International Code List Management – The Case of Land Administration

Erik STUBKJÆR, Denmark, Jesper M. PAASCH, Sweden, Volkan ÇAĞDAŞ, Turkey, Peter van OOSTEROM, the Netherlands, Scott SIMMONS, USA, Jenny PAULSSON, Sweden and Christiaan LEMMEN, the Netherlands

Key words: Code lists, Semantic Web, Land Administration Domain Model (LADM), ISO 19152, OGC Land and Infrastructure Conceptual Model Standard (LandInfra), OGC InfraGML

SUMMARY

Standardization reduces technical barriers to trade and foster dissemination of innovations. Within the domain of information technology, standardization enhances semantic interoperability of systems and services. In order to achieve the potential of standardization, IT solutions must be localized to adapt to local needs. To reduce localization costs, software developers, but also standards develop and adopt internationalization principles and best practices, cf. the W3C Internationalization (I18n) Activity, the ISO 639 Language Codes, which provides an example for code lists and code list management, and the coding of coordinate reference systems.

For the domain of Land Administration, the localization issue extends from language names to the various organizations and institutions dealing with interests in land. Paasch et al (2013) propose code lists as a mean of internationalization by which the classes of the ISO 19152:2012 Land Administration Domain Model (LADM) may be related to the concerned jurisdiction. The issue of code lists has been addressed by the OGC as well, namely in terms of the document 17-050r1 Code List Manifesto by Paul Scarponcini. Motivations for the study include that various OGC standards have encoded enumerations and code lists differently, as realized during the development of the InfraGML standard, which regards land and civil engineering infrastructure facilities, and thus share part of its scope with LADM.

Aiming at harmonization of standards within the domain of Land Administration, the present paper proposes a joint management of the code lists which are specified by ISO LADM and by OGC LandInfra / InfraGML, respectively. The FIG motivated the ISO standard LADM and moreover framed research on code lists. It seems therefore appropriate to join with this organization of surveying professionals, also to benefit from sharing of expertise and cost of the management activities. The paper the outlines the tasks of code list management by drawing on the mentioned Code List Manifesto and resuming research supporting code list management, e.g. terminological theory and semantic tools. The setup of a possible code list management system is discussed, and summarized in terms of a draft Memorandum of Understanding.

International Code List Management – The Case of Land Administration

Erik STUBKJÆR, Denmark, Jesper M. PAASCH, Sweden, Volkan ÇAĞDAŞ, Turkey, Peter van OOSTEROM, the Netherlands, Scott SIMMONS, USA, Jenny PAULSSON, Sweden and Christiaan LEMMEN, the Netherlands

1. INTRODUCTION

Standardisation is an international task, especially after the International Organization for Standardization (ISO) was founded in 1947. Generally, it reduces technical barriers to trade and foster dissemination of innovations. Within the domain of information technology, standardization enhances semantic interoperability of systems and services.

In order to achieve the potential of standardisation, IT solutions must be localized to adopt local language, format of dates and currency, etc. This implies that IT solutions, which are costly to localise, have lower potential to be reused across borders. To reduce localization costs, software developers, but also standards develop and adopt internationalization principles and best practises, cf. the W3C Internationalization (I18n) Activity. Taking language as an example, ISO 639 Language Codes provides an example for code lists and code list management in terms of a Joint Advisory Committee. Comparable to the coding of languages world-wide is the coding of coordinate reference systems (CRSs) applied world-wide. The International Association of Oil & Gas Producers (IOGP) provides for the EPSG Geodetic Parameter Dataset, where codes are assigned to CRSs, coordinate transformations, and their component entities (datums, projections, etc.). Further examples include code lists of the Electronic Data Interchange standards (GS1 EDI) for electronic business messaging, and the INSPIRE code list register, containing the code lists and their values, as defined in the INSPIRE implementing rules on interoperability of spatial data sets and services.

For the domain of Land Administration, the localization issue extends from language names to the various organisations and institutions dealing with interests in land. Paasch et al (2013) propose code lists as a mean of internationalization by which the classes of the ISO 19152:2012 Land Administration Domain Model (LADM) may be related to legal concepts of the jurisdiction concerned. LADM was proposed to ISO/TC211 by the International Federation of Surveyors (FIG) in January 2008, through a parallel voting in ISO TC211 and CEN TC287. At the ISO/TC211 plenary meeting in Wellington, New Zealand, the revision of LADM was furthered by appointment of the coordinator to initiate the Stage 0 project (ISO-TC211, 2017).

The issue of code lists has been addressed by the OGC as well (Scarponcini, 2017). Motivations for the study include that various OGC standards have encoded enumerations and code lists differently, as realized during the development of the InfraGML standard, which regards land and civil engineering infrastructure facilities, and thus share part of its scope with LADM. The Manifesto recommends that OGC assume ownership for code lists, including maintenance of the original list and deciding the format and location of the OGC lists.

Harmonization of standards within the domain of Land Administration was proposed at the OGC TC meeting, March 2017 in Delft (Stubkjær, Scarponcini, 2017). The present paper in addition draws on the mentioned concern for code lists, and proposes a joint management of the code lists specified by ISO LADM and by OGC LandInfra / InfraGML, respectively. The FIG motivated the ISO standard LADM and moreover framed research on code lists. It seems therefore appropriate to join with this organisation of surveying professionals, also to benefit from sharing of expertise and cost of the management activities.

The paper provides an introduction to code lists and elaborate on this, drawing on the mentioned Code List Manifesto (section 2). Research supporting code list management, e.g. terminological theory and standards, and semantic tools are resumed (section 3). Recent experiences with code list management within the Land Administration domain are reported, and components of a code list management system, potential users, and the organizational setup are outlined (section 4), which also attempts to assess the value chain of the proposed effort. A conclusion closes the paper.

2. CODE LISTS OF STANDARDS – CONCEPTS AND PURPOSE

2.1 Code list basics

The use of code lists has grown fairly recently, motivated by the wish to enhance interoperability among information systems. Code lists are used in the context of modelling information in a way that allows for exchange of information between systems, which are maintained by different parties, and where the models are understandable for both machines and humans. The Unified Modelling Language, adopted as a standard by the Object Management Group in 1997, is widely used for this modelling[1].

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. Some classes refer to objects in the real world; these classes are marked 'FeatureType'. It is often necessary to provide detailed information on the attributes of these classes. If the attribute can be measured, a number may provide this information, but otherwise words, concepts, terms or names have to be used. For this purpose, other classes or ‘classifiers’ are introduced, which are marked either 'enumeration' or 'Code list'. (The marking of UML classes is dubbed stereotypes). Thus, a code list provides information on the various attributes of an object, more specifically the observed or selected values of the attribute[2].

The idea of establishing code lists emerged in the 1960s in order to overcome limitations in storage space of the computers of the time. Names of countries, institutions, persons, or roads were listed and supplemented by a code of few digits. For example, in a file with e.g. bill records, each record would contain only the say 4-digit road code, while the full road name was provided through a corresponding code list. In the USA, this practice was standardized for description of books and other items catalogued by libraries in terms of the MAchine-

[1] https://www.omg.org/spec/UML/About-UML/
[2] https://github.com/ISO-TC211/UML-Best-Practices/wiki

Readable Cataloging (MARC) standard in 1971, which in 1999 was replaced by the international MARC 21. A related resource is the GeoNames database, which among others provides names of places in various languages[3].

The use of code lists is not trivial when the item coded has more names or interpretations. For example, an author may have one or more pseudonyms, and you would expect the library information system to guide you to the very author. This may be achieved by what in the MARC context is called an ‘authority file’, which relates the various pseudonyms of an author to the version of the name, which is adopted in the code list.

Not only librarians have to address the problem of interpretations of coded items. ISO 19103 Geographic information – Conceptual schema language, which provides rules and guidelines concerning geographic information, includes a Requirement 7, which reads: As the values of enumerated types are concepts, each value shall have a definition for the value. This requirement applies to the items of code lists as well[4]. Whether such definition should be included into the ‘authority file’ or be recorded otherwise is left open here. The essential is that the codes and the corresponding text or label need be supplemented with information, which informs on alternative renderings, on definitions, and on the bodies, who authorize the chosen wordings. Such body may be called a ‘code list authority’ (Greiner, U., 2006).

2.2 The OGC Code List Manifesto

The recent development of the OGC InfraGML standard motivated a review of the use of enumeration and code list classifiers in OGC standards. The review resulted in the Code List Manifesto by Paul Scarponcini, a document pending on OGC consideration (Scarponcini, 2017). The Manifesto provides guidelines for specification and use of enumerations and code lists. It does so by summarizing content of the above-mentioned ISO 19103:2015, which regards UML schemas describing geographic information, and by proposing four code list types or cases, based on an analysis of the OGC encoding standards GML 3.2.1 and GML 3.3. The analysis is not rendered here, only the cases with motivation.

Case 0 (No list)

For this case, there is no list of values. The property datatype is specified as UML CharacterString (XML “string”). The creator of a valid XML document based on the .xsd schema may enter any character string value of his choosing. Of course other data types, such as integer follow the same situation, if there is no restriction on a range of allowable integer values. For example, in InfraGML Part 0, a professional’s company (the optional name of the company through which the professional offers service) is specified as being of type “string”.

Case 1 (Enumeration)

For this case, the list of values is complete (non-extensible) in the .xsd schema for a given version of the encoding standard. The stereotype of the UML classifier will most likely be <<Enumeration>> but may be <<CodeList>>, but content shall only be the complete list of

[3] http://www.geonames.org/about.html
[4] https://github.com/ISO-TC211/UML-Best-Practices/wiki/Definitions

values. The creator of a valid XML document based on the .xsd schema shall select only values from the schema encoded list.

For this case, changes to the list can only be made via updates to, or extensions of, the standard containing the .xsd schema definition. A corrigendum can alter any erroneous values. Revisions to the standard can add new values to the list and can then be backward compatible.

It should be noted that for situations where the UML uses the stereotype <<CodeList>>, but where the intent is that the list cannot be extended either by an XML document creator or an external organization, then the recommendations in Case 1 shall apply. This is especially true when the code list values differentiate which other parts of the standard apply. For example, the LandDivision SpatialUnit.dimension, satisfied by the DimensionType <<CodeList>> in the LandInfra UML, restricts the types of BoundaryElement a SpatialUnit can use.

<u>Case 2 (Union Type)</u>
For this case, the list of values contains none or only a few of the allowed (known to the .xsd schema) values, such as a likely set, or an initial set. The stereotype of the UML classifier shall be <<CodeList>>. The content of the classifier shall only be the list of values known when the .xsd schema is standardized. Therefore, in accordance with ISO 19103, no tagged value "codeList" is present.

For this case, the creator of a valid XML document based on the .xsd schema can freely add additional unique values. These additional values should not replace an existing code by changing the name or definition, or have the same definition as an existing value.

The classifier shall be encoded as the union of an enumeration type and a pattern. The enumeration type covers the UML listed values and is encoded as in Case 1 above. The pattern allows the user to define an unlisted value, preceded by "other: ".

<u>Case 3 (Reference Type)</u>
For this case, the list of values contains none or only a few of the allowed (known to the .xsd schema) values, such as a likely set, or an initial set. The stereotype of the UML classifier shall be <<CodeList>>. The content of the classifier can be the list of values known when the .xsd schema is standardized. In accordance with ISO 19103, a tagged value "codeList" is present as the last entry on the list to signify that an external authority manages the allowable list of values. GML3.3 stipulates that if the classifier has an "asDictionary" tagged value, then the value shall be "true". Then the GML 3.3 asDictionary-type Requirement holds, stating that the code list shall be represented by an external dictionary, vocabulary or ontology, using any suitable syntax or encoding.

For this case, the creator of a valid XML document based on the .xsd schema is limited to the values on the external authority list. This of course begs the question of why any values were also listed in the UML, unless the UML list was intended to be offered by some authority, perhaps the OGC. Definitions provided by external authorities may already be packaged for delivery in various ways, both online and offline. In order that they may be referred to from GML documents it is merely necessary that a URI be available to identify for each definition.

In the .xsd schema, the GML 3.2.1 ReferenceType is used. In response, the XML would provide an xlink href to the external list URI and enumerated values. For additional clarity, an xlink title can provide a more meaningful form of the selected value.

The above description summarizes knowledge on the meaning and use of enumerations and code lists. The Manifesto provides examples on the corresponding coding of .xsd schemas.

3. CODE LIST MANAGEMENT – THEORIES, TOOLS AND APPLICATIONS

3.1 Terminological theory

Code lists contain words to be communicated, either by humans or technical systems. These words only make sense in communication, if they are understood by both parties in the communication. Terminology is therefore of vital importance for any communication attempt. In order to achieve a thorough understanding of a fact, a semantic network of events, or a problem, the understanding must comprise not only the nature of the case and what it consists of. The understanding must also draw our own thoughts and mental images, as the symbolism employed when speaking is, according to Ogden and Richards (1923), partly caused by the reference that is made to the fact and partly by social and psychological factors. Furthermore, different professions may use a specialised terminology. For example, professional legal languages may even differ within themselves (Jackson, 1995). The legal domain is therefore not a homogenous body, but a patchwork of different legal domains based on different national legislation and cultural heritage.

Therefore, the basic terminological components used in communication through e.g. standardized code lists have to be addressed, namely: object, concept, characteristic, definition and term. They are closely interrelated. An *object* is anything that is conceivable or perceivable. Objects can be material (e.g. a piece of land), immaterial (e.g. an urban planning zone) or imagined (e.g. a unicorn). A *concept* is a mental construction of the real world formed in our own mind. A concept, however, does not stand alone, but is part of a system of concepts which are related to each other according to specific rules. It is the *characteristics* of objects which make us identify the ‘real world’, when we create our vision of it in our mind as a concept. Objects, concepts, and characteristics are not sufficient to communicate effectively, because the concept is bound to each individual. Therefore, we must describe what we mean, i.e. produce a definition describing the (mental) concept we have in our minds. A *definition* has to be as precise as possible to avoid misunderstandings and confusions. It would however be rather tiresome and complicated always to use definitions when communicating. On the other hand, ambiguity of words makes it difficult to express precisely what is meant and to ensure the correct understanding of the texts and diagrams describing the topic subject for the description. We therefore communicate through *terms* to express the definitions. A term must have a specific meaning, based on the definition describing and delimiting a concept. It would otherwise mean different things to different people. Any term has to be based on the discussion of our mental pictures of real world objects, and delimited by characteristics which are mandatory for the object in question (Suonuuti, 2001 and Annex 3).

3.2 Terminology standards

The definitions derived from terminological theory provide a basis for terminology standards, some of which accounts for that concepts do not 'stand alone, but [are] part of a system of concepts which are related to each other ...', cf. above. The code lists specified by ISO LADM and OGC LandInfra present simple list of terms or values for various attributes. These code lists do not include alternative terms, definitions of terms, or semantic relationships between the terms (e.g. hierarchical, associative). The more sophisticated Knowledge Organization System (KOS) supported by semantic technologies (e.g. SKOS, Linked Data), however, provide more comprehensive schemas for the specification of code list values (cf. Paasch et al, 2015). A KOS is a general term which refers to tools that present the organized interpretation of knowledge structures (Zeng, 2004, p. 377). It covers all types of schemes for organizing information and promoting knowledge management, such as (i) term lists (e.g. glossaries, dictionaries and gazetteers), (ii) classifications and categories (e.g. subject headings and taxonomies), and (iii) relationship lists (e.g. thesauri and ontologies) (Hodge, 2000, p. 4). They are also referred to as controlled vocabularies, structured vocabularies, value vocabularies, concept schemes, semantic, and classification by various standards (Golub et al, 2014, p. 1902).

There are a number of international standards for the development of above mentioned different types of KOSs. ANSI/NISO Z39.19-2005 is an American standard related to all kind of monolingual controlled vocabularies, whereas ISO 25964-1:2011 and ISO 25964-2:2013 are international standards specified for the development and maintenance of thesauri (both monolingual and multilingual), and other types of controlled vocabularies, respectively. ISO 25964-1:2011 includes a data model and an XML schema for data exchange, but ANSI/NISO Z39.19-2005 does not provide a data model, nor addresses multilingual vocabularies or other aspects of interoperability problems, such as mapping between KOSs (Chatterjee, 2016, p. 487).

ISO has also published standards related to different aspects of terminology (see Appendix 3). Moreover, ISO provided a number of domain vocabularies organized under the ISO 01.040 Standard Catalogue for different domains, e.g. Natural and applied sciences, Metrology and measurement, Information technology, Agriculture[5].

As for the development of terminology or KOS within geographic information domain, ISO 19135-1:2015 'Geographic information – Procedures for item registration – Part 1: Fundamentals' has been published. This standard specifies procedures to be followed in establishing, maintaining, and publishing registers of unique, unambiguous and permanent identifiers and meanings that are assigned to items of geographic information. ISO 19135-1:2015 makes a distinction between the terms of registry and register. Accordingly, a *registry* is an information system on which a register is maintained, while a *register* is set of files containing identifiers assigned to items of geographic information with descriptions of the

5 https://www.iso.org/ics/01.040/x/

associated items. ISO 19135-1:2015 provides a conceptual register model, as well as defines roles and responsibility of relevant organizations (e.g. register owner, register manager, control body and submitting organization), and procedures for the management of registers (e.g. submission, approval and appeal procedures). ISO 19135-1:2015 has been applied for the development of the INSPIRE Registry, the DGIWG registries, and ISO Geodetic Registry.

Code lists as well as any other controlled vocabularies or KOSs can be represented by a number of human-readable (e.g. HTML, CSV) and machine-readable formats (e.g. XML, RDF, OWL). For the latter, World Wide Web Consortium (W3C) has specified Resource Description Framework (RDF) which is a graph-based data model for expressing information about things (e.g. documents, people, physical objects, and abstract concepts) in Semantic Web environment (Manola et al, 2014). Based on RDF technology, W3C has also developed a common data model, Simple Knowledge Organization Systems (SKOS) for expressing the structure and content of concept schemes such as thesauri, classification schemes, subject heading lists, taxonomies, and other similar types of controlled vocabulary. SKOS is an application of the RDF, thus it enables machine-readable representation of a KOS and allows sharing and linking different KOSs through Linked Data approach (Isaac and Summers, 2009; Miles and Bechhofer, 2009; Baker et al, 2013). INSPIRE (2017a) proposes use of SKOS for modelling of INSPIRE registers and register items. Similarly, a draft guidelines 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. By following these proposals, the current version of INSPIRE Registry includes SKOS representations of code lists used in INSPIRE application schemas.

There are also open source software tools for publication of KOSs, i.e. the Re3gistry, an open source tool for the management of reference codes developed by the European Commission's Joint Research Centre (JRC)[6] through the ARE3NA[7] action of the ISA programme.

3.3 Terminology (KOS) registries within the domain of geographic information

Over the last decade a large amount of controlled vocabularies in terms of or Knowledge Organization Systems (KOS) (e.g. code lists, taxonomies, thesauri, ontologies, have been published online. Also 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). In the following, some of these registries related to geographic information domain are briefly mentioned.

The Defence Geospatial Information Working Group (DGIWG) is the multi-national body responsible for geospatial standardization for the defence organizations of member nations. It defines information components for use in the development of product specifications and application schemas for military geospatial data. DGIWG maintains a number of registry including DGIWG Terminology Registry, DGIWG Feature and Attribute Data Registry,

[6] https://ies-svn.jrc.ec.europa.eu/projects/registry-development/wiki
[7] https://joinup.ec.europa.eu/collection/are3na

DGIWG Geodetic Codes and Parameters Registry. Among these, DGIWG Terminology Registry includes terms, their definitions and acronyms used in DGIWG community[8]. These registries are compliant with ISO 19135-1:2015 which is detailed in the next section.

INSPIRE 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 ISO 19135-1:2015, a number of INSPIRE registers has 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. These centrally managed registers are accessible in different formats (e.g. XML, RDF/XML, JSON, Atom, CSV) through the INSPIRE Registry service which uses the Re3gistry software[9]. Moreover, INSPIRE Maintenance and Implementation Group (2017a) has provided general guidance and best practices for setting up registers supporting INSPIRE implementation and for sharing the content of national or community registers. Accordingly some countries have developed national INSPIRE registries which extend INSPIRE vocabulary according to national requirements, e.g. Italian INSPIRE Registry, Austrian INSPIRE registry and BRGM Registry in France, GDI-DE Registry in Germany.

The European Petroleum Survey Group (EPSG) Geodetic Parameter Registry[10] 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' and has been implemented in GML through 'ISO 19136 Geographic information – Geographic markup language (GML)'. The EPSG Registry is maintained by the Geodesy Subcommittee of International Association of Oil and Gas Producers' (IOGP) Geomatics Committee (OGP, 2016).

Another registry example within the geographic information domain is the ISO Geodetic Registry developed by the ISO Technical Committee 211 Geographic information/Geomatics. The ISO Geodetic Registry is a database which provides parameters defining global and regional CRSs and transformations between these CRSs. It conforms to other relevant ISO standards, e.g. ISO 19111:2007, ISO/TS 19127:2005, ISO 19135-1:2015 and ISO 19135-2:2012. Its demo version is available online at https://iso.registry.bespire.eu.

[8] http://www.dgiwg.org/Terminology/
[9] http://inspire.ec.europa.eu/registry
[10] http://www.epsg-registry.org/

4. CODE LIST MANAGEMENT FOR THE LAND ADMINISTRATION DOMAIN

4.1 Categories of code list management

The ISO LADM standard has provided the basis for country experts to extend the current informative annexes F „Legal Profiles" and J „Code lists" into country profiles, especially by adding more content and „structure". Code list management thus implied the registration and publication of code lists (van Oosterom and Lemmen, 2015). This conceptual modelling was recently extended by implementing the conceptual model level LADM classes into computer-processable model descriptions (Kalogianni et al, 2017; Kim and Heo, 2017). In this context, cross-country code list issues were addressed; they are summarized in section 4.2.

The publication of OGC Land and Infrastructure Conceptual Model Standard (LandInfra) in 2016 and OGC InfraGML Encoding Standards, including Part 7 on Land Division in 2017 calls for a more complex code list management. While ISO LADM and OGC LandInfra standards are related, as described in Annex D.2 of the LandInfra standard, they have modelled the domain differently, with implications for the code lists. The corresponding management challenges are addressed in sections 4.3-6 below.

4.2 Cross-country reflections on code lists

The conceptual schema language INTERLIS has been successfully applied in the Swiss Cadastre System for several decades, and became a Swiss standard in 1998. Since 2007, it has been part of the Swiss Federal Act on Geoinformation, and all data models of the Swiss NSDI have to be described with the standard by law (Kalogianni et al., 2017).

INTERLIS was selected as the modelling language to obtain a prototype implementation of a proposed Multi-purpose Land Administration System (MLAS) for Greece (Kalogianni, 2015). The implementation of the proposed LADM-based model with INTERLIS included drawing particular attention to the explicit formulation of constraints, code lists and enumeration values. Moreover, a recently developed a Colombian LADM profile was implemented using INTERLIS and thus, an INTERLIS-based COL-LADM data model was developed, which will be applied in World Bank-financed pilot projects related to a new Multipurpose Cadastre (Jenni et al, 2017).

The findings reported compare to the code list cases described in section 2: In case of fixed values of the Greek country profile, the values of the model would be defined as enumeration types, cf. Case 1. For values that can be extended, a catalogue table with referential integrity is used to express code lists, potentially with nesting of lists of values, cf. Case 2, while code lists from external catalogues can be referenced from the model and imported into the database, cf. Case 3.

Other findings include the practice of providing a unique identifier for each code list and description attributes. The advantage of this type of code list is that its value can be updated, and it can also be versioned when adding the attributes "beginDateTime" and "endDateTime" (Kalogianni et al, 2017).

4.3 Cross-standard management of code lists

The code lists of ISO LADM and OGC LandInfra standards are summarized in Appendix 2. With the single exception of LA_DimensionType vs. DimensionType, it appears that while the code lists regard largely the same domain, the chosen names of the lists provide no certain meaning of mutual relationships. Obviously, this calls for application of one or more of the semantic tools, mentioned in section 3.3. In fact, a thesaurus is available for the Land Administration domain, as already noted in (Kalogianni et al, 2017), namely the Cadastre and Land Administration Thesaurus (CaLAThe)[11] (Çağdaş and Stubkjær, 2015). CaLAThe is inspired by and derived from ISO 19152 LADM, and has the potential of providing a framework for relating the two sets of code lists in a consistent way because of the SKOS technology applied. The present version 2 of CaLAThe needs an update, especially with the terms and definitions of OGC LandInfra. Moreover, it needs being integrated with the two sets of code lists, cf. next section, and the multi-language potential has to be developed.

OGC's overall Knowledge Management strategy includes the registering and maintenance of all code lists developed in the OGC. The data are all being managed in a triple store and will be accessible through any number of semantic web, textual, or serial encoded methods and formats. Based on the underlying triple store framework, references and links to externally managed code lists may be provided as well. The registry software/services will be available to ISO/TC 211 as well. This provides that the code lists of LADM and of LandInfra / InfraGML may become available through the same web portal.

Users world-wide may thus query the web portal, but should in addition have the option of consulting a body, staffed with standardization and domain expertise. Such body is needed anyway to frame the establishment and maintenance of the code list portal. Moreover, the consultation dialogue may provide information on needs for revision(s) of the standards. Revisions may imply code list related changes and the portal must be updated with these changes, e.g. implications of the present revision of LADM 19152:2012.

4.4 Potential users

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 the pertinent agency, e.g. in the context of renovation of existing information systems or establishment of new. Code list management thus has to provide the agency and supporting companies with an overview of available code list options and - where available – information on similar implementations and trade-offs.

Land administration agencies appear as the primary end users of the proposed code list web portal, supplemented with companies and NGOs who provide software and services for updating procedures. However, countries with federal government structure (e.g. Australia, Brazil, Canada, Germany, India, Switzerland, USA, etc.) may have a special interest, because generally their states have a mandate to localize code list at their discretion, while economy of scale suggest a shared and interoperable solution. The above-mentioned code list management body and the various federal units in charge of the land administration domain thus have a

[11] http://cadastralvocabulary.org

common concern for standards-related harmonization. Consultations on code list localizations may thus develop into specification and eventually implementation of further standards-related services.

4.5 The organizational setup

The body framing the code list portal has to include the standardization bodies, ISO and OGC. The FIG motivated the adoption of the LADM standard by ISO/TC 211 and moreover framed research on code lists. It seems therefore appropriate to join with this organisation of surveying professionals, also to benefit from sharing of expertise and cost of the management activities.

The International Office for Cadastre and Land Records (OICRF)[12] is a permanent body of FIG, a study and documentation center for cadaster, land administration and affiliated fields of interest. OICRF is, among others, charged with the tasks of providing information and advice on all cadasters and land registration systems to all interested persons and institutions for the purposes of study or to help countries wishing to set up a cadaster or land registration system or improve an existing system. The OICRF is hosted by the Netherlands Cadastre, Land Registry and Mapping Agency (in short: Kadaster).

The mandate of OICRF seems to include the proposed consultations on code list issues which could motivate formal participation into the framing code list body. However, the main activity so far appears to be a valued library function, providing access to published papers within the domain.

The OGC Land Administration DWG (Domain Working Group)[13] was established in 2016. The charter members of this 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. The announced main action is to draft a White paper on OGC Land Administration architecture and relationship with other standards and Standards Developing Organizations. The latest revision of the corresponding charter, June 2017, mentions 20 charter members, including Esri, Intergraph, Leica Geosystems, and Trimble, as well as AdV Germany, Land Information New Zealand (LINZ), and Ordnance Survey, UK. The related Landadmin.dwg mailing list is public and intended for Land Administration discussion. As of February 2018, 48 landadmin.dwg subscribers are found[14]. Last entry was a call in January 2017 for development of a posted outline of the mentioned White paper.

The present charter of the Landadmin DWG is wide in scope, perhaps wider than the tasks charged on the OICRF. While the charter does not refer explicitly to code list management, it may be subsumed, e.g. under issue 4: to help the technology and user community in land administration understand and align on the use of open geospatial standards for land

12 https://www.oicrf.org/about-oicrf
13 http://www.opengeospatial.org/projects/groups/landadmin
14 https://lists.opengeospatial.org/mailman/listinfo/landadmin.dwg

administration. The modest activity on the mail list, etc. does not indicate a strong support for code list management. On the other hand, it seems problematic to establish uncoordinated mailing lists and groups, which all address standardization aspects of land administration.

4.6 The value chain of the proposed effort

The parties who engage in establishing information technology standards are concerned that the standards are implemented and in fact provides for better interoperability. Therefore, a follow-up or outreach-like activity may be acceptable, in addition to the very specification and approval of the standards, but such follow-up has to be limited in time and effort, in order to be feasible. In the present case, it might be adequate to specify a time limit to about two years, and a meeting activity for signing parties amounting to about one meeting a year and occasional correspondence on complex and/or fundamental code list issues, raised by land administration agencies in the process of implementation of standards.

The federal units in charge of the land administration domain, mentioned above in section 5.2, seem to have a more permanent mandate, and corresponding funding, to engage in code list management. Together with national land administration agencies, who engage in the international development of the domain, e.g. Kadaster of the Netherlands, they may request a revision of the proposed organizational setup. Most likely, such initiative by one or more federal units depends on dedicated information and motivation.

5. CONCLUSION

The present paper proposes a joint management of the code lists which are specified by ISO LADM and by OGC LandInfra / InfraGML, respectively. The paper the outlines the tasks of code list management by introducing the code list facility as part of modelling activities and drawing on the OGC Code List Manifesto. Research supporting code list management, e.g. terminological theory and semantic tools is resumed, experiences within the domain of Land Administration are reported, and the setup of a possible code list management system is outlined, comprising a triple store framework, provided by the OGC, yet widely accessible, including the SKOS based Cadastre and Land Administration Thesaurus. A co-operation among the ISO/TC211, the OGC, the FIG, and perhaps the OICRF is proposed as the organizational framework for code list management, thereby furthering harmonization of standards and reduction of interoperability problems within the domain of Land Administration.

REFERENCES

Baker, T., Bechhofer, S., Isaac, A., Miles, A., Schreiber, G, 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, pp. 35-49.

Çağdaş, V., Stubkjær, E. (2015). A SKOS vocabulary for Linked Land Administration: Cadastre and Land Administration Thesaurus. Land Use Policy, 49, pp. 668–679.

Chatterjee, A. (2016). Elements of information organization and dissemination. Chandos Publishing.

GML 3.2.1 OGC 07-036 OpenGIS® Geography Markup Language (GML) Encoding Standard.

GML 3.3 OGC 10-129r1 OGC® Geography Markup Language (GML) Extended schemas and encoding rules.

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

Golub, K., Tudhope, D., Zeng, M.L., Žumer, M. (2014). Terminology registries for knowledge organization systems: Functionality, use, and attributes. Journal of the Association for Information Science and Technology. 65 (9), pp. 1901–1916.

Greiner, U. (2006). Business Documents - Concepts and Techniques.
The Athena Consortium, 2006. Slide 35/48 at http://slideplayer.com/slide/6925575/.

Hodge, G. (2000). Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. The Digital Library Federation, Washington, DC.

Isaac, A., Summers, E. (2009). SKOS Simple Knowledge Organization System. Primer, World Wide Web Consortium (W3C).

INSPIRE, 2017a. Best Practices for registers and registries & Technical Guidelines for the INSPIRE register federation. INSPIRE Maintenance and Implementation Group, 2017-05-31, Version 1.0.

INSPIRE (2017b). Guidelines for the RDF encoding of spatial data, 2017-07-17 at http://inspire-eu-rdf.github.io/inspire-rdf-guidelines/.

ISO 19103:2015 Geographic information -- Conceptual schema language.

ISO 25964-1:2011. Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval.

ISO 25964-2:2013. Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies.

ISO/TC211, 2017, Resolutions from the 45th ISO/TC211 Geographic information/Geomatics plenary meeting in Wellington, New Zealand, 2017-11-30-2017-12-01.

Jackson, B.S. (1995). Making Sense in Law. Legal Semiotics Monographs. Deborah Charles Publications, Liverpool.

Jenni, L., Germann, M., Eisenhut, C., Guarin, L.A., Bajo, V.M. (2017). LADM Implementation in Colombia – Process, Methodology and Tools developed and applied. In Proceedings of the FIG Working Week, Helsinki, Finland, 29 May-2 June 2017.

Kalogianni, E. (2015). Design of a 3D Multipurpose Land Administrative System for Greece in the Context of Land Administration Domain Model (LADM). Master's Thesis, National Technical University of Athens, Athens, Greece, 2015.

Kalogianni, E., Dimopoulou, E., Quak, W., Germann, M., 3, Lorenz Jenni, L., Van Oosterom, P. (2017). INTERLIS Language for Modelling Legal 3D Spaces and Physical 3D Objects by Including Formalized Implementable Constraints and Meaningful Code Lists. ISPRS International Journal of Geo-information, 6, 319, p. 36.

Kim, S., Heo, J. (2017). Development of 3D underground cadastral data model in Korea: Based on Land Administration Domain Model. Land Use Policy, 60 () pp. 123–138.

Ledl, A., Voß, J. (2016). Describing Knowledge Organization Systems in BARTOC and JSKOS. Available online at http://eprints.rclis.org/29366/1/Ledl_Voss_TKE2016_final_version_20160518.pdf.

Manola F., Miller E., McBride B. (2014). RDF 1.1 Primer. W3C Working Group Note, 24 June 2014, https://www.w3.org/TR/rdf11-primer/.

Miles, A., Bechhofer, S. (2009). SKOS Simple Knowledge Organization System Reference. W3C recommendation, W3C.

Ogden, C., Richards, I. (1923). The Meaning of Meaning. Brace & World, Harcourt, New York.

OGP (2016). Geomatics Guidance Note 7, part 3 EPSG Geodetic Parameter Registry - Developer Guide.

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, Abuja, Nigeria, 6-10 May 2013. International Federation of Surveyors (FIG), Copenhagen, Denmark.

Scarponcini, P. (2017). Code List Manifesto. OGC Pending document, 17-050r1, 2017-09-04.
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. Land Admin DWG, 2017-03-21.

Suonuuti, H. (2001). Guide to Terminology. NORDTERM 8, second edition, originally printed 1997. Published by The Finnish Centre for Technical Terminology, 2001, Helsinki.

Van Oosterom, P., Lemmen, C. (2015). The Land Administration Domain Model (LADM): Motivation, standardisation, application and further development. Land Use Policy, 49 pp. 527–534.

Zeng, M.L. (2004). Trends and Issues in Establishing Interoperability among Knowledge Organization Systems. Journal of the American Society for Information Science and Technology, 55 (5) pp. 377–395.

APPENDICES

1. *Draft* Memorandum of Understanding

[This draft is prepared as an academic exercise and does in no way imply the consent of the parties mentioned].

Memorandum of Understanding
between the signing parties:
International Organization for Standardization (ISO/TC 211),
Open Geospatial Consortium (OGC),
International Federation of Surveyors (FIG),
+???
concerning joint code list management

Purpose

The signing parties establish a joint unit for code list management within the domain of land administration and cadastre, in order to harmonize the implementation and revisions of standards within the domain and to develop best practices for code list adoption to localization needs.

Background

The number of standards within information technology increases with the implication that an application domain more likely is addressed by more standards. This is in fact also the case for the domain of land administration and cadastre, which saw the issuing of ISO 19152:2012 Geographic information -- Land Administration Domain Model (LADM) and - four years later - OGC Land and Infrastructure Conceptual Model Standard (LandInfra) and in 2017 the corresponding encoding standards, OGC InfraGML, Parts 0-7, where especially Part 7 on Land Division regards the said domain. While addressing largely the same domain, the two standards have different initiating contexts and various scopes. However, implementation practices and revision changes ought to opt for a more integrated solution and the signing parties intends to support such effort.

Moreover, implementation of standards has to adapt to localization needs and this applies especially to the code lists specified by the standards. The code lists of LADM and the domain relevant code lists of LandInfra/ InfraGML are listed in Appendix 2. The signing parties find that the envisioned more integrated solution is promoted by supporting implementing agencies with information and suggestions/ advice from a joint unit, established by the signing parties.

<u>Definitions</u>
(if needed)

<u>Parties</u>
(Name, address, contact, etc. for each of the parties)

<u>Structure</u>
The signing parties establish a joint code list management unit, which offers information and advice on localization of code lists within the domain of cadastre and land administration.
The unit consists of a management group, a secretariat, and an optional group of user representatives.

The signing parties appoint each x person(s) to the management group. The members of the group may elect a chairperson and another person in charge of the secretariat.

The secretariat provides for a web portal, which informs on standards, especially the code lists of the standards, and on advices provided. Localized code list implementation may be published as well, if supplied by the authority concerned.

Federal government units in charge of state agencies for cadastre and land administration may nominate persons to a group of user representatives. Members of this group may request issues included into the agenda of management group meetings.

<u>Activities</u>
This MoU applies to two years of activity.

The management group meets twice a year, if so preferred through online meetings. A summary is published.
During 3rd and 4th meeting, code list management activities are evaluated and continuation efforts assessed.

The secretariat answers requests from standard implementing agencies and companies. If needed, a motivated response draft is submitted to members of the management group for deliberation and communication to the requesting party.

Preconditions of cooperation
The MoU implies no settlement between the signing parties. Each of the signing parties covers its own costs related to the operation of the management group and the secretariat.

Signatures
…

---

2. Code Lists within the domain of Land Administration

LADM Code Lists	LandInfra Code Lists
6.3.4 Party Package LA_PartyRoleType LA_GroupPartyType LA_PartyType **6.4.9 Administrative Package** LA_AdministrativeSourceType LA_MortgageType LA_RightType LA_RestrictionType LA_ResponsibilityType LA_AvailabilityStatusType LA_BAUnitType **6.5.8 Spatial Unit Package** LA_BuildingUnitType LA_AreaType LA_VolumeType LA_SurfaceRelationType LA_DimensionType LA_UtilityNetworkStatusType LA_RegisterType LA_UtilityNetworkType LA_LevelContentType LA_StructureType LPIS..SubParcelType **6.6.6 Surveying and Representation Subpackage** LA_MonumentationType LA_SpatialSourceType LA_InterpolationType LA_PointType	**7.2 Core** DocumentType ProfessionalType **7.8 Survey** SurveyType SurveyResultType **7.9 LandFeature** LandElementType **7.10 LandDivision** DimensionType EasementType ImplicitSurface LandParcelCurrentLandUse LandParcelPlannedLandUse LandParcelState SigningRole StatementType StringDirection StringType SuperficieObjectType SurveyMonumentType **7.11 Condominium** BuildingPartType CondominiumUseType

3. ISO standards related to different aspects of terminology

ISO 1087-1:2000 Terminology work – Vocabulary – Part 1: Theory and application.
ISO 860:2007 Terminology work – Harmonization of concepts and terms.
ISO 1951:2007 Presentation/representation of entries in dictionaries – Requirements, recommendations and information.
ISO 704:2009 Terminology work – Principles and methods.

ISO 23185:2009 Assessment and benchmarking of terminological resources – General concepts, principles and requirements.
ISO 29383:2010 Terminology policies – Development and implementation.
ISO 10241-1:2011 Terminological entries in standards – Part 1: General requirements and examples of presentation.
ISO 26162:2012 Systems to manage terminology, knowledge and content – Design, implementation and maintenance of terminology management systems.
ISO 22274:2013 Systems to manage terminology, knowledge and content – Concept-related aspects for developing and internationalizing classification systems.
ISO 24156-1:2014 Graphic notations for concept modelling in terminology work and its relationship with UML – Part 1: Guidelines for using UML notation in terminology work.
ISO 19104:2016 Geographic information – Terminology.

BIOGRAPHICAL NOTES

Erik Stubkjær is emeritus professor, having served as professor of cadastre and land law at Department of Development and Planning, Aalborg University 1977 – 2008. Recently, he engaged in standardization activities, contributing to OGC standards LandInfra and InfraGML (2016/17). He initiated in 2001 the research project 'Modelling Real Property Transactions,' supported by European Science Foundation as COST action G9. Taking an institutional economics perspective, the project framed the modelling of transactions in real estate, e.g. the purchase of real property, mortgaging, and cadastral processes, as basis for assessing transaction costs. He graduated as land surveyor in 1964. During 1979-1988, he was member of the Tribunal of the Danish Association of Chartered Surveyors.

Jesper M. Paasch is a senior lecturer/associate professor at the University of Gävle and coordinator of research in geographic information at Lantmäteriet, the Swedish mapping, cadastral and land registration authority, Gävle, Sweden. He holds a MSc degree in Surveying, planning and land management, a Master of Technology Management degree in Geoinformatics, both from Aalborg University, Denmark, and a PhD degree in Real Estate Planning from the KTH Royal Institute of Technology, Stockholm, Sweden. His thesis concerned the development of the Legal Cadastral Domain Model. He is a Swedish delegate in FIG, Commission 3, and was a delegate in the drafting team of ISO 19152:2012 LADM. He is a member of the FIG joint commission 3 and 7 working group on „3D-Cadastres‟.

Volkan Çağdaş has been working in Yildiz Technical University (YTU), Department of Geomatic Engineering, Istanbul / Turkey. He obtained his Ph.D. degree in 2007, and then studied as a post-doc researcher at Aalborg University under the supervision of Prof. Dr. Erik Stubkjær. In 2010, he became an assistant professor in YTU, and in 2014 he was awarded an associate professorship in cadastre and land administration. He has been teaching cadastre, immovable property law, land re-adjustment, immovable property valuation, and land information management systems at undergraduate and graduate levels. His research interest covers both the technical and the institutional aspects of cadastre and land administration.

Peter van Oosterom obtained an MSc in Technical Computer Science in 1985 from Delft University of Technology, the Netherlands. In 1990 he received a PhD from Leiden University. From 1985 until 1995 he worked at the TNO-FEL laboratory in the Hague. From 1995 until 2000 he was senior information manager at the Dutch Cadastre, where he was involved in the renewal of the Cadastral (Geographic) database. Since 2000, he is professor at the Delft University of Technology, and head of the 'GIS Technology' Section, Department OTB, Faculty of Architecture and the Built Environment, Delft University of Technology, the Netherlands. He is the current chair of the FIG Working Group on '3D Cadastres'. He was one of the co-editors of ISO 19152:2012 LADM (together with Christiaan Lemmen and Harry Uitermark).

Scott Simmons is Executive Director of the Standards Program of the Open Geospatial Consortium. In this role, he coordinates member-driven standards development activities and ensures that all OGC standards progress through the organization's consensus process to approval and publication. His OGC-related research has focused on data lifecycle management, integration, and dissemination. He has also served as an Adjunct Professor at the College of Wooster, on the Advisory Board of the GIS Cluster at the Rocky Mountain Ionosphere, and as a subject matter expert for seminars at universities and conferences around the World. He holds a Bachelor of Science degree in Geology from the University of Texas and a Master of Science degree in Geology from Southern Methodist University.

Jenny Paulsson is an Associate Professor at the Department of Real Estate and Construction Management of the KTH Royal Institute of Technology, Stockholm, Sweden. She holds a MSc degree in Surveying and a PhD degree in Real Estate Planning, both from the KTH Royal Institute of Technology. Her PhD thesis concerned 3D property rights. She is a member of the FIG joint commission 3 and 7 working group on „3D-Cadastres".

Christiaan Lemmmen is full Professor Land Information Modeling at the Faculty of Geo-Information Science and Earth Observation of the University of Twente in the Netherlands. His other main job is as Senior Geodetic Advisor at Kadaster International, the international branch of the Netherlands Cadastre, Land Registry and Mapping Agency. He is director of the OICRF, the International Office of Cadastre and Land Records, one of the permanent institutions of the International Federation of Surveyors (FIG). He is chairing the Working Group Fit-For-Purpose Land Administration of the Commission 7, Cadastre and Land Management of the International Federation of Surveyors (FIG). He is contributing editor of GIM International, the worldwide magazine on Geomatics. He is co-editor of the International Standard for the Land Administration Domain, ISO 19152 and the designer of the Social Tenure Domain Model (in co-operation with UN HABITAT and FIG). He holds a PhD from Delft University of Technology, the Netherlands. Title of his thesis is 'A Domain Model for Land Administration'.

CONTACTS

Erik Stubkjær
Aalborg University, Department of Planning
Rendsburggade 14, niveau 3
DK-9000 Aalborg
DENMARK
Phone: +45 23319553
E-mail: est@land.aau.dk
Website: http://people.plan.aau.dk/~est/

Jesper M. Paasch
University of Gävle / Lantmäteriet
SE-801 76 Gävle / SE-801 82 Gävle
SWEDEN
Phone:+4620154701/+46 26633001
E-mail: jesper.paasch@hig.se / jesper.paasch@lm.se
Website: www.hig.se/jesperpaasch / www.lantmateriet.se

Volkan Çağdaş
Yıldız Technical University, Department of Geomatic Engineering
34210 Esenler,
Istanbul
TURKEY
Phone: +90 212 383 5313
Fax: +90 212 383 5274
E-mail: volkan@yildiz.edu.tr
Website: http://yildiz.edu.tr/~volkan/

Peter van Oosterom
Delft University of Technology
Julianalaan 134
2628 BL Delft
THE NETHERLANDS
Phone: +31 15 2786950
E-mail: P.J.M.vanOosterom@tudelft.nl
Website: www.gdmc.nl

Scott Simmons
Executive Director, OGC Standards Program
Colorado
USA
E-mail: ssimmons@opengeospatial.org
Website: http://www.opengeospatial.org/ogc/organization/staff/ssimmons

Jenny Paulsson
KTH Royal Institute of Technology, Real Estate Planning and Land Law
Teknikringen 10B, 10044 Stockholm
SWEDEN
Phone: +46 87906661
E-mail: jenny.paulsson@abe.kth.se
Website: http://www.kth.se/profile/paulsson

Christiaan Lemmen
University of Twente, Faculty of Geo-Information Science and Earth Observation/ITC
P.O. Box 217
7500AE Enschede
THE NETHERLANDS
Phone: + 31 6 52481717
E-mail: c.h.j.lemmen@utwente.nl
Website: www.itc.nl
and
Cadastre, Land Registry and Mapping Agency, Kadaster International
P.O. Box 9046
7300 GH Apeldoorn
THE NETHERLANDS
Phone: +31 88 183 4417
Email: Chrit.Lemmen@kadaster.nl
Website: www.kadaster.nl