

SKOS-structured referencing between a domain thesaurus,
international standards, and large controlled vocabularies

- The case of CaLAThe

Workshop on Terminology and Shared Concepts

Virtual

Erik Stubkjær

13. April 2021

Cadastre and Land Administration Thesaurus (CaLAThe)



Professor emeritus, Dr.
Erik Stubkjær

Department of Planning,
Aalborg University, Denmark



Associate professor, Dr.
Volkan Çağdaş

Department of Geomatic Engineering, Yildiz Technical
University, Turkey

Editors with research area: Cadastre, Land administration, Immovable property law

Joint publications include:

- Alignment of Standards through Semantic Tools - The Case of Land Administration,
Land Use Policy, Volume 104, May 2021, 105381,
available for free at <https://authors.elsevier.com/a/1chc%7EyDvMI8nM>
until 24. April (and then via library)
- [A SKOS vocabulary for Linked Land Administration: Cadastre and Land Administration Thesaurus](#)
Land Use Policy, 49, 2015, p. 668–679
- [Core immovable property vocabulary for European linked land administration](#)
Survey Review, 47, 2015, 340, p. 49-60.

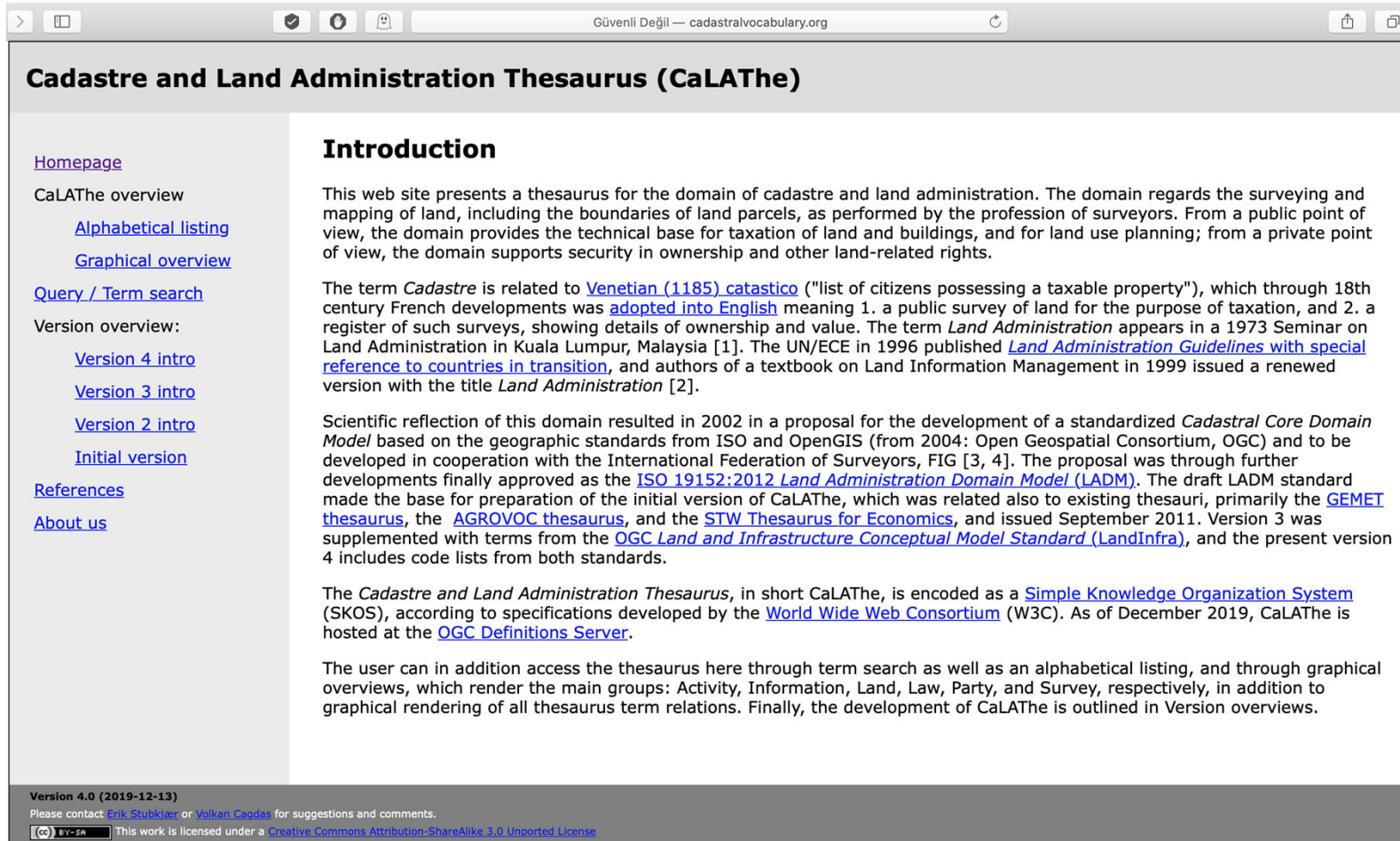
Agenda

1. Introduction to Cadastre and Land Administration Thesaurus (CaLAThe)
2. The issue of code lists
3. The present nesting of two SKOS Concept Schemes
4. Problem: Management of two related Concept Schemes
5. Other issues: BARTOC, schema matching, ..

1. CaLAThe - Introduction

- The first domain-specific thesaurus that represents cadastre and land administration concepts and their relationships in the SKOS format of W3C.
- Introduced 2011, based on then draft ISO Land Administration Domain Model (LADM).
 - Ver 2 added dynamic aspects.
 - Ver 3 supplemented with concepts from OGC Land and Infrastructure Conceptual Model Standard (LandInfra).
 - Ver 4 (Dec. 2019) survey-related terms and code lists of OGC LandInfra (Survey, Land division, and Condominium sections) and a few code lists of LADM
- Six top concepts: Land, Party, Law, Information, Survey and Activity; total ~ 250 concepts, relating to concepts from other thesauri, e.g. AGROVOC, GEMET, STW Thesaurus for Economics, ..

Available online at cadastralvocabulary.org



The screenshot shows a web browser window with the URL cadastralvocabulary.org. The page title is "Cadastral and Land Administration Thesaurus (CaLAThe)". The main content area is titled "Introduction" and contains three paragraphs of text. The first paragraph describes the website's purpose. The second paragraph discusses the historical context of the term "Cadastral" and its relation to French developments. The third paragraph describes the development of the CaLAThe thesaurus based on geographic standards and existing thesauri. A sidebar on the left contains navigation links. At the bottom, there is a footer with contact information and a Creative Commons license.

Cadastral and Land Administration Thesaurus (CaLAThe)

[Homepage](#)

CaLAThe overview

- [Alphabetical listing](#)
- [Graphical overview](#)

[Query / Term search](#)

Version overview:

- [Version 4 intro](#)
- [Version 3 intro](#)
- [Version 2 intro](#)
- [Initial version](#)

[References](#)

[About us](#)

Introduction


This web site presents a thesaurus for the domain of cadastre and land administration. The domain regards the surveying and mapping of land, including the boundaries of land parcels, as performed by the profession of surveyors. From a public point of view, the domain provides the technical base for taxation of land and buildings, and for land use planning; from a private point of view, the domain supports security in ownership and other land-related rights.

The term *Cadastral* is related to [Venetian \(1185\) catastico](#) ("list of citizens possessing a taxable property"), which through 18th century French developments was [adopted into English](#) meaning 1. a public survey of land for the purpose of taxation, and 2. a register of such surveys, showing details of ownership and value. The term *Land Administration* appears in a 1973 Seminar on Land Administration in Kuala Lumpur, Malaysia [1]. The UN/ECE in 1996 published [Land Administration Guidelines with special reference to countries in transition](#), and authors of a textbook on Land Information Management in 1999 issued a renewed version with the title *Land Administration* [2].

Scientific reflection of this domain resulted in 2002 in a proposal for the development of a standardized *Cadastral Core Domain Model* based on the geographic standards from ISO and OpenGIS (from 2004: Open Geospatial Consortium, OGC) and to be developed in cooperation with the International Federation of Surveyors, FIG [3, 4]. The proposal was through further developments finally approved as the [ISO 19152:2012 Land Administration Domain Model \(LADM\)](#). The draft LADM standard made the base for preparation of the initial version of CaLAThe, which was related also to existing thesauri, primarily the [GEMET thesaurus](#), the [AGROVOC thesaurus](#), and the [STW Thesaurus for Economics](#), and issued September 2011. Version 3 was supplemented with terms from the [OGC Land and Infrastructure Conceptual Model Standard \(LandInfra\)](#), and the present version 4 includes code lists from both standards.

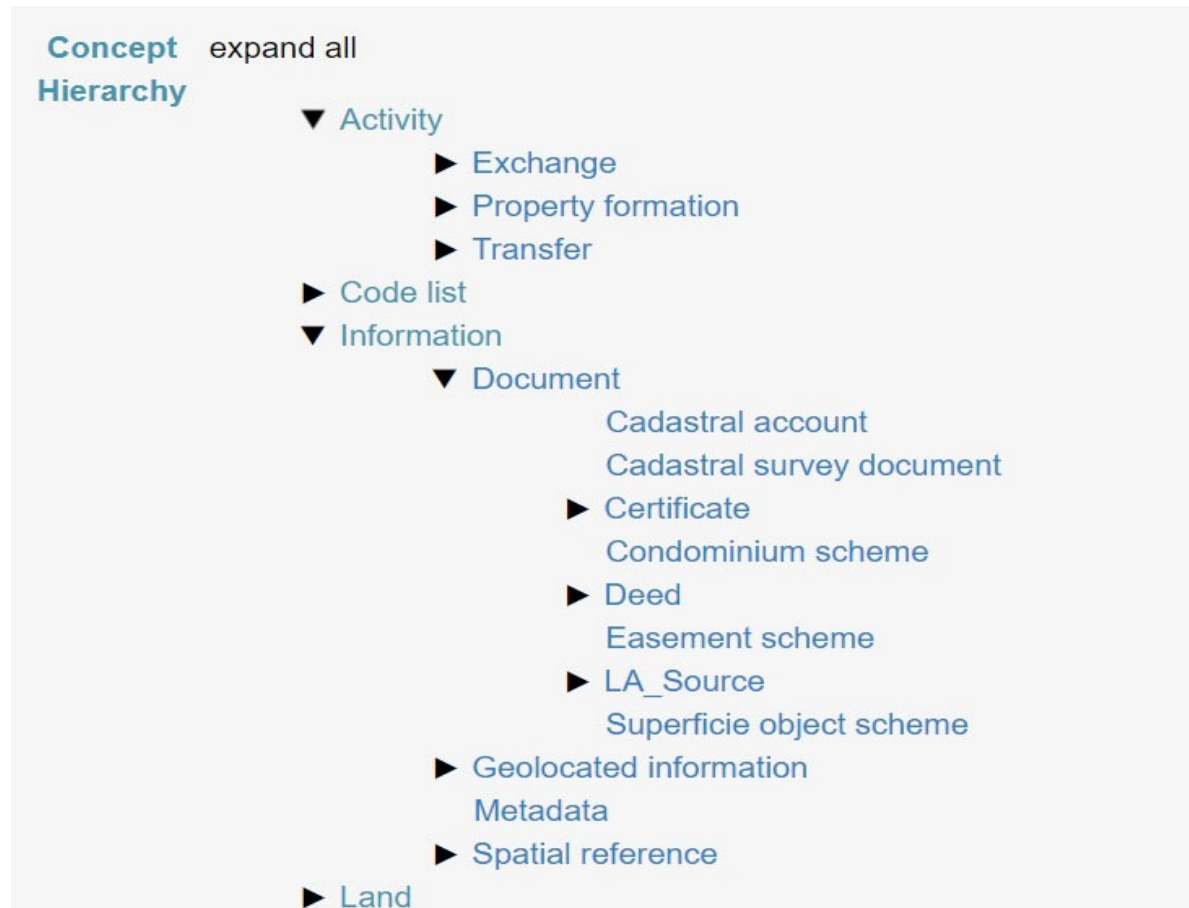
The *Cadastral and Land Administration Thesaurus*, in short CaLAThe, is encoded as a [Simple Knowledge Organization System \(SKOS\)](#), according to specifications developed by the [World Wide Web Consortium \(W3C\)](#). As of December 2019, CaLAThe is hosted at the [OGC Definitions Server](#).

The user can in addition access the thesaurus here through term search as well as an alphabetical listing, and through graphical overviews, which render the main groups: Activity, Information, Land, Law, Party, and Survey, respectively, in addition to graphical rendering of all thesaurus term relations. Finally, the development of CaLAThe is outlined in Version overviews.

Version 4.0 (2019-12-13)
Please contact: [Erik Stubkjær](#) or [Volkan Cagdas](#) for suggestions and comments.
 This work is licensed under a [Creative Commons Attribution-ShareAlike 3.0 Unported License](#)

Available online at <https://www.opengis.net/def/CaLAThe>

Screenshot
excerpt



Available online at <http://bartoc-skosmos.unibas.ch/CaLAThe/en/>

The screenshot displays the CaLAThe web interface. At the top, there is a red header with the text 'CaLAThe' on the left, a 'Content language' dropdown menu set to 'English', and a search box with a 'Search' button. Below the header, there are two tabs: 'Alphabetical' and 'Hierarchy'. The 'Alphabetical' tab is active, showing a hierarchical tree of concepts. The tree is organized as follows: CaLAThe 4.0 > Code list > ISO LADM > LA_PartyRoleType > Surveyor. The 'Surveyor' concept is highlighted in bold. The right side of the interface shows the details for the 'Surveyor' concept. It includes a breadcrumb trail: Party > Party role type > Natural person role > Professional > Surveyor. Below this, there are sections for 'PREFERRED TERM' (Surveyor), 'BROADER CONCEPT' (LA_PartyRoleType, Professional), 'NARROWER CONCEPTS' (LA_CertifiedSurveyor, Licensed surveyor, LI_LandSurveyor), 'CHANGE NOTE' (Established with version 1.), 'URI' (http://www.cadastralvocabulary.org/CaLAThe/Surveyor), and 'Download this concept:' (RDF/XML, TURTLE, JSON-LD).

2. The issue of code lists

- Code lists supplement the classes, attributes, relations, and methods of UML modeling
- Code lists have names, and a code list consist of a number of values: In the present context, the values are terms.
- The ISO Land Administration Domain Model (LADM) has to be supplemented by national profiles, to allow for LADM implementation (no InfraGML-like encoding)
- CaLAtHe was supplemented with code lists to provide a mediating platform.

Examples of codelists in LADM and LandInfra

LADM (Figure J.1)	LandInfra (Figure 65)
LA_PartyRoleType <ul style="list-style-type: none">• bank• certifiedSurveyor• citizen• conveyor• employee• farmer• moneyProvider• notary• stateAdministrator• surveyor• writer	SigningRole <ul style="list-style-type: none">• owner• landSurveyor• seller• buyer• neighbor• otherParty

CaLAtHe extension might support the user community in obtaining more consistency within the domain.

- 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.

Source: LUP, 2021; sec. 4

4. The present nesting of two SKOS Concept Schemes

- SKOS recommendation (W3C, 2009) offers
 - 4. Concept Schemes , and
 - 9. Concept Collections

The latter ‘compares to the hierarchical structure of code list values. However, elements of the basic SKOS ‘concept schemes’ cannot refer to elements of ‘concept collections’.’

The former ‘allow[s] 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.’

The two-Concept Scheme option applied in only one file

- `<!--` The following part contains code list names and values, and refers to two Concept Schemes, as demonstrated with `CodeListTest2208.rdf`, presented at OGC-NA Telecon on LandInfra Codelists, 22 August 2019. The code list names are all in the `CaLATheCodeList` scheme, but the code list values may refer to concepts of `CaLAThe` proper, that is from the `CaLAThe` scheme.
- Basically, the code list structure is hierarchical. This might suggest that `skos:Collection` is applied, and not `skos:ConceptScheme`.
- However, LandInfra could specify a number of values in terms of `/codeList` URI `[0,1]`, as specified for code list `SurveyMonumentType`, which would break the strict hierarchical structure. To keep flexibility, `skos:ConceptScheme` is chosen. `-->`

- `#` Additional `ConceptScheme`, reflecting the :

```
<skos:ConceptScheme rdf:about="http://www.cadastralvocabulary.org/v4_draft/CaLATheCodeList">
```

CaLAtThe CodeList rendered by OGC Def. Server/ VocPrez

CaLAtThe - Codelists

URI
<https://www.opengis.net/def/CaLAtTheCodeList>

Concept Hierarchy expand all

- ▶ Activity
- Administrative unit
- Association
- Cadastral parcel
- ▼ Code list
 - ▼ ISO LADM
 - ▶ LA_GroupPartyType
 - ▼ LA_PartyRoleType
 - Bank
 - Citizen
 - Conveyor
 - Employee
 - Farmer
 - LA_CertifiedSurveyor
 - MoneyProvider
 - Notary
 - StateAdministrator
 - ▶ Surveyor
 - Writer
 - ▶ LA_PartyType
 - ▶ OGC LandInfra - Land

4. Problem: Management of two related Concept Schemes

The rendering of the content of two Concept Schemes within one file has caused troubles. How can these be remedied?

- Structure the content into two files, one for each Concept Scheme
 - How can you refer between Concept Schemes/ files?
 - Is a SKOS formalism available for such referencing?
 - Have ad hoc solutions been found and reported?
- Not all hosts have reported of troubles
 - Ask host editors, how CaLAtHe with code list was integrated
- Is cross-reference between CaLAtHe and CaLAtHeCodeList needed?

5. Other issues: BARTOC, schema matching, ..

BARTOC

- The **Basic Register of Thesauri, Ontologies & Classifications (BARTOC)** is a database of Knowledge Organization Systems and KOS related registries. The main goal of BARTOC is to list as many Knowledge Organization Systems as possible at one place BARTOC includes any kind of KOS from any subject area, in any language, any publication format, and any form of accessibility. BARTOC's search interface is available in 20 European languages ..

The geospatial community ought to be visible there

Thank you !

Erik Stubkjær

est@plan.aau.dk

Aalborg University, Denmark

From: Erik Stubkjær
Sent: 13. april 2021 11:58
To: 'Rob Atkinson' <ratkinson@ogc.org>;

I wonder, whether it is SKOS compliant to list the name of a Concept Scheme within a Collection, as 'member'. If so, we could have

Collection CadastralVocabulary, with CaLAtthesaurus and CaLAcodelist, where the two terms refer to Concept Schemes in individual files.

Similarly, we could have

Collection LandVocabulary with 'mappings available against standards ' and also CadastralVocabulary

And moreover: Collection GeospatialVocabulary with LandVocabulary and much more

The latter seems to be SKOS compliant, cf. 9.6.3. Nested Collections <https://www.w3.org/TR/skos-reference/#collections>