**2022-07-11**

ICEG Legal Thesauri

**OVERVIEW**

|  |  |
| --- | --- |
| Start date | Q3 2022 |
| End date | TBD |
| Chairman | TBD |
| Project team | TBD |
| Scheduled meetings | TBD |
| Decision criterion | Unanimity minus one (U-1)  (each party has one vote) |
| License | Existing thesauri (and other materials) contributed to the WG remain property of (and keep the license determined by) the respective owners.  Creative Commons Zero will be used for new mappings and other deliverables of the working group |
| License documentation | <https://creativecommons.org/publicdomain/zero/1.0/> |
| Issue logging | [Issues · belgif/thematic (github.com)](https://github.com/belgif/thematic/issues)  (use Label "Legal Thesauri") |

# Context

## WHAT

Via this initiative, the various stakeholders wish to semantically map (parts of) the open legal thesauri being used in Belgium, and promote the reuse of these thesauri to a wider audience.

More specifically, the Universele Thesaurus Universel (UTU) and the thesaurus of the Court of Cassation - both used by the federal JuPortal.be – are to be published (and if possible mapped) using the widely used semantic specification W3C SKOS. While the mappings will be available under Creative Commons Zero, the thesauri themselves remain the property of their respective owners and can be licensed differently.

It is envisaged that some parts of these thesauri can also be mapped to other international thesauri like the EUROVOC thesaurus of the European Publication Office, though it may not be feasible to map each and every term, due to the sheer size of the thesauri. It may, however, be beneficial to at least map the “top terms”.

In addition, participants may suggest other improvements and additional translations, or usage notes on the terms of various thesauri. It is up to the maintainers of said thesauri whether or not these suggestions will be integrated in the thesauri.

## WHY

Law and justice are pillars of democracy. It is therefore of great importance that legal documents (court decisions, legal doctrine…) can be easily found, compared and used.

As an aid to achieving this goal, thematic classification or enrichment of texts with keywords from one or more thesaurus can be used to improve search engines and knowledge systems. Since high-quality thesauri already exist and are being maintained by various parties, it is useful to promote them to a wider audience, and - if possible - to map the relationships between the keywords of these different, independent thesauri.

In this way, lawyers and software companies can build new applications that make maximum use of these thesauri, and even combine them with thesauri in other countries.

## USe cases

There are various use cases for which the mapping effort provides added value. These use cases will be discussed in the first business workshop.

### First use case

*Users, which may include non-profit organizations and commercial entities, could use these thesauri to build databases and new search engines on (soon to be published) public and anonymized judgments of various Belgian courts.*

### Second use case

*Academics from other countries may perform additional analysis on Belgian judgments using EUROVOC-mappings and the published thesauri.*

# Scope

The following items are within the scope:

* Publication of UTU and Court of Cassation thesauri as SKOS file
* Mapping between these thesauri using SKOS (at least top level terms)
* Mapping between these thesauri and EUROVOC (at least top level terms)
* Promoting the work to a large(r) audience
* Suggestions for tooling, publication processes

In the workshops the different use cases and input of the stakeholders will be used to fine-tune the scope and deliverables.

# Stakeholders

The stakeholders of this process include:

|  |  |
| --- | --- |
| **Stakeholder** | ***Type*** |
| Court of Cassation | *Legal experts and maintainer of thesaurus* |
| UTU VZW | *Legal experts and maintainer of thesaurus* |
| Instituut Gerechtelijke Opleiding (IGO-IFJ) | *Users and legal experts* |
| Other courts | *Potential users of thesauri* |
| FPS Justice | *Maintainers of JuPortal and users of the thesauri* |
| FPS BOSA DG DT | *Linked (open) data experts and open data sponsor* |
| FPS Economy | *Potential users of thesauri (economic court decisions)* |
| Regions | *Users of thesauri* |
| OpenJustice ASBL | *Interested third party (NPO)* |
| CogniZone | *Interested third party (SME)* |

# SuccesS criteria

This process will be considered a success when the deliverables are widely used and applied. Initially within the Government in Belgium, but also beyond. In particular, we list the following criteria:

1. There is maximum coordination with all stakeholders - mentioned in point 3 - who are represented in at least one of the workshop sessions
2. The workshops (and work between the workshops) results in a series of reusable deliverables that represent a consensus of all participants

# Deliverables

The workshops will deliver the following deliverables:

* a business workshop with stakeholders to validate and further expand the information needs.
* 4 workshops composed of domain experts to processing the feedback
* a linked data (SKOS) version of the UTU thesauri
* a linked data (SKOS) version of the Court of Cassation thesauri
* mapping(s) between terms of these two thesauri (if it is not feasible to map all the terms due to the number of terms in these thesauri, the top level terms will be mapped)
* mapping(s) between terms of the thesauri and EUROVOC (if it is not feasible to map all the terms due to the number of terms in these thesauri, the top level terms will be mapped)
* a communication to inform a larger audience (professionals not part of the working group) about the thesauri and the work of the working group

# Milestones and timing

|  |  |
| --- | --- |
| **Date[[1]](#footnote-2)** | **Milestone** |
| **2022-10** | ICEG meeting: go/no-go |
|  | Prepare a letter of intent and invite interested parties to the first business workshop. |
|  | Business workshop with stakeholders to validate the information needs and further refine scope. |
|  | Validation of charter by Workshop Data standards |
| **2022-10** | Workshop 1 |
| **2022-12** | Workshop 2 |
| **2023-02** | Workshop 3 |
| **2023-04** | Workshop 4 |
|  | Start of public review period - Recognition "Candidate Standard" – Workshop Data Standards |
| **2023-05** | Public review period |
|  | End of public review period - Recognition of "Standard" - Workshop Data Standards |
| **2023-07** | Standard communication to the ICEG Committee |
| **2023-07** | Additional communication to a larger audience / the public |

# dependencies

* UTU vzw thesauri
* Court of Cassation thesauri

1. Specific data to be confirmed [↑](#footnote-ref-2)