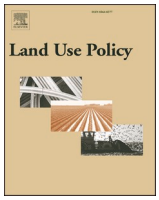
Land Use Policy 104 (2021) 105381

Contents lists available at ScienceDirect

Land Use Policy

journal homepage: www.elsevier.com/locate/landusepol



Alignment of standards through semantic tools – The case of land administration

Erik Stubkjær [a,*,1], Volkan Çağdaş [b]

[a] Aalborg University, Department of Planning, Aalborg DK-9000, Denmark
[b] Yildiz Technical University, Department of Geomatic Engineering, Istanbul 34220, Turkey

ARTICLE INFO

Keywords:
ISO LADM
OGC LandInfra/InfraGML
Alignment of standards
E-government
Interoperability
Knowledge Organization Systems (KOS)
Cadastre and Land Administration Thesaurus (CaLAThe)

ABSTRACT

The fact that both the ISO 19152:2012 Land Administration Domain Model (LADM) and the OGC 15–111r1 Land and Infrastructure Conceptual Model Standard (LandInfra) exist within the domain of Land Administration, calls for a mediating platform, which could support an alignment of the standards. The purpose of this article is to demonstrate, how semantic tools have been used to establish such platform.

Semantic tools include thesauri. The Cadastre and Land Administration Thesaurus (CaLAThe) was issued in 2011 and based on the then draft version of the ISO LADM standard. CaLAThe was further developed to include also terms of the OGC LandInfra standard of 2016. The recent version (December 2019) integrates code list names and values, reflecting both the relevant Survey and Land division parts of LandInfra and, as proof of concept, some of the code lists mentioned in the Annex J Code lists of LADM. CaLAThe thus constitutes a shared universe of discourse, a frame for joint management of the code lists of the mentioned standards.

Taking a problem-based approach to the design of artefacts like standards, the mentioned alignment of standards is part of a problem hierarchy. In this hierarchy, development of information standards provides a means for improving public administration, more specifically its aspects of e-government and interoperability. These aspects have spurred interest in semantic tools; for example, standardization organizations have recently publicized their terminological resources in terms of the ISO/TC211 Geolexica glossary and the OGC Definitions Server. The definitions of CaLAThe and the mentioned code lists are by December 2019 included into the OGC Definitions Server. Besides servicing the Land Administration communities, CaLAThe illustrates the use of thesauri as a cross-domain means for supporting the alignment of standards.

1. Introduction

The notion of semantic tools reflects the fact that the internet provides for access to data, rather than to documents, allowing data to be shared and reused across applications, organizations, and international communities. Data access requires common schemes, leading to the development of Knowledge Organization Systems (KOS). Ontologies were developed a decade before (Gruber, 1993), but the notion of KOS allowed for comparing heterogeneously structured systems, including controlled vocabularies, thesauri, and ontologies (Zeng and Chan, 2004; Zeng and Mayr, 2019).

The development and use of such semantic tools for mediating among standards is a complex endeavor. The design research methodology of Peffers et al. (2008) suggests a problem-based approach to the design of artefacts like standards, which seems to compare to the methodology approach applied for development of core vocabularies for e-government through The Interoperability Solutions for European Public Administrations (ISA) Programme of the European Commission.[1] The combined specification of Core Business, Core Location and Core Person Vocabulary was issued in 2012.[2] The semantic aspect addressed here may be seen in the context of the Reference Model for Open and Distributed Processing[3] (RM-ODP), which presents five perspectives: the

* Correspondence to: Aalborg University, Department of Planning, Rendsburggade 14, niveau 3, DK-9000 Aalborg, Denmark.
E-mail address: est@plan.aau.dk (E. Stubkjær).
[1] https://vbn.aau.dk/en/persons/erik-stubkjaer
[1] https://ec.europa.eu/isa2
[2] https://joinup.ec.europa.eu/solution/core-location-vocabulary/about
[3] http://www.rm-odp.net

https://doi.org/10.1016/j.landusepol.2021.105381
Received 26 June 2020; Received in revised form 23 February 2021; Accepted 24 February 2021
Available online 6 March 2021

enterprise viewpoint, the engineering viewpoint, the information viewpoint, the computational viewpoint and the technical viewpoint, where the 'information viewpoint' includes semantic assets as the above-mentioned core vocabularies (Vandenbroucke et al., 2020). The following develops within the information viewpoint.

The research in KOSs, the availability of the then draft ISO 19152:2012 Land Administration Domain Model (LADM; ISO, 2012b) and the search for a sound, scientific basis for cadastral research triggered the development of the Cadastre and Land Administration Thesaurus (CaLAThe), which was issued in 2011 (Çağdaş and Stubkjær, 2015a). CaLAThe was used as basis for proposing 'A Core immovable property vocabulary for European linked land administration' (Çağdaş and Stubkjær, 2015b). The proposal extended the Core Location Vocabulary and reflected European projects concerning land registries, including European Land Information Service[4] (EULIS) and EULIS Project Land Information for Europe (LINE). Moreover, it pointed to the challenges of establishing interoperability among the registries within the domain of Land Administration, besides cadastre and land registry also address registry or building and dwelling registry and the property tax registry. CaLAThe benefitted from a vocabulary of legal terms, prepared within the EULIS project. The European Land Registers Association[5] (ELRA) framed the development of further projects: IMOLA I[6] and IMOLA II,[7] and specified a European Land Register Document[8] (ELRD). The ELRD provides an interoperability solution to the variations found in individual land registries and the different formats they use. Unfortunately, contact to a vocabulary component of IMOLA projects failed.

The issuing of OGC LandInfra encoding in InfraGML in 2017 (OGC, 2016; OGC, 2017), meant that two international standards were available within the domain of Land Administration, in addition to regional standardization efforts, e.g. the Australian and New Zealand ICSM ePlan Protocol. Early, a harmonization concern was raised (Stubkjær and Scarponcini, 2017), suggesting that the upcoming revision of LADM might serve as a frame for this. The idea was further developed with reference to Paasch et al. (2013), who propose both a refined class diagram and code lists as a mean of internationalization by which the classes of LADM may be related to national jurisdictions (Stubkjær et al., 2018). In order to support such joint code list management, new versions of CaLAThe were issued during 2019, which include also the relevant Survey and Land division parts of the OGC LandInfra standard, as well as code list names and values (OGC, 2016). As proof of concept, CaLAThe includes also a few of the code lists mentioned in the informative Annex J of LADM. CaLAThe with code lists are now available for users worldwide through BARTOC,[9] and also through the OGC Definitions Server.[10]

Drawing on Stubkjær et al. (2018) and Stubkjær et al. (2019), the article introduces the semantic tools, which were used in the context of CaLAThe development. It unfolds how CaLAThe provides a frame for joint management of the code lists of the mentioned standards. This shared universe of discourse may support an alignment of standards within the domain of Land Administration. The following Section 2 introduces the semantic basis, pertinent standards, and registries, as background for motivating the development path of CaLAThe (Section 3). The implementation of code lists from LADM and LandInfra/InfraGML standards in CaLAThe is documented in Section 4. Discussion and Conclusion sections close the article.

[4] https://eulis.eu
[5] https://www.elra.eu
[6] https://www.elra.eu/imola/
[7] https://www.elra.eu/imola-iii/
[8] https://e-justice.europa.eu/content_land_registers_at_european_level-108–maximize-en.do
[9] https://bartoc-skosmos.unibas.ch/CaLAThe/en/ (select Hierarchy)
[10] https://www.opengis.net/def/CaLAThe/4.0

2. Semantic Web structures and tools

2.1. Knowledge structures

The alignment of standards is a complex task, which among others includes the development of vocabularies used or implied by the standards that are going to be aligned. Technical standards, methodologies and tools developed within the Semantic Web domain are available for the development, maintenance and harmonization of such vocabularies. This section introduces to Knowledge Organization System (KOS) related standards, methodologies and tools, in order to support further initiatives aiming at alignment of standards within the domain of land administration.

A KOS is a general term, which refers to tools that present the organized interpretation of knowledge structures (Zeng and Chan, 2004, p. 377). It covers all types of schemes for organizing information and promoting knowledge management, such as glossaries, dictionaries, gazetteers, subject headings, taxonomies, thesauri and ontologies (Hodge, 2000, p. 4). They are also referred to as controlled vocabularies, structured vocabularies, value vocabularies, concept schemes, semantic assets, and classification by various standards (Golub et al., 2014).

Based on these schemes, the information to be represented may be structured with increasing detail in terms of taxonomies, thesauri, and ontologies, respectively. The following overview is based on Breitman et al. (2007), p 17ff:

A taxonomy classifies terms hierarchically, using the father-son (generalization, is-a, or type-of) relationship. Indeed, taxonomies allow only the father-son relationship, ruling out other relationships, such as part-of, cause-effect, association, and localization.

A thesaurus contains a set of relationships among concepts, organized in a taxonomic way, together with a set of semantic relationships, such as equivalence, broader or narrower, and association, which hold among the concepts.

"An ontology is a formal, explicit specification of a shared conceptualization" (Gruber, 1993). In other terms, an ontology is a representation of the knowledge of a domain, where a set of objects and their relationships are described by a vocabulary. However, in addition to the concept hierarchy, comparable to the taxonomy, supplemented with the relationships, comparable to the thesaurus, the ontology also comprises axioms, expressed in an appropriate logical language, e.g. describing causal relationships between concepts.

In order to render the increasing detail, or in other words: stronger semantics, cf. Fig. 1, a number of encoding standards are available: XML, RDFS, SKOS, OWL. They are introduced in Section 2.2.

The increased level of detail of these knowledge structures may imply the view that less detailed KOSs are not needed when an ontology is available. Kless et al. (2015) objects to such view and based on rigorous comparison states that 'results show that thesauri and ontologies need to be treated as 2 orthogonal kinds of models with superficially similar structures.' Soon (2013), Sladić et al. (2013) and Shi and Roman (2018) have addressed the issue of land administration ontology. Köpke (2019) demonstrates the use of ontologies in the context of XML-schema matching.

The functionality of KOSs may be assessed by a number of criteria and metrics, for instance findability, accessibility, interoperability and reusability which are so-called FAIR principles (Wilkinson et al., 2016); plus functionality, impactful and transformability which are so-called FAIR+FIT principles (Zeng and Clunis, 2020). See also: https://fairsharing.org.

2.1.1. Lists and code lists

The simplest type of knowledge structures is the List, cf. Fig. 1. These are applied in the ISO LADM and the OGC LandInfra standards in terms of code lists, cf. Table 1 below. Compliant with ISO 19103:2015 'Geographic information — Conceptual schema language' these standards apply the Unified Modelling Language (UML; ISO, 2015a) of the

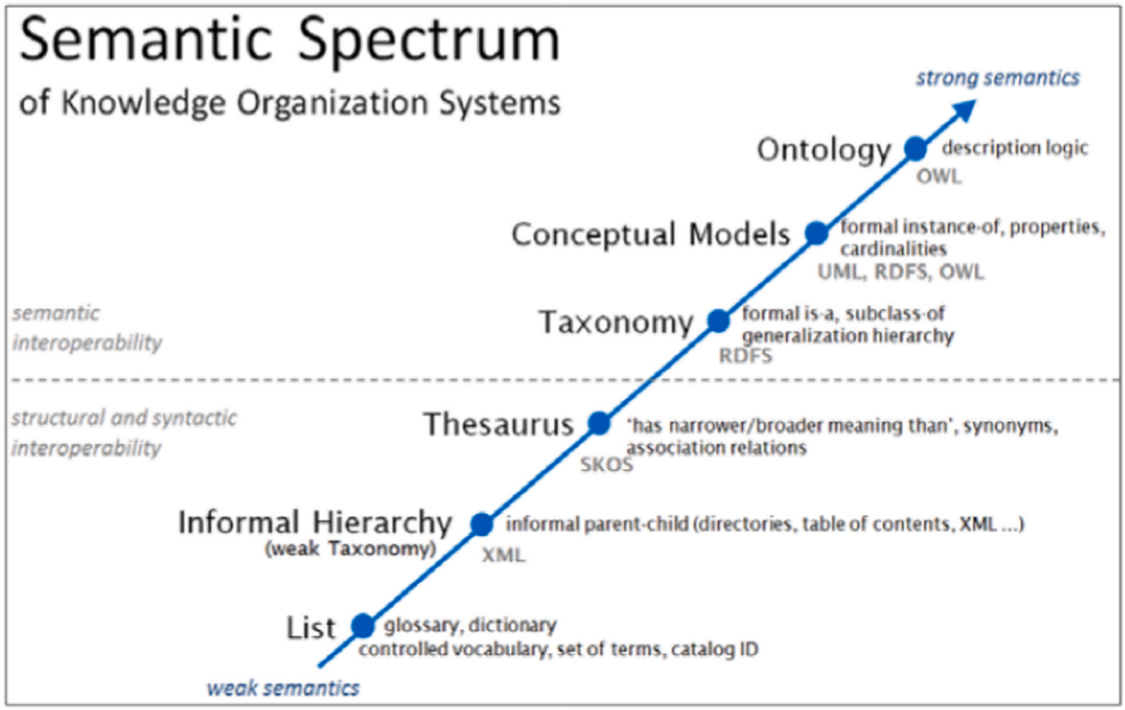


Fig. 1. The Semantic spectrum of Knowledge Organization Systems.
Source: Taxonomy and Ontology overview by Geoff Gross, OSTHUS GmbH (EMMC-CSA, 2017, p. (b) 7), adopted from Obrst (2010).

Table 1
Examples of related LADM and LandInfra code lists with values. The LADM informative Annex J show possible examples of values for these code lists. User communities have to define and manage their own values when implementing LADM. LandInfra code lists may be extended by the user community. Section 3 presents a mediating platform for these code lists.

LADM (Figures J.1,.2,.4)	LandInfra (Figure 65)
LA_PartyRoleType	**SigningRole**
• bank	• owner
• certifiedSurveyor	• landSurveyor
• citizen	• seller
• conveyor	• buyer
• employee	• neighbor
• farmer	• otherParty
• moneyProvider	
• notary	
• stateAdministrator	
• surveyor	
• writer	
LA_AdministrativeSourceType	**StatementType**
• agriConsent	• parcelEstablishment
• agriLease	• parcelAcquisition
• agriNotaryStatement	• cadastralAccount
• deed	• condominiumSchemeEstablishment
• mortgage	• condominiumAcquisition
• title	• encumbranceSchemeEstablishment
	• easementEstablishment
	• surveyMonumentEstablishment
LA_SpatialSourceType	
• fieldSketch	
• gnssSurvey	
• orthoPhoto	
• relativeMeasurement	
• topoMap	
• video	
LA_MonumentationType	**SurveyMonumentType**
• beacon	• boundaryMark
• cornerstone	• trigonometricMark
• marker	• siteMark
• notMarked	• otherMark
	• codeList: URI [0.1]

Object Management Group.[11] UML comprises several types of diagrams, including the Class diagram, which shows the classes and associations between them, as well as the classes' names and attributes. To provide detailed information on the attributes of these classes, code lists may be applied. Generally, a standard includes several code lists, the name of which again provides the base for a number of code values, each with a label (a term) and possibly a definition (ISO, 2015a, 6.5.1–3; cf. NISO, 2017; European Commission, 2018). Selic (2004) asserts that from a semantic point of view, a problem with UML is that the information relating to semantics is scattered throughout the text making it difficult to obtain a global understanding of how the various fragments fit together. Section 4.1 below develops on a solution.

2.2. Standards for Knowledge Organization Systems (KOSs)

There are a number of international standards concerning development and representation of KOSs, as well as for the mapping between cross-domain KOSs. KOSs can be represented by human-readable (e.g. HTML, CSV) and machine-readable formats (e.g. XML, RDF, OWL). For the latter, World Wide Web Consortium[12] (W3C) has specified Resource Description Framework[13] (RDF) which is a graph-based data model for expressing information about things (e.g. documents, people, physical objects, and abstract concepts) in a Semantic Web environment (W3C, 2014). Based on RDF technology, W3C has also developed a common data model, Simple Knowledge Organization Systems[14] (SKOS) for expressing the structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, and other similar types of KOSs. SKOS enables machine-readable representation of KOSs and allows sharing and linking different KOSs through Linked Data approach. Fig. 2, below, illustrates a part of a SKOS file. It

[11] https://www.omg.org
[12] https://www.w3.org
[13] https://www.w3.org/RDF/
[14] https://www.w3.org/2004/02/skos/

```
<skos:Concept rdf:about="http://www.cadastralvocabulary.org/CaLAThe/Surveyor">
   <skos:prefLabel>Surveyor</skos:prefLabel>
   <skos:changeNote>Established with version 1.</skos:changeNote>
   <skos:exactMatch rdf:resource="http://aims.fao.org/aos/agrovoc/c_26e30c73"/>
   <skos:inScheme rdf:resource="http://www.cadastralvocabulary.org/CaLAThe"/>
   <skos:broader rdf:resource="http://www.cadastralvocabulary.org/CaLAThe/Professional"/>
   <skos:narrower rdf:resource="http://www.cadastralvocabulary.org/CaLAThe/LicensedSurveyor"/>
   <skos:narrower rdf:resource="http://www.cadastralvocabulary.org/CaLAThe/LI_LandSurveyor"/>
   <skos:narrower rdf:resource="http://www.cadastralvocabulary.org/CaLAThe/LA_CertifiedSurveyor"/>
</skos:Concept>
```

Fig. 2. Excerpt of a SKOS file, which refers to data at http://www.cadastral vocabulary.org/CaLAThe/. The file refers to a thesaurus, as the information links the concept Surveyor to broader Professional, and three narrowers: LicensedSurveyor, etc. The concept is defined by stating exactMatch with an AGROVOC concept.

realizes Linked Data principles by making information on concept relations available on the web with an open license, identifying concepts with URIs, and applying open standards, providing machine-readable structured data (W3C, 2009a, 2009b; Baker et al., 2013).

INSPIRE (2017a) proposes use of SKOS for modelling of INSPIRE registers and register items. Similarly, a draft guideline for the RDF encoding of spatial data sets in INSPIRE (INSPIRE, 2017b), suggests that INSPIRE code lists - and extensions - shall be represented as SKOS concept schemes, and their codes as SKOS concepts.

In addition to these specifications, ISO has published many standards related to different aspects of terminology, such as those organized under the ISO 01.020 Terminology (Principles and Coordination).[15] Among these standards ISO 704:2009, ISO 10241-1:2011, ISO 10241-2:2012, ISO 860:2007, ISO 19146:2018 and ISO 19135–1:2015 have potential for the development and management of a harmonized terminology for the domain of land administration (ISO, 2009, 2011, 2012a, 2007b, 2018, 2015b). ISO 704:2009 'Terminology work - Principles and methods' defines the basic principles and methods for preparing and compiling terminologies (ISO, 2009). It focuses on concepts and their characteristics (e.g. necessary, sufficient, and essential characteristics), concept relationships (e.g. hierarchical, associative relations) and concept designations (e.g. terms, symbols). ISO 704:2009 also presents detailed specifications for concept definitions, and states that intensional definitions are the most explicit and precise method of concept definition (ISO, 2009, p. 22). An intensional definition consists of two parts. A basic part states the superordinate concept to which the concept belongs (genus), and a second part enumerates the delimiting characteristics (differentia), which distinguish this concept from its coordinate or related concepts (ISO, 2011, p. 26).

2.3. KOS registries

Over the last decade, a large number of KOSs has been published online. In addition, terminology registries have been developed to list, describe, identify and point to sets of vocabularies available for use in information systems and services. These registries allow discovery of suitable schemes for information or, potentially, use, by exposing rich metadata about them for navigation and retrieval (Golub and Tudhope, 2009; Ledl and Voß, 2016). Here, only the most relevant are mentioned.

The Basel Register of Thesauri, Ontologies and Classifications[16] (BARTOC) is a terminology registry developed by Basel University Library to describe KOSs in a uniform way, visualize them and to make them browsable for humans, foster interoperability and machine readability by utilizing Semantic Web standards. Currently, BARTOC groups metadata of more than 2800 KOSs and 87 other terminology registries in one place (Waeber and Ledl, 2019). BARTOC is a general or reference registry that covers a huge amount of KOSs from different domains, yet there are also various registries developed within the geographic information domain, as briefly introduced below.

INSPIRE[17] is a European Union (EU) initiative which aims at establishing an infrastructure for spatial information in Europe to support environmental policies and activities. The INSPIRE infrastructure involves a number of items, e.g. themes, code lists, application schemas or discovery services. Based on the ISO 19135-1:2015 standard, a number of INSPIRE registers have been developed for assigning unique identifiers to and consistently managing different versions of items used in the INSPIRE infrastructure. These include application schema register, code list register, enumeration register, feature concept dictionary, glossary, layer register, media-types register, metadata code list register, reference document register and theme register. Moreover, INSPIRE (2017a) provides general guidance and best practices for setting up registers supporting INSPIRE implementation. Accordingly, some national INSPIRE registries which extend INSPIRE vocabulary according to national requirements were developed, e.g. Italian INSPIRE Registry, Austrian INSPIRE registry and BRGM Registry in France, GDI-DE Registry in Germany.

The European Petroleum Survey Group[18] (EPSG) Geodetic Parameter Registry[19] provides an online repository for parameters required to define Coordinate Reference Systems (CRSs) and transformations between CRSs. Its geodetic model has been developed in accordance with ISO 19111:2007 'Geographic information – Spatial referencing by coordinates' (ISO, 2007a) and has been implemented in GML through ISO 19136 'Geographic information – Geographic markup language (GML)' (ISO, 2020). The EPSG Registry is maintained by the Geodesy Subcommittee of International Association of Oil and Gas Producers' (IOGP) Geomatics Committee (OGP, 2016).

The ISO Technical Committee 211 Geographic information/Geomatics[20] provides for a number of semantic resources. The ISO Geodetic Registry[21] is a database, which provides parameters defining global and regional CRSs and transformations between these CRSs. Moreover, Geolexica[22] is an online glossary for geographic information technology. Terminology entries in Geolexica fully reflect entries given in the ISO/TC 211 Multi-Lingual Glossary of Terms (MLGT), where its authoritative English terms originate from ISO/TC 211 standards.

The OGC Definitions Server is an online registry that allows for the management of resources such as terms, definitions, vocabularies and other related resources that are defined in OGC standards. The resources

[15] https://www.iso.org/ics/01.020/x/
[16] https://bartoc.org
[17] https://inspire.ec.europa.eu
[18] http://www.epsg.org
[19] http://www.epsg-registry.org
[20] https://www.isotc211.org
[21] https://geodetic.isotc211.org
[22] https://www.geolexica.org

are recorded in registers that conform to Linked Data principles and are published to provide for definitions service. The Definitions Server is intended to facilitate semantic interoperability between different systems that use OGC standards.[23] It was announced in 2018 and is being extended to contain much of the knowledge incorporated in OGC documents as well as to provide a single reference site for CRSs definitions, Discrete Global Grid Systems (DGGS), sensor models, and other specialized catalogues or ontologies.[24]

3. CaLAThe – Development and motivation for changes

The Cadastre and Land Administration Thesaurus, shortly CaLAThe, is a first domain-specific thesaurus that presents and relates core concepts of the cadastral domain both from the legal and technical aspects in the SKOS format. It was originally intended as a core terminology, to alleviate terminological inconsistencies and support the development of a coherent and universal cadastral theory or cadastral ontology (cf. Çağdaş and Stubkjær, 2009). However, in its present state, CaLAThe is not intended to be normative, differently from the use of thesauri in library settings, where thesauri assist in directing the location and the search for literature. Rather, CaLAThe accommodates related concepts, until harmonization is achieved and thereby constitutes a semantic tool for the alignment of standards in the land administration domain, cf. Fig. 3 below (Stubkjær et al., 2018; Stubkjær et al., 2019).

CaLAThe's initial version was based on the then draft version of LADM. However, adopting a Linked Data principle, it included also terms from other thesauri, such as the AGROVOC,[25] the GEMET[26] with INSPIRE Spatial Data Themes,[27] the STW Thesaurus for Economics,[28] Cycorp's OpenCyc ontology, and United Kingdom's Integrated Public Sector Vocabulary.[29]

LADM renders a static, database-derived view of the land administration domain, which marked version 1 of CaLAThe. The parts of CaLAThe were adopted from LADM, with the addition of a Documentation part, which reflect 'Source': 'document providing legal and/or administrative facts' (LADM, 4.1.21). Version 2 of CaLAThe, issued 2012, was supplemented with dynamic aspects. The extension was based on outcomes from the European research activity 'Modelling Real Property Transactions'[30] (ESF/COST G9, 2001–05) and a related Nordic project (Kort og Matrikelstyrelsen, 2006). Çağdaş and Stubkjær (2015a) provides for more information about the preparation and two first versions of CaLAThe. The extension during 2019 with about 10 LADM terms and about 30 terms adopted from the Land Division and Condominium parts of LandInfra (version 3), and with about 25 survey-related terms (version 4) raised the number of CaLAThe terms to almost 250, organized through its top concepts that characterize the domain covered. The six top concepts of CaLAThe parts are: Activity, Information, Land, Law, Party, and Survey.

Generally, the 2019 extension was unproblematic as the mentioned parts of LandInfra in some cases suggested more appropriate terms for existing concepts. Thus, the existing term, Administrative feature, was replaced with Land Division, which is precisely the feature addressed. Similarly, the existing term, Spatial representation type, was replaced with Spatial unit, as Spatial unit is the abstract entity which provides the conceptual base for an array of boundary representations in 2 or 3 dimensions.

The 2019 extension however included a more basic change: the

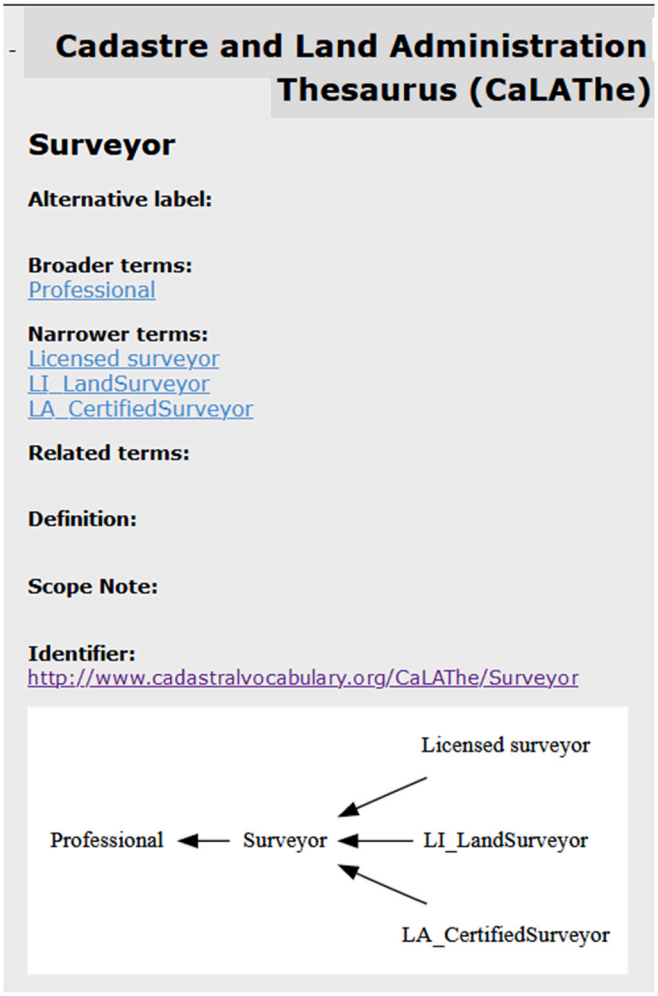


Fig. 3. Screenshot excerpt of CaLAThe term Surveyor, as rendered by a dedicated web server at Yildiz Technical University, Istanbul, Turkey.

Documentation part of CaLAThe was restructured by introducing Information as a top concept in version 4. For version 1, a top concept denoted Documentation was motivated by claiming that land administration is based on documents (Çağdaş and Stubkjær, 2015a). LADM refers to documentation in more packages. 'We prefer to establish an additional concept collection under the heading of Documentation, which is adopted from the STW thesaurus [for Economics]. It refers to 'Information science' as a broader term.' The concept collection of Documentation unfolds, among others, 'in terms of various data collections, e.g. Map and Land registry, subsumed under the heading of Spatial Data Infrastructure.' (Çağdaş and Stubkjær, 2015a).

From the chosen cadastre and land administration perspective, it seems defendable to consider spatial data infrastructure as a means or documentation resource. However, the SKOS thesaurus notation does not allow for such qualification of concept relations, as the semantic relations are restricted to either broader or narrower (both with transitive) or related (W3C, 2009a). Therefore, as information and next spatial data infrastructures have many more application domains than cadastre and land administration, these concepts must be located high in the concept hierarchy. As mentioned above in Section 2.1, the *ontology* knowledge structure allows for more varied concept relations.

CaLAThe is encoded as a SKOS, cf. Section 2.2. It is presently based on a dedicated web server, drawing SKOS-structured content from a

[23] https://github.com/opengeospatial/NamingAuthority
[24] http://www.opengeospatial.org/blog/2922
[25] http://aims.fao.org/vest-registry/vocabularies/agrovoc
[26] https://www.eionet.europa.eu/gemet/en/about/
[27] https://inspire.ec.europa.eu/Themes/Data-Specifications/2892
[28] http://zbw.eu/stw/version/latest/about
[29] http://id.esd.org.uk/list/subjects
[30] http://costg9.plan.aau.dk

MySQL database by means of php code. Moreover, concept relations are depicted as graphs, provided through the GraphViz[31] visualization software, as illustrated in (Çağdaş and Stubkjær, 2015a). The thesaurus is available online at cadastralvocabulary.org, which allows for term search and alphabetical listing of terms. Each concept is rendered in the context of its broader and narrower concepts, see Fig. 3. Finally, the development of CaLAThe is outlined in Version overviews.

The SKOS representation of CaLAThe is hosted by the BARTOC system, and can be accessed through BARTOC Skosmos Browser.[32] As of December 2019, it is also published on the OGC Definitions Server,[33] see Fig. 4. Again, the Linked Data principle of linking to related semantic resources is manifestly demonstrated. As of December 2020, CaLAThe was recorded byfairsharing.org, which was mentioned above in Section 2.1. (see on CaLAThe, https://fairsharing.org/bsg-s001551/).

As unfolded in Section 2, concept relations are essential to a thesaurus, while concept definitions are optional. The first versions of CaLAThe had not many definitions, and the few present were often taken from the mentioned existing thesauri. From version 3 this has changed, as more definitions are provided, while CaLAThe concepts increasingly are related to concepts in other thesauri through the predicates skos: exactMatch and skos:closeMatch.

4. A mediating platform for code list management

ISO LADM and OGC LandInfra include a number of code lists, which are not structured semantically because their role is to supplement the basic classes and relations of standards. Table 1, above, compared a selection of related code lists. Paasch et al. (2015), Stubkjær et al. (2018), and van Oosterom et al. (2019) call for more explicit semantics of code list values. The following briefly describes efforts towards semantic management of code lists specified by LADM and LandInfra through CaLAThe. It also presents the outcome of experiments aiming at harmonization of land administration related standards at code list level.

As CaLAThe provides for a semantic structuring of the terms applied by both LADM and LandInfra, an extension of the CaLAThe vocabulary with the name of the various code lists and their corresponding labels might assist the user community to obtain more consistency within the domain. The SKOS recommendation includes a) a provision for handling 'concept collections', which compares to the hierarchical structure of code list values. However, elements of the basic SKOS 'concept schemes' cannot refer to elements of 'concept collections'. b) The SKOS recommendation includes an Appendix B on SKOS eXtension for Labels (SKOS-XL), which might be applied for labelling code list values systematically; a similar approach would be use of the skos:notation option. This option has to be supplemented with domain-specific labelling conventions. Finally, c) the SKOS recommendation's 'concept scheme' marks an aggregation of one or more SKOS concepts with their semantic relationships. However, it is allowed to establish more 'concept schemes' addressing the same set of SKOS concepts, and thereby allow for both the network view and the hierarchical (code list) view of the same set. Experiments, testing how code lists could best be integrated into the SKOS-based CaLAThe suggested the adoption of the third option, namely the use of one concept scheme for the domain vocabulary, and an accompanying concept scheme for the related code lists, both from the LADM and the LandInfra standards.

The code list names and values address the same domain as the CaLAThe vocabulary; therefore, the objective was to integrate the code list names and values as far as reasonable, that is: adding code list values directly to CaLAThe as ordinary concepts with broader, narrower and optional definition. Some code list values are added with prefixes, indicating the pertinent standard, e.g. LI_ for LandInfra, and LA_ for LADM, awaiting mutual alignment. If the code list value is a term, which cannot be properly defined, e.g. OtherParty, this value is not added to CaLAThe, but is recorded only in the code list concept scheme and without prefix. Three LADM code lists with values are included as proof of concept in the present version, comprising both CaLAThe proper and the code lists, as demonstrated by Fig. 5, which presents code values of the LA_PartyRoleType code list.

The LandInfra parts of CaLAThe proper: Activity, Information, Land, etc. are visible as well. The intension is to add the code lists of the revised LADM in a subsequent version of CaLAThe.

The proposed approach to increase the semantics of code list values leaves an open issue, because in a thesaurus, concept relations are hierarchical: broader or narrower. Opposite, the values of a list do not have to comply with such restriction. For example, the LA_PartyRoleType list includes both surveyor and certifiedSurveyor, as well as bank and moneyProvider (see Table 1). Fig. 4 illustrates a rendering, which is faithful to the source, while Fig. 5 conveys the thesaurus view, that Surveyor is broader than LA_Certified Surveyor. Both views are obtainable in CaLAThe; the more complex thesaurus view increases the semantic information.

The code lists are available in several ways: a) In a file in SKOS format, available at request/downloadable from the CaLAThe site, b) the SKOS file visualized through SKOS Play!, which is a visualization tool for controlled vocabularies[34], and c) through the BARTOC Skosmos Browser,[35] and d) through the OGC Definitions Service.[36] Fig. 5 is a screen dump of the BARTOC presentation of CaLAThe with code lists. Similarly, the Definitions Server renders the values corresponding to typed or selected code list names. The Server allows for several views: all, basic, description, SKOS. Being simple and complete, the SKOS view is probably the most adequate, cf. Fig. 4.

5. Discussion

The fact that users worldwide have access to the code lists of domain standards, raises the question of how to facilitate use and management of code lists at the national level. The land administration domain has a noted administrative and judicial component. Therefore, standards within this domain are likely to be implemented through provisions provided by a land administration (cadastre, land registry) agency, e.g. in the context of recurrent overhaul of existing information systems or establishment of new. Code list management thus has to provide the agency and supporting companies and NGOs with an overview of available code list options and related information.

A 'code list management body', staffed with standardization and domain expertise, may be established at international and/or at regional level.

The OGC Land Administration DWG[37] (Domain Working Group) was established in 2016. The charter members of the DWG 'seek to identify enabling standards and best practices to guide countries in a programmatic way to establish more cost effective, efficient and interoperable land administration capability, to upgrade current manual to semi-automated processes, and to suggest solutions that are more automated and flexible to new data sources technologies' (Reichardt and Soliz, 2016). An OGC White Paper on Land Administration (OGC, 2019) refers to the need for registries to provide code lists, and states that 'ISO and OGC will cooperate on LADM in order to accelerate development. Apart from the development of the next edition, there will be attention to the management of code lists in order to include formal ontology and semantics' (5.3).

[31] https://www.graphviz.org
[32] https://bartoc-skosmos.unibas.ch/CaLAThe/en/
[33] https://www.opengis.net/def/CaLAThe
[34] http://labs.sparna.fr/skos-play/
[35] https://bartoc-skosmos.unibas.ch/CaLAThe/en/
[36] https://www.opengis.net/def/CaLAThe CodeList/4.0/CodeList
[37] https://www.ogc.org/projects/groups/landadmin

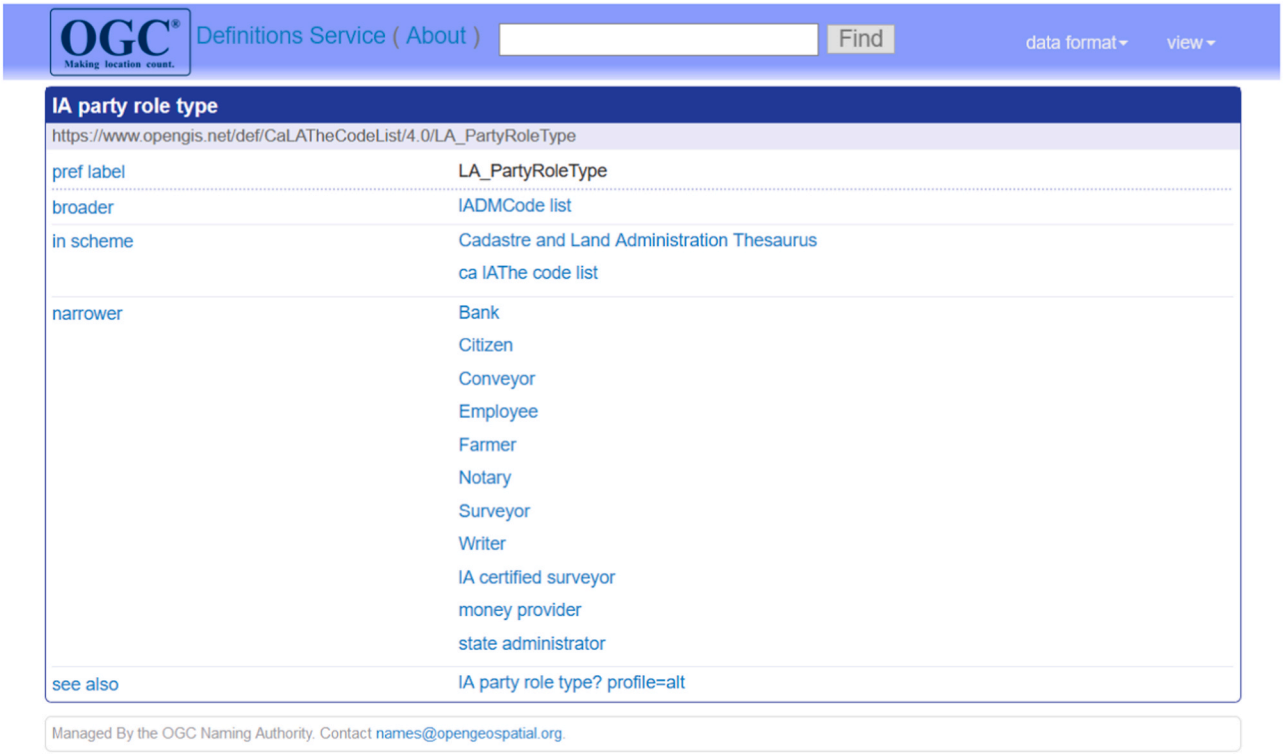


Fig. 4. Screenshot excerpt of CaLAThe with code lists, as rendered by OGC in SKOS view. The LADM code list PartyRoleType is shown. (lA should be read as LA).

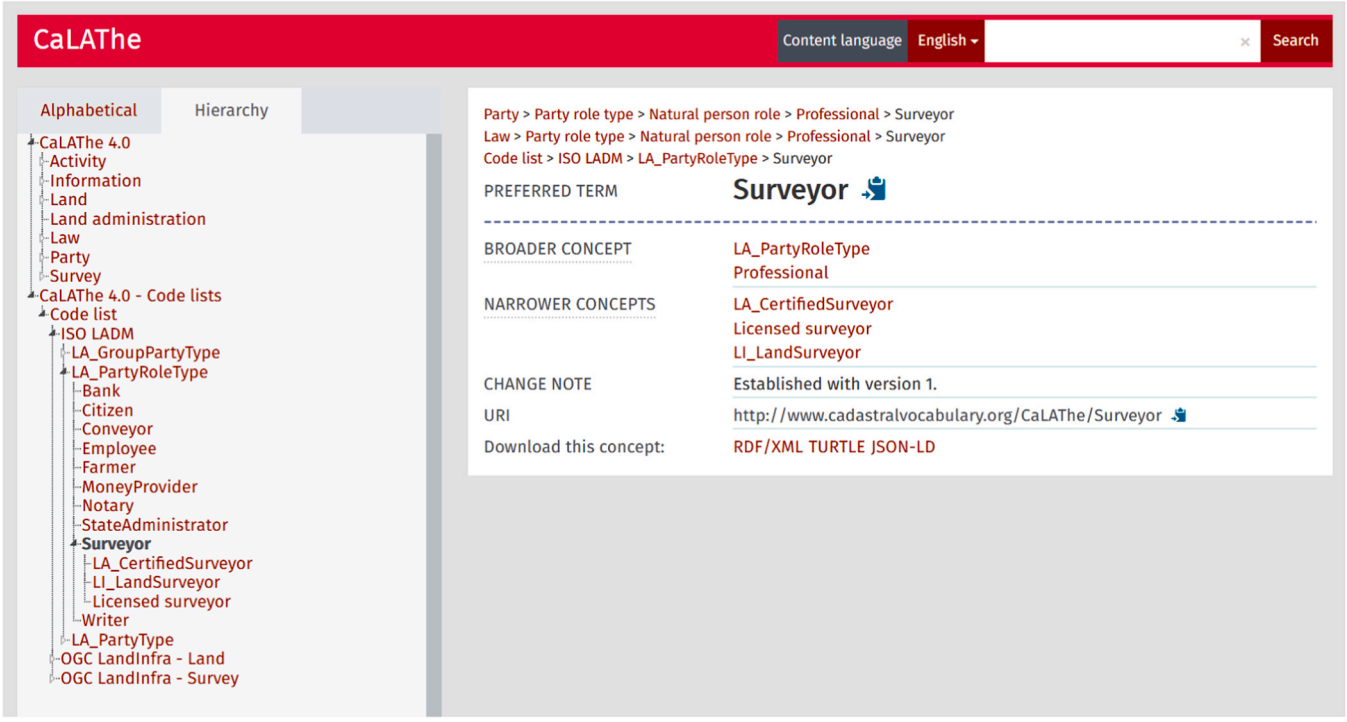


Fig. 5. Screenshot excerpt of CaLAThe with code lists, as rendered by BARTOC Skosmos Browser. Clicking the black and white triangles enables navigation in the hierarchy.

Code list management may be framed by regional associations like the Intergovernmental Committee on Surveying and Mapping (ICSM), which is a Standing Committee of ANZLIC – the Australian and New Zealand Spatial Information Council,[38] as well as the Permanent Committee on Cadastre in the European Union[39] (PCC). For Latin America, the Comité Permanente sobre el Catastro en Iberoamérica similarly might stimulate cooperation in this issue. Moreover, large countries with federal governmental structure, like Brazil, Canada, China, Germany, India, Russia and the USA could pave the way for others in demonstrating the benefits of a joint unit for code list management. Generally, their states have a mandate to localize code list at their discretion, while economy of scale suggest a shared and interoperable solution.

CaLAThe covers only the land part of the Land and Infrastructure standard. A Dictionary of Construction, Surveying and Civil Engineering is available (Gorse et al., 2020). This dictionary provides isolated explanation of terms, but superficial tests suggest that the dictionary does not replace a thesaurus of the domain. Given the fact, that the Land and Infrastructure standard exists, it should be a manageable task to develop a thesaurus for the infrastructure domain, especially when the definitions of the LandInfra standard becomes available through the OGC Definitions Server.

The Integrated Digital Built Environment[40] (IDBE) joint working group under buildingSMART International[41] (bSI) and OGC have issued a report to coordinate the development of the relevant data standards for the domain of built environment, focusing on CityGML, IFC (Industry Foundation Classes regarding building and infrastructure data), and LandInfra (Gilbert et al., 2020). Among others, the report mentions 'Proposed action points for overcoming integration challenges' to '- Derive and make publicly available a shared vocabulary or definition dictionary'. The report develops on this issue, noting 'a shared vocabulary with agreed definitions could be introduced, although this carries the risk of breaking backwards compatibility; alternatively, a shared resource for identifying synonymous terms in different domain vocabularies might be more feasible.' The above description of CaLAThe, presenting related designations regarding the surveyor compares to the report's needs and thus further confirms the relevance of CaLAThe development.

6. Conclusion

Considering the diversity of interests, initiatives and technology advancements, the alignment of standards within a domain is an ongoing process. Within the domain of cadastre and land administration, the CaLAThe thesaurus has been established as a semantic tool, which facilitates the alignment process in two ways, reflecting the fact that standards comprise a set of classes with their relations, and are supplemented by a set of code lists.

As regards the code lists, the contribution of this article is to report the availability of CaLAThe, supplemented with the code lists of the relevant parts of LandInfra, while selected code lists of LADM are included as proof of concept. The availability of CaLAThe with code lists, both at the OGC Definitions Server and BARTOC, facilitates the management of national code lists. Regional associations are mentioned, which might support the implementation and coordination process.

As regards the core of the standards, CaLAThe is made up from concepts defined in both LADM and LandInfra. This implies that neighboring and perhaps overlapping concepts are presented next to each other, facilitating discussions on possible revisions. We conclude that a joint initiative is motivated for the revision of the definitions used in LADM and LandInfra according to ISO 704:2009 etc., observing that intensional definitions, based on genus-differentia, are the most precise method of concept definition.

The domain of cadastre and land administration is multidisciplinary, which may have encouraged CaLAThe's reference to a wide array of other semantic tools. The described successful development path of CaLAThe may thus be of inspiration for researchers of other domains, especially within the geospatial domain.

[38] https://www.icsm.gov.au/about
[39] http://www.eurocadastre.org
[40] https://www.ogc.org/projects/groups/idbesc
[41] https://www.buildingsmart.org

CRediT authorship contributor statement

Erik Stubkjær: Conceptualization, Methodology, Project administration, Writing - original draft, Writing - review & editing. **Volkan Çağdaş:** Methodology, Writing - original draft, Writing - review & editing.

Acknowledgements

As editors of CaLAThe we would like to thank Open Geospatial Consortium staff and community for their prompt support and generous contribution towards the alignment of standards in the domain of land administration. The authors would like to acknowledge anonymous reviewers for their constructive comments to this article.

References

Baker, T., Bechhofer, S., Isaac, A., Miles, A., Schreiber, G., and Summers, E., 2013. Key choices in the design of Simple Knowledge Organization System (SKOS). Web Semantics: Science, Services and Agents on the World Wide Web, 20, 35–49. https://doi.org/10.1016/j.websem.2013.05.001 (Accessed 25 May 2020).

Breitman, K., Casanova, M.A., Truszkowski, W., 2007. Semantic Web: Concepts, Technologies and Applications. ASA Monographs in Systems and Software. Springer-Verlag, London.

Çağdaş, V., Stubkjær, E., 2009. Doctoral research on cadastral development. Land Use Policy 26 (4), 869–889. https://doi.org/10.1016/j.landusepol.2008.10.012.

Çağdaş, V., Stubkjær, E., 2015a. A SKOS vocabulary for linked land administration: cadastre and land administration thesaurus. Land Use Policy 49, 668–679. https://doi.org/10.1016/j.landusepol.2014.12.017.

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

EMMC-CSA, 2017. Report on first horizontal workshop on Metadata/Interoperability and Roadmap contributions. European Materials Modelling Council: Deliverable Report D2.11 for H2020-NMBP-CSA-2016 NMBP-24–206. Version final. https://emmc.info/wp-content/uploads/2018/01/EMMC-CSA_-D2.11_final-public.pdf. (Accessed 1 February 2021).

European Commission, 2018. Guidelines for the Use of Code Lists. https://joinup.ec.europa.eu/sites/default/files/document/2018-05/Guidelines%20for%20the%20Use%20of%20Code%20Lists%20v1.00.pdf. (Accessed 25 May 2020).

Gilbert, T., Rönsdorf, C., Plume, J., Simmons, S., Nisbet, N., Gruler, H.-C., Kolbe, T., van Berlo, L., Mercer, A., 2020. Built Environment Data Standards and Their Integration: An Analysis of IFC, CityGML and LandInfra. Open Geospatial Consortium & buildingSMART International. https://portal.ogc.org/files/?artifact_id=92634. (Accessed 25 May 2020).

Golub, K., Tudhope, D., Zeng, M.L., Žumer, M., 2014. Terminology registries for knowledge organization systems: functionality, use, and attributes. J. Assoc. Inf. Sci. Technol. 65 (9), 1901–1916. https://doi.org/10.1002/asi.23090.

Golub, K.,Tudhope, D., 2009. Terminology Registry Scoping Study (TRSS): Final report. http://www.ukoln.ac.uk/projects/trss/dissemination/trss-report-final.pdf (Accessed 25 May 2020).

Gorse, C., Johnston, D., Pritchard, M., 2020. A dictionary of construction, surveying, and civil engineering, 2nd edition.,. Oxford University Press, Oxford. ISBN: 9780198832485.

Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowl. Acquis. 5 (2), 199–220. https://doi.org/10.1006/knac.1993.1008.

Hodge, G., 2000. Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. The Digital Library Federation, Washington, DC. https://www.clir.org/wp-content/uploads/sites/6/pub91.pdf (Accessed 25 May 2020).

INSPIRE, 2017a. Best Practices for registers and registries & Technical Guidelines for the INSPIRE register federation. INSPIRE Maintenance and Implementation Group. https://inspire.ec.europa.eu/id/document/tg/registers-and-register-federation (Accessed 25 May 2020).

INSPIRE, 2017b. Guidelines for the RDF encoding of spatial data. http://inspire-eu-rdf.github.io/inspire-rdf-guidelines/ (Accessed 25 May 2020).

Terminological entries in standards — Part 1: General Requirements and Examples of Presentation, 2011. International Organization for Standardization (ISO), Switzerland.

Terminological Entries in Standards — Part 2: Adoption of Standardized Terminological Entries, 2012. International Organization for Standardization (ISO), Switzerland.

Geographic Information — Conceptual Schema Language, 2015. International Organization for Standardization (ISO), Switzerland.
Geographic Information – Spatial Referencing by Coordinates, 2007. International Organization for Standardization (ISO), Switzerland.
Geographic Information — Procedures for Item Registration — Part 1: Fundamentals, 2015. International Organization for Standardization (ISO),, Switzerland.
Geographic Information — Geography Markup Language (GML) — Part 1: Fundamentals, 2020. International Organization for Standardization (ISO), Switzerland.
Geographic Information — Cross-domain Vocabularies, 2018. International Organization for Standardization (ISO), Switzerland.
Geographic information – Land Administration Domain Model (LADM), 2012. International Organization for Standardization (ISO), Switzerland.
Terminology Work — Principles and Methods, 2009. International Organization for Standardization (ISO), Switzerland.
Terminology Work — Harmonization of Concepts and Terms, 2007. International Organization for Standardization (ISO), Switzerland.
Kless, D., Milton, S., Kazmierczak, E., Lindenthal, J., 2015. Thesaurus and ontology structure: formal and pragmatic differences and similarities. J. Assoc. Inf. Sci. Technol. 66 (7), 1348–1366. https://doi.org/10.1002/asi.23268.
Köpke, J., 2019. Annotation paths for matching XML-schemas. Data Knowl. Eng. 122, 25–54. https://doi.org/10.1016/j.datak.2017.12.002.
Kort og Matrikelstyrelsen, 2006. Property Formation in the Nordic countries - Denmark. Introduction and Comparison chapters, and chapters on Denmark. Danish Geodata Agency. https://gst.dk/media/2916021/propertyformationinthenordiccountries.pdf (Accessed 25 May 2020).
Ledl, A., Voß, J., 2016. Describing Knowledge Organization Systems in BARTOC and JSKOS. 12th International Conference on Terminology and Knowledge Engineering, Copenhagen, 22- 24 June (Unpublished, presentation). http://eprints.rclis.org/29366/1/Ledl_Voss_TKE2016_final_version_20160518.pdf (Accessed 25 May 2020).
NISO, 2017. Issues in Vocabulary Management. National Information Standards Organization, Technical Report TR-06–2017. https://groups.niso.org/topics/tl/BibliographicRoadmap/ (Accessed 25 May 2020).
Obrst, L., 2010. Ontological architectures. In: Poli, R., Healy, M., Kameas, A. (Eds.), Theory and Applications of Ontology: Computer Applications. Springer, Dordrecht, pp. 27–66.
OGC, 2016. OGC Land and Infrastructure Conceptual Model Standard (LandInfra). http://docs.opengeospatial.org/is/15-111r1/15-111r1.html (Accessed 25 May 2020).
OGC, 2017. OGC InfraGML 1.0 - Part 7 – LandInfra Land Division - Encoding Standard. https://docs.opengeospatial.org/is/16-107r2/16-107r2.html (Accessed 25 May 2020).
OGC, 2019. OGC White Paper on Land Administration. https://docs.opengeospatial.org/wp/18-008r1/18-008r1.html (Accessed 25 May 2020).
Paasch, J.M., van Oosterom, P., Lemmen, C., Paulsson, J., 2015. Further modelling of LADM's rights, restrictions and responsibilities (RRRs). Land Use Policy 49, 680–689. https://doi.org/10.1016/j.landusepol.2014.12.013.
Paasch, J., van Oosterom, P., Paulsson, J., Lemmen, C., 2013. Specialization of the Land Administration Domain Model (LADM) - An Option for Expanding the Legal Profiles. In: Proceedings of the FIG Work Week 2013, 6–10 May 2013, Abuja, Nigeria.
Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2008. A design science research methodology for information systems research. J. Manag. Inf. Syst. 24 (3), 45–77. https://doi.org/10.2753/MIS0742-1222240302.
Reichardt, M., Soliz, J.P., 2016. Land Administration (LandAdmin) Domain Working Group. OGC Project Document 16–082r3. https://www.ogc.org/projects/groups/landadmin (Accessed 25 May 2020).
Selic, B.V., 2004. On the Semantic Foundations of Standard UML 2.0. In: Bernardo, M., Corradini, F. (Eds.), Formal Methods for the Design of Real-Time Systems. Lecture Notes in Computer Science, Vol 3185. Springer, Berlin, Heidelberg. LNCS, volume 3185.
Shi, L., Roman, D., 2018. Ontologies for the Real Property Domain. 8th International Conference on Web Intelligence, Mining and Semantics, 25–27 June 2018, Novi Sad, Serbia. https://dl.acm.org/doi/pdf/10.1145/3227609.3227661 (Accessed 25 May 2020).
Sladić, D., Govedarica, M., Pržulj, D., Radulović, A., Jovanović, D., 2013. Ontology for real estate cadastre. Surv. Rev. 45 (332), 357–371. https://doi.org/10.1179/1752270613Y.0000000042.
Soon, K.H., 2013. Representing roles in formalizing domain ontology for land administration. 5th Land Administration Domain Model Workshop, 24–25 September 2013, Kuala Lumpur, Malaysia.
Stubkjær, E., Gruler, H.C., Simmons, S., Çağdaş, V., 2019. Code list management supported through a controlled domain vocabulary. 8th Land Administration Domain Model Workshop, 1 - 3 October 2019, Kuala Lumpur, Malaysia.
Stubkjær, E., Paasch, J., Çağdaş, V., van Oosterom, P., Simmons, S., Paulsson, J., Lemmen, C., 2018. International code list management - The Case of Land Administration. 7th Land Administration Domain Model Workshop, 11–13 April 2018, Zagreb, Croatia.
Stubkjær, E., Scarponcini, P., 2017. Harmonization of standards - The ISO 19152:2012 LADM-family. OGC March '17 Technical and Planning Committee Meeting - Delft, The Netherlands. https://vbn.aau.dk/files/268457453/Harmonization_20of_20standards_LandAdminDWG_20170321_.pptx (Accessed 25 May 2020).
van Oosterom, P., Kara, A., Kalogianni, R., Shnaidman, A., Indrajit, A., Alattas, A., 2019. Joint ISO/TC211 and OGC Revision of the LADM: Valuation Information, Spatial Planning Information, SDG Land Indicators, Refined Survey Model, Links to BIM, Support of LA Processes, Technical Encodings, and Much More on Their Way! FIG Working Week 2019, Hanoi, Vietnam, April 22–26, 2019.
Vandenbroucke, D., Olijslagers, M., Boguslawski, R., Borzachiello, M.T., Perego, A., Smith, R.S., 2020. Architectures and Standards for Spatial Data Infrastructures and Digital Government. Technical Report, European Commission Joint Research Center. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC121025/jrc121025_jrc121025_architectures_and_standards_for_sdis_and_digital_government.pdf (Accessed 26 October 2020).
W3C, 2009a. SKOS Simple Knowledge Organization System Reference. W3C Recommendation 18 August 2009. https://www.w3.org/TR/skos-reference/ (Accessed 25 May 2020).
W3C, 2009b. SKOS Simple Knowledge Organization System Primer. W3C Working Group Note 18 August 2009. https://www.w3.org/TR/skos-primer/ (Accessed 25 May 2020).
W3C, 2014. RDF 1.1 Primer. W3C Working Group Note, 24 June 2014. https://www.w3.org/TR/rdf11-primer/ (Accessed 25 May 2020).
Waeber, J., Ledl, A., 2019. A semantic web SKOS vocabulary service for open Knowledge Organization Systems. In: Garoufallou, E., Sartori, F., Siatri, R., Zervas, M. (Eds.), Metadata and Semantics Research, Vol. 846. Springer, Cham, pp. 3–12. https://doi.org/10.1007/978-3-030-14401-2.
Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (1), 1–9.
Zeng, M.L., Chan, L.M., 2004. Trends and issues in establishing interoperability among Knowledge Organization Systems. J. Am. Soc. Inf. Sci. Technol. 55 (5), 377–395. https://doi.org/10.1002/asi.10387.
Zeng, M.L., Clunis, J., 2020. FAIR+ FIT: guiding principles and functional metrics for Linked Open Data (LOD) KOS products. J. Data Inf. Sci. 5 (1), 93–118.
Zeng, M.L., Mayr, P., 2019. Knowledge Organization Systems (KOS) in the Semantic Web: a multi-dimensional review. Int. J. Digit. Libr. 2019 (20), 209–230. https://doi.org/10.1007/s00799-018-0241-2.