RDA Urban Quality of Life Indicators (UQoL): Proposal for work on the UQoL Data Dictionary

# Background

The JRC is the technical coordinator of the INSPIRE Directive, which is the legal framework in Europe (adopted and being implemented across all 28 EU countries) establishing an infrastructure for spatial information in Europe. The purpose of INSPIRE is to make the data necessary to support environmental policies, or policies that affect the environment, discoverable, accessible, shareable, and interoperable. INSPIRE covers 34 data themes, many of which will be relevant also to UQOL. They are organised in three groups or annexes:

**Key data themes addressed by INSPIRE**

|  |  |
| --- | --- |
| **Annex I**  Coordinate reference systems  Geographical grid systems  Geographical names  Administrative units  Addresses  Cadastral parcels  Transport networks  Hydrography  Protected sites  **Annex II**  Elevation  Land cover  Ortho-imagery  Geology | **Annex III**  Statistical units  Buildings  Soil  Land use  Human health and safety  Utility and governmental services  Environmental monitoring facilities  Production and industrial facilities  Agricultural and aquaculture facilities  Population distribution – demography  Area management/restriction /regulation zones & reporting units  Natural risk zones  Atmospheric conditions  Meteorological geographical features  Oceanographic geographical features  Sea regions  Bio-geographical regions  Habitats and biotopes  Species distribution  Energy Resources  Mineral resources |

To develop interoperability across these 34 data themes (in 28 countries, and 24 languages) is not a trivial process. The approach taken has been to define for each theme (with the help of many experts across Europe) what are the key variables that are necessary to support environmental policy, develop a generalised data model, and then develop methods and tools so that the many organisations responsible for data falling in each of these themes can map the schemas of their data to the commonly agreed European one. For more information see <http://inspire.ec.europa.eu/index.cfm> and more specifically <http://inspire.ec.europa.eu/index.cfm/pageid/2>

As part of this process, a key issues has been to agree on terms and definitions and publish them in the INSPIRE registry. <http://inspire.ec.europa.eu/registry/>

Many of these terms can contribute to the creation of an Urban Dictionary as set out in the WG case statement:

*The WG will assess the feasibility of developing a dedicated urban data dictionary, leveraging existing thesauri and dictionaries such as the INSPIRE feature concept dictionary, the General Environmental Multilingual Thesaurus, EuroVOC, etc. and augmented by folksonomies derived from social networks data mining.*

As part of the JRC research activities, we have already developed bridges between the INSPIRE Feature Concept Dictionary and other thesauri and vocabularies. With this in mind, we propose to start from this core set of aligned vocabularies and explore its feasibility as a framework for an UQOL Urban data dictionary.

# Proposal

To assess the feasibility of developing a dedicated urban data dictionary, we propose to set up an online environment for the creation and maintenance of a knowledge base (KB), where WG members can:

* contribute existing thesauri
* identify relevant terms from existing thesauri
* contribute / create alignments across relevant terms from different thesauri
* create new thesauri.

From the technological side, the proposal is to use Linked Data and Semantic Web tools and languages. In particular we propose to use SKOS[[1]](#footnote-1) for the representation of the KB’s content, and a triple store for its storage, maintenance, and access. The rationale behind this proposal is that (a) SKOS is currently a widely used standard for the representation of thesauri and semantic relationships, and (b) triple stores are specifically designed for storing Linked Data (as SKOS-encoded content is) and it supports a query language, SPARQL[[2]](#footnote-2), offering WG members the ability to effectively explore the KB’s semantic graph.

# Our contribution

JRC can offer a contribution at two levels:

## Setting up and maintaining the triple store to store the knowledge base

We plan to use OpenLink Virtuoso[[3]](#footnote-3), which offers the basic functionalities required for the knowledge base environment (KBEnv).

WG members will be given specific access privileges to edit the KB. These operations will be carried out manually by using the Virtuoso frontend. Potentially, the knowledge base can also be edited via the SPARUL API (SPARQL Update[[4]](#footnote-4)). The availability of this option will be considered in a later phase, based on use cases and requirements.

## Contributing in the creation of the knowledge base

JRC can contribute a number of thesauri and alignments as a basis for the KB.

In particular, we can contribute all the alignments created in the framework of the EU project EuroGEOSS[[5]](#footnote-5), which cover the following thesauri:

* INSPIRE themes[[6]](#footnote-6) and feature concepts[[7]](#footnote-7)
* GEMET[[8]](#footnote-8): The General Multilingual Environmental Thesaurus
* GEOSS Societal Benefit Areas[[9]](#footnote-9)
* GEOSS Earth Observation vocabulary
* EuroGEOSS Drought vocabulary[[10]](#footnote-10)
* GCMD Earth science keywords[[11]](#footnote-11)
* Cadastre and Land Administration Thesaurus (CaLAThe)[[12]](#footnote-12)

These alignments, currently used in GEOSS, are available through the JRC’s SemanticLab SPARQL endpoint[[13]](#footnote-13) (HTML preview: <http://bit.ly/1z1hLSu>).

Moreover, we can include in the KB the alignments developed by the EU Publications Office between EuroVoc[[14]](#footnote-14) and a number of other thesauri - including, AGROVOC[[15]](#footnote-15), GEMET and the INSPIRE themes. These alignments are available from the EU Open Data Portal.[[16]](#footnote-16)

1. <http://www.w3.org/TR/skos-reference/> [↑](#footnote-ref-1)
2. <http://www.w3.org/TR/rdf-sparql-query/> [↑](#footnote-ref-2)
3. <http://virtuoso.openlinksw.com/> [↑](#footnote-ref-3)
4. <http://www.w3.org/Submission/SPARQL-Update/> [↑](#footnote-ref-4)
5. <http://www.eurogeoss.eu/> [↑](#footnote-ref-5)
6. <http://inspire.ec.europa.eu/theme/> [↑](#footnote-ref-6)
7. <http://inspire.ec.europa.eu/featureconcept/> [↑](#footnote-ref-7)
8. <http://www.eionet.europa.eu/gemet/> [↑](#footnote-ref-8)
9. <http://en.wikipedia.org/wiki/Societal_Benefit_Areas> [↑](#footnote-ref-9)
10. <http://ijsdir.jrc.ec.europa.eu/index.php/ijsdir/article/view/264> [↑](#footnote-ref-10)
11. <http://gcmd.nasa.gov/learn/keyword_list.html> [↑](#footnote-ref-11)
12. <http://www.cadastralvocabulary.org/> [↑](#footnote-ref-12)
13. <http://semanticlab.jrc.ec.europa.eu:4433/sparql> [↑](#footnote-ref-13)
14. <http://eurovoc.europa.eu/> [↑](#footnote-ref-14)
15. <http://aims.fao.org/agrovoc> [↑](#footnote-ref-15)
16. <https://open-data.europa.eu/en/data/dataset/eurovoc> [↑](#footnote-ref-16)