks, addresses, cadastral sectors, buildings, or plots. These data were available using WGS84 (EPSG:4326) and MAGNA-SIRGAS (EPSG:4686) as coordinate reference systems. Regarding data licenses, both datasets (National and local) are licensed for free distribution following a Creative Commons Attribution 4.0 license.
- OGC WFS. The WFS services (Shekhar, Xiong, and Zhou 2017) are an interface defined by OGC that permits geographic data exchange across the Web. This kind of service determines the rules to request and retrieve geographic information using HTTP.
With respect to the IGAC, we used two WFS services from its geoportal. On the one hand, a service that provides information related to land, buildings, blocks, and veredas, and on the other hand, a service with cartographic information on a scale of 1:2,000. Additionally, we selected a WFS service from the Bogota cadaster that contained land information about the use, altitude, social stratum, etc. Also, the Bogota SDI provided a WFS service related to reference information, where diverse information layers (e.g. transport network, water bodies, cadastral sectors, etc.) can be requested. Both WFS services are available using WGS84 (EPSG:4326) as a default coordinate reference system.

5.2. Data conversion

The data collected from two aforementioned land administration agencies were transformed to RDF using our developed ontology network to obtain a set of integrated knowledge graphs. We adopted graphs as a knowledge representation since knowledge graphs work as a lingua franca between humans and machines, in that it is structured enough for machines to handle and ingest with semantics and is intuitive sufficient for humans (Kejriwal 2019). In addition, we applied RDF as the underlying knowledge representation model to adapt distinct formats of our datasets, circumvent proprietary formats, and provide interoperable representations (Vilches-Blázquez and Saavedra 2019; Seneviratne et al. 2018).

In order to generate knowledge graphs, we utilized OpenRefine and our GeoLOD framework (Vilches-Blázquez and Saavedra 2019) for datasets related to tables, and shapefiles and WFS services, respectively. Figure 6 shows an overview of the generated knowledge graphs, where we integrated different and heterogeneous datasets mentioned from both agencies through some elements of our ontology network.

To handle tables, we used Open Refine and, more concretely, the RDF Refine extension (Verlic 2012). This tool and its extension provide an interface to import data from various sources like XML, CSV, etc., allowing importing ontologies, modeling data structure by aligning the columns with components of imported vocabularies, and the result is built as RDF. Listing 1 depicts an excerpt of RDF data generated from the Bogota land property (predio) table, where several data details are displayed (gray shapes) in Figure 6.

**Listing 1. Land property RDF data from tables.**

```xml
resource:AAA0207NHNX a bog:Predio;
  bog:chip “AAA0207NHNX”;
  bog:idPredio ”76287”; bog:preChip “AAA0207NHNX”;
  [...]
  bog:address “CL 32 13 52 GS 2112”.
```

Figure 6. An excerpt of land administration knowledge graphs.
For transforming shapefiles, geodatabases, and WFS services into RDF according to our ontology network, we used two elements of our GeoLOD framework, specifically, SHP2GeoSPARQL and WFS2GeoSPARQL (Vilches-Blázquez and Saavedra 2019). It is important to note that we performed an extension of SHP2GeoSPARQL for dealing with geodatabases. Both elements work as web applications, where geometry and spatial relations associated with land administration features were converted into RDF using our ontology’s diverse elements (classes, relationships, and attributes). In this way, we generated knowledge graphs (RDF) from the aforementioned shapefiles, geodatabases, and WFS services, where a common element establishes the connection between datasets, that is, the identifier (ID) of land administration features. Listing 2 and Listing 3 present some excerpts of generated RDF data and how we developed our knowledge graphs with diverse data sources. For instance, the graph of a specific block (manzana) (http://datos.igac.gov.co/id/catastro/bogota/008108006) is built with geometric information (Listing 2) and associated uses and values (Listing 3) from shapefiles and WFS service, respectively. Some details of these RDF data are depicted in Figure 6 as green (shapefiles and geodatabases) and blue (WFS services) shapes.

Listing 2. An excerpt of RDF data from shapefiles.

```
<http://datos.igac.gov.co/id/catastro/bogota/008108006>
  a ladmcManzana, bog:Manzana, geo:Feature ;
  bog:id "008108006" ;
  geo:defaultGeometry <http://datos.igac.gov.co/id/catastro/bogota/Geometry/7fb92fc5c5e6e168d90cfc06791dac99ba495ecb> a sf:Polygon ;
  geo:asWKT "POLYGON ((-74.0685102599994 4.619458722000047, -74.0685102599994 4.619458722000047, -74.0685102599994 4.619458722000047, -74.0685102599994 4.619458722000047));"^^sf:literal .
```

Listing 3. An excerpt of RDF data from WFS.

```
<http://datos.igac.gov.co/id/catastro/bogota/008108006>
  bog:valorum2 "2167897,0000000000000000"^^xsd:double ;
  bog:grupoUso "RESIDENCIAL" .
```

5.3. Data linking

We also applied additional SHP2GeoSPARQL and WFS2GeoSPARQL functions tools to enrich and interconnect the generated graphs (RDF data) through spatial connections between different land administration features. These tools carry out topological analyses, and when a link between two features is located, an spatial relation is defined employing the GeoSPARQL vocabulary. Listing 4 shows a spatial relation between Block (manzana) and Neighborhood (barrio) through geo:sWithin relation, and an example of these relations is presented in Figure 6.

Listing 4. Spatial relations between RDF data.

```
<http://datos.igac.gov.co/id/catastro/bogota/008108006>
  a ladmcManzana, bog:Manzana, geo:Feature ;
  geo:sWithin <http://datos.igac.gov.co/id/catastro/bogota/008108006> .
```

Considering that the value of data and its benefits increase when it is more connected with other data (Heath and Bizer 2011), we also decided to interrelate these graphs with several datasets of the Linked Open Data cloud in a two-fold way: On the one hand, we applied SILK. This open-source framework allows us to connect some specific parts of our graphs with DBpedia (Lehmann et al. 2015); a community project obtains structured, multilingual knowledge from Wikipedia and where existing connections to this knowledge base from multiple datasets published on the Web make one of the reference hubs in the Linked Open Data cloud. We set up SILK using various similarity metrics, and obtained an interlinking process between our graphs and DBpedia, materialized with owl:sameAs links. An example of the connections between both data sources is depicted in Listing 5.

Listing 5. RDF data interlinked to DBpedia.

```
<http://datos.igac.gov.co/id/catastro/bogota/008108006>
  a bog:Barrio, geo:Parameter ;
  geo:sWithin <http://datos.igac.gov.co/id/catastro/bogota/Geometry/be5032c1a a6898d65504be794f8356d6d754eb4e> .
```

On the other hand, we utilized spatial information associated with our land administration knowledge graphs to enrich them with new data about points of interest (POIs) from LinkedGeoData and GeoNames. Thus, we used our GeoLOD framework to discover POIs associated with the land parcel RDF data. The framework employed geometric information and addresses to execute spatial analyses. It allowed setting diverse topological relations using these attributes by means of the GeoSPARQL vocabulary (e.g. contain, intersect, touch, etc.) between considered datasets. Listing 6 shows two examples of this enrichment process using spatial relations, where land administration knowledge graphs were enriched with POIs, for instance, a library (GeoNames) and a supermarket (LinkedGeoData).
5.4. Exploiting the land administration integration

Once knowledge graphs have been generated, we deployed them through a triple store that allows accessing, querying, and exploiting the land administration graphs. Our knowledge graphs were published using the Parliament triple store, which was chosen because it deals with the GeoSPARQL standard and can exploit the data's spatial component. The knowledge graphs are available on http://20.115.164.119:8089/parliament/.

In addition, we extended a component of our GeoLOD framework to allow querying and displaying our knowledge graphs. This component utilizes LOD4WFS, Leaflet, and Apache Jena libraries and permits us to exploit land administration graphs using GeoSPARQL and provide visualizations of the outputs of each semantic query.

Both components (triple store and visualization) enable us to exploit the main benefits of our knowledge graphs, that is, to query and display data that were handled separately by two mentioned governmental agencies through different land administration systems. Then, as a first example, we demonstrate an integrated view of our knowledge graphs with the GeoSPARQL query of Figure 7. Here, we requested buildings with more than three floors from two neighborhoods of Bogota (Bogota cadaster) and Soacha (IGAC) using our developed ontology network. The obtained results are displayed with a blue color on the map.

On the other hand, we can also retrieve information that was initially in different data sources within the same agency. For instance, Figure 8 presents an example where multiple Bogota cadaster data sources are queried and visualized. Concretely, the query allows recovering plots (lotes) with residential use, classified as levels 1 and 2 with respect to its social stratum (a Colombian socio-economic classification), and including its valuation information within a specific neighborhood (La Candelaria) in Bogota. The mentioned figure also depicts results...
obtained on a map highlighting in blue color those elements of our knowledge graphs that satisfied the query criteria.

6. Discussion and conclusions

Land administration data and systems’ well-known heterogeneity entails relevant constraints for performing data interpretation, visualization, and analysis by land researchers, managers, or related specialists. It is remarkably outstanding when data come from distributed and multi-jurisdictions agencies since heterogeneity issues are connected to changing definitions, historical legacies, standards across jurisdictions, etc. (Chen et al. 2018).

Considering this heterogeneity scenario, we constructed a graph-based representation of knowledge for integrating multiple data sources belonging to two Colombian agencies within a decentralized and multi-jurisdictional land administration scenario. The developed knowledge graphs overcome semantic heterogeneity, breaking down data-silos by means of an ontology network built using multi-jurisdictions models and national and international standards. Once data were semantically integrated (knowledge graphs) and enriched (linked to diverse data sources in the Web of Linked Data), visualization and analysis operations may be carried out using the developed ontology network as a crucial component, without requiring hardcoded logic to fix distinct structures of data inputs. This fact evidences a significant benefit,
especially when dealing with data from decentralized and multi-jurisdiction agencies. We presented the process to generate land administration knowledge graphs through a real case study in the previous sections. Some lessons were learned from this work and we explain them here.

Land administration modeling is the base of cadastral information systems, where every data component that is significant in land administration is represented and eventually has its records stored (Kalantari et al. 2015). In the knowledge graph context, these elements are modeled in ontologies represented in the graph alongside the data level. In this way, the representation is enriched to handle the real world’s complexity and enable learning, reasoning, and inference abilities (Bellomarini, Sallinger, and Vahdati 2020). Therefore, it is fundamental to understand domain knowledge (land information) and ontological engineering scenario (ontology definitions, terms, usages, and tools) to accomplish this task.

According to (Chen et al. 2018), the best approach for creating an ontology is a close collaboration between domain knowledge experts and ontology engineers. Often, if it does not perform in this manner, several anomalies or pitfalls emerge in the ontological resources, manifesting that they did not adhere to ontology modeling best practices. Thereby, we adopted the proposal of Chen et al. (2018), and this expert combination allowed addressing several pitfalls and completing the ontological resource provided by Soon (2013).

The construction of knowledge graphs can benefit land administration from improved interoperability using open standards. Improved interoperability decreases deployment time, moderated system and data lifecycle costs, increases flexibility and scalability, enhances decisions from technologies, and improves the capacity to exchange, share, and integrate information related to land administration (Lemmen et al. 2018). Furthermore, improved interoperability involves achieving semantic interoperability and increasing the availability of open government data. It can make more accessible the understanding of existing heterogeneity and supports a new efficient and valuable land administration scenario (Rajabifard, Ho, and Soheil 2016).

We want to remark that we did not provide data concerning ownership in this work since it concerns personal data privacy. In this sense, land administration agencies need to protect issues related to this kind of data’s privacy and security.

Additionally, we believe that our approach will be helpful to government organizations who manage land administration information across the country and at different administration levels. However, we know that our approach’s deployment can entail some additional issues due to an operational implementation inside land administration agencies. Although it is out of the scope of our work due to strategic decisions that involve a change of technological paradigm within the governmental agencies; nevertheless, based on our experience, we recommend performing an implementation of our approach based on pilots to test them and move up gradually considering the organizational context and learned lessons during pilot projects.

This paper has presented a graph-based representation of knowledge for integrating multiple and heterogeneous land administration data sources (tables, shapefiles, geodatabases, and WFS services) belonging to two mentioned Colombian agencies (Bogota cadastral and IGAC). In this way, we carried out this work without modifying the current processes, models, and vocabularies used by two considered agencies and where the utilizing of the developed ontology network was the crucial component to resolve the aforementioned heterogeneity. Additionally, we have tested this work’s results exploiting an integrated and harmonized view of land administration using SPARQL queries in order to evidence the potential of knowledge graphs developed.

The scope of the investigation was restricted to the Colombian land administration domain. Nevertheless, other land administration systems can consider the model’s suitability for different scenarios and attempt the LADM standard implementation. In addition, the proposed data cycle can be applied to other decentralized land administration agencies to accomplish an integrated land administration system since it entails the additional benefit without modifying existing processes, models, or vocabularies inside agencies. However, the complexity of managing these issues requires that stakeholders from varied disciplines collaborate in the development and operationalization of this new land administration scenario (Lemmen et al. 2018).

Notes

1. http://ideca.gov.co/.
2. http://idesc.cali.gov.co/.
3. https://www.medellin.gov.co/geomedellin.
4. http://www.icde.org.co/.
5. https://www.igac.gov.co/sites/igac.gov.co/files/ladm col_v.0_5_17082016.pdf.
6. The used dataset in this work belongs to Bogotá Spatial Data Infrastructure (IDEECA) – http://www. ideca.gov.co/.
7. https://www.w3.org/RDF/.
8. https://www.w3.org/2001/sw/wiki/RDFS.
9. https://www.w3.org/OWL/.
10. https://www.w3.org/TR/owl2-overview/.
11. https://googleblog.blogspot.com/2012/05/introdu cing-knowledge-graph-things-not.html.
12. https://www.w3.org/TR/sparql11.query/.
13. http://landregistry.data.gov.uk/.
temporal data analysis and mining. He is co-author of more than 70 refereed articles in peer-reviewed journals, book chapters, conferences, and workshops.

Johnny Saavedra received his M.Sc. in Information Technology from the Universidad Politécnica de Madrid (Spain), in 2011. He is currently a Ph.D. researcher at the Technical School of Engineers in Topography, Geodesy, and Cartography of the Universidad Politécnica de Madrid. His research interests include Spatial Data Infrastructures and Linked Data, specifically in cadastral and biodiversity domains.

**ORCID**

Luis M. Vilches-Blázquez [http://orcid.org/0000-0001-5799-469X](http://orcid.org/0000-0001-5799-469X)

**Data availability statement**

The knowledge graphs associated with this contribution are available on [http://20.115.164.119:8089/parliament/](http://20.115.164.119:8089/parliament/).

**References**

Bank, W. 2019. Colombia – Multipurpose Cadastre Project. Washington, D.C.: World Bank Group. Accessed 30 October 2021. [https://documents.worldbank.org/en/publication/documents-reports/documentdetail/378461551361739801/colombia-multipurpose-cadastre-project](https://documents.worldbank.org/en/publication/documents-reports/documentdetail/378461551361739801/colombia-multipurpose-cadastre-project).

Bellomarini, L., E. Sallinger, and S. Vahdati. 2020. “Knowledge Graphs: The Layered Perspective.” In Knowledge Graphs and Big Data Processing, edited by V. Janev, D. Graux, H. Jabeen, and E. Sallinger, 20–34. Springer. LNCS 12072. Accessed December 13 2021. doi:10.1007/978-3-030-53199-7_2.

Berners-Lee, T. 2006. “Linked Data Design Issues.” Accessed 26 October 2021. [https://www.w3.org/DesignIssues/LinkedData.html](https://www.w3.org/DesignIssues/LinkedData.html).

Borzacchiello, M.T., and M. Craglia. 2013. “Estimating Benefits of Spatial Data Infrastructures: A Case Study on e-Cadastres.” Computers, Environment and Urban Systems 41: 276–288. doi:10.1016/j.compenvurbs.2012.05.004.

Çağdaş, V., and E. Stubkjær. 2015a. “A SKOS Vocabulary for Linked Land Administration: Cadastre and Land Administration Thesaurus.” Land Use Policy 49: 668–679. doi:10.1016/j.landusepol.2014.12.017.

Çağdaş, V., and E. Stubkjær. 2015b. “Core Immovable Property Vocabulary for European Linked Land Administration.” Survey Review 47 (340): 49–60. doi:10.1179/1752270614Y.0000000093.

Chen, Y., S. Sabri, A. Rajabifard, and M.E. Agunbiade. 2018. “An Ontology-based Spatial Data Harmonisation for Urban Analytics.” Computers, Environment and Urban Systems 72: 177–190. doi:10.1016/j.compenvurbs.2018.06.009.

Cudré-Mauroux, P. 2020. “Leveraging Knowledge Graphs for Big Data Integration: The XI Pipeline.” Semantic Web 11 (1): 13–17. doi:10.3233/SW-190371.

Díaz, A.A., and L.M. Vilches-Blázquez. 2014. “Cadastral Linked Data according to INSPIRE.” In INSPIRE Conference. Aalborg.

Doan, A., A.Y. Halevy, and Z.G. Ives. 2012. Principles of Data Integration. Amsterdam: Morgan Kaufmann.
Ehrlinger, L., and W. Wöß. 2016. “Towards a Definition of Knowledge Graphs.” SEMANTiCS, Vol. 48. Accessed 26 October 2021. http://ceur-ws.org/Vol-1695/paper4.pdf

FGDC (Federal Geographic Data Committee). 2008a. “Cadastral Data Content Standard for the National Spatial Data Infrastructure.” Version 1.4. Subcommittee on Cadastral Data Federal Geographic Data Committee. Accessed 30 October 2021. http://nationalcad.org/download/cadastral-data-content-standard-ver-1-4.pdf

FGDC (Federal Geographic Data Committee). 2008b. “Geographic Information Framework Data Standard Part 1, Cadastral.” Subcommittee on Cadastral Data Federal Geographic Data Committee. Accessed 26 October 2021. https://www.fgdc.gov/standards/projects/framework-data-standard/GIFrameworkDataStandard_Part1_Cadastral.pdf

FIG - International Federation of Surveyors. (1998). “Cadastr 2014: A Vision for Future Cadastral Systems.” Accessed 06 November 2021. https://www.fig.net/resources/publications/figpub/cadastr2014/translation/c2014-english.pdf

Folmer, E., W. Beek, and L. Rietveld. 2018. “Linked Data Viewing as Port of the Spatial Data Platform of the Future.” International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-4/W8: 49–52. doi:10.5194/isprs-archives-XLII-4-W8-49-2018.

Fonseca, F.T., M.J. Eigenhofer, P. Agouris, and G. Câmara. 2002. “Using Ontologies for Integrated Geographic Information Systems.” Transactions in GIS 6 (3): 231–257. doi:10.1111/1467-9671.00109.

Giuliani, G., P. Lacroix, Y. Guigoz, R. Roncella, L. Bigagli, M. Santoro, A. Lehmann, S. Nativi, N. Ray, and A. Lehmann. 2017. “Bringing GEOSS Services into Practice: A Capacity Building Resource on Spatial Data Infrastructures.” Transactions in GIS 21 (4): 811–824. doi:10.1111/tgis.12209.

Golshan, B., A.Y. Haley, G.A. Mihaila, and W. Tan. 2017. “Data Integration: After the Teenage Years.” In Proceedings of the 36th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, PODS 2017, 101–106. Chicago, IL, USA. May 14–19. doi:10.1145/3034786.3056124.

Gruber, T. 1993. “Toward Principles for the Design of Ontologies Used for Knowledge Sharing.” International Journal Human-Computer Studies 43: 907–928. doi:10.1006/ijhc.1995.1081.

Guarin, A., L. Baron, L. Jenni, and S. Salamanca. 2017. “LADM-A Tool for Land Administration in Post-Conflict Colombia.” In FIG Working Week 2017. Accessed 30 October 2021. https://www.fig.net/resources/proceedings/fig_proceedings/fig2017/papers/ts505e/Ts505e_jenni_baron_et_al_8853_abs.pdf

Hakimi, O., J.L. Gelpi, M. Krällinger, F. Curi, D. Repchevsky, and M.P. Ginebra. 2020. “The Devices, Experimental Scaffolds, and Biomaterials Ontology (DEB): A Tool for Mapping, Annotation, and Analysis of Biomaterials Data.” Advanced Functional Materials 30 (16): 1909910. doi:10.1002/adfm.201909910.

Hakimpour, F., and S. Timpf. 2001. “Using Ontologies for Resolution of Semantic Heterogeneity in GIS.” In Proceedings of 4th AGILE Conference on Geographic Information Science, 385–395. Brno, Czech Republic, April 19–21.

Heath, T., and C. Bizer. 2011. Linked Data: Evolving the Web into a Global Data Space. San Rafael, CA: Morgan & Claypool.

Hess, C., and M. de Vries. 2006. “From Models to Data: A Prototype Query Translator for the Cadastral Domain.” Computers, Environment and Urban Systems 30 (5): 529–542. doi:10.1016/j.compenvurbs.2005.08.008.

Huang, W., and L. Harrie. 2019. “Towards Knowledge-based Geovisualisation Using Semantic Web Technologies: A Knowledge Representation Approach Coupling Ontologies and Rules.” International Journal of Digital Earth 13 (9): 976–997. doi:10.1080/17538947.2019.1604835.

INSPIRE. 2014. Data Specification on Cadastral Parcels – Technical Guidelines. version 3.1. Accessed 30 October 2021. https://inspire.ec.europa.eu/id/document/tg/cp

ISO. 2005. “ISO 19110, Geographic Information – Methodology for Feature Cataloguing.” Accessed 30 October 2021. https://www.iso.org/iso/iso_catalogue/catalogue_detail.htm?csnumber=51206

ISO. 2015. “ISO 19150-2:2015 Geographic Information – Ontology – Part 2: Rules for Developing Ontologies in the Web Ontology Language (OWL).” Accessed 30 October 2021. https://www.iso.org/iso/iso_catalogue/catalogue_detail.htm?csnumber=57466.html

Jenni, L., M. Germann, C. Eisenhut, A. Guarin, and M. V. Bajo. 2017. “LADM Implementation in Colombia – Process, Methodology and Tools developed and applied.” In FIG Working Week 2017, Helsinki, Finland. Accessed 30 October 2021. https://www.fig.net/resources/proceedings/fig_proceedings/fig2017/papers/ts505e/Ts505e_jenni_baron_et_al_8853_abs.pdf

Jung, C.-T., C.-H. Sun, and M. Yuan. 2013. “An Ontology-enabled Framework for a Geospatial Problem-solving Environment.” Computers, Environment and Urban Systems 38: 45–57. doi:10.1016/j.compenvurbs.2012.10.008.

Kalantari, M., K. Dinsmore, J. Urban-Karr, and A. Rajabifard. 2015. “A Roadmap to Adopt the Land Administration Domain Model in Cadastral Information Systems.” Land Use Policy 49: 552–564. doi:10.1016/j.landusepol.2014.12.019.

Kejiwiål, M. 2019. “What Is a Knowledge Graph?” In Domain-Specific Knowledge Graph Construction, 1–7. Cham: Springer. Springer. Briefs in Computer Science. doi:10.1007/978-3-030-12375-8_1.

Krotzsch, M., and V. Thost. 2016. “Ontologies for Knowledge Graphs: Breaking the Rules.” In International Semantic Web Conference, 376–392. Cham: Springer. doi:10.1007/978-3-319-46523-4_23.

Lehmann, J., R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P.N. Mendes, S. Hellmann, et al. 2015. “DBpedia—a Large-scale, Multilingual Knowledge Base Extracted from Wikipedia.” Semantic Web 6 (2): 167–195. doi:10.3233/SW-140134.

Lemmen, C., P. van Oosterom, M. Kalantari, E.-M. Unger, and C. de Zeeuw. 2018. OGC White Paper on Land Administration. Accessed 30 October 2021. https://portal.opengeospatial.org/files/artifact/id=77835&version=2

Lenzerini, M. 2002. “Data Integration: A Theoretical Perspective.” In Proceedings of the Twenty First ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 233–246. Madison, Wisconsin, USA, June 3–5. doi:10.1145/543613.543644.

Li, M., X. Zhu, C. Shen, D. Chen, and W. Guo. 2012. “The Development of Cadastral Domain Model Oriented at Unified Real Estate Registration of China Based on Ontology.” ISPRS Annals of the
Methods for Handling Environmental Information: Special Aspects of Processing Space and Time. Magdeburg, Germany.

Suárez-Figueroa, M.C., A. Gómez-Pérez, E. Motta, and A. Gangemi, Eds. 2012. Ontology Engineering in a Networked World. Heidelberg: Springer. doi:10.1007/978-3-642-24794-1.

Tandy, J., L. van den Brink, and P. Barnaghi. (Ed.). 2017. Spatial Data on the Web Best Practices, W3C Working Group Note. Accessed 30 October 2021. https://www.w3.org/TR/sdw-bp/

UNECE. 1996. Land Administration Guidelines. With Special Reference to Countries in Transition. Geneva, Switzerland: United Nations/Economic Commission for Europe.

van den Brink, L., P. Janssen, W. Quak, and J. Stoter. 2017. “Towards a High Level of Semantic Harmonisation in the Geospatial Domain.” Computers, Environment and Urban Systems 62: 233–242. doi:10.1016/j.compenvurbsys.2016.12.002.

van Oosterom, P., and C. Lemmen. 2015. “The Land Administration Domain Model (LADM): Motivation, Standardisation, Application and Further Development.” Land Use Policy 49: 527–534. doi:10.1016/j.landusepol.2015.09.032.

van Oosterom, P., C. Lemmen, T. Ingvarsson, P. van der Molen, H. Ploeger, W. Quak, J. Stoter, and J. Zevenbergen. 2006. “The Core Cadastral Domain Model.” Computers, Environment and Urban Systems 30 (5): 627–660. doi:10.1016/j.compenvurbsys.2005.12.002.

Verlic, M. 2012. "LODGrefine-LOD-enabled Google Refine in Action." In I-SEMANTICS (Posters & Demos), Graz, Austria, 31–37.

Vilches-Blázquez, L.M., B. Villaizán-Terrazas, O. Corcho, and A. Gómez-Pérez. 2014. “Integrating Geographical Information in the Linked Digital Earth.” International Journal of Digital Earth 7 (7): 554–575. doi:10.1080/17538947.2013.783127.

Vilches-Blázquez, L.M., and J. Saavedra. 2019. “A Framework for Connecting Two Interoperability Universes: OGC Web Feature Services and Linked Data.” Transactions in GIS 23: 22–47. doi:10.1111/tgis.12496.

Wache, H., T. Voegele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Huebner. 2001. “Ontology-based Integration of Information – A Survey of Existing Approaches.” In Proceedings of the workshop on Ontologies and Information Sharing at the International Joint Conference on Artificial Intelligence (IJCAI), 108–117. Accessed 30 October 2021. http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-47/wache.pdf

Williamson, I.P., S. Enemark, J. Wallace, and A. Rajabifard. 2010. Land Administration for Sustainable Development. Redlands, CA: ESRI Press.