**2024-04-29**

ICEG

Person

**OVERVIEW**

|  |  |
| --- | --- |
| Start date | TBD |
| End date | TBD |
| Chairman | TBD |
| Project team | * Jonas Demeulenaere * Vincent Turine * Liesbet D’Hondt * Mathieu Tulpinck |
| Scheduled meetings | TBD |
| Decision criterion | Unanimity minus one (U-1)  (each party has one vote) |
| License | TBD |
| License documentation | TBD |
| Issue logging | [Issues · belgif/thematic (github.com)](https://github.com/belgif/thematic/issues)  (use Label "Person") |

# Context

The Belgian Crossroads Bank for Social Security (CBSS) is spearheading an initiative to redesign the distribution of natural person data across multiple services. The modeling effort presented here fits in that context. The goal is to develop a formalized person data model which can be used as a common basis across these services. This formalization is necessary to address the current challenges of fragmented and non-interoperable data sources which lead to inefficiencies in data provision and consumption. By standardizing the way person data is structured, the CBSS ensures that data is not only accessible and reusable but also consistent across various government platforms, facilitating better service delivery and decision-making.

A key aspect of this initiative is the detailed understanding of the relationships between different facets of person data, which is a fundamental precondition for the design of effective REST APIs. These relationships enable the APIs to efficiently query and manipulate data according to the needs of consumers. Without a clear and structured data model, it would be challenging to achieve the level of interoperability and flexibility required for modern digital services. Thus, the redesign of person data services towards REST JSON APIs, guided by a robust and formalized data model, is essential for the future-proofing of public digital infrastructures.

The distribution of person data currently occurs at multiple governmental levels, making ICEG an ideal forum to flesh out the person data model.

## WHAT

Via this initiative, the various stakeholders propose to semantically model the different data flows and standardize the structure of the data for the following subject:

* Designation of the natural person (identification number, names, …)
* Life events (birth, death, ...)
* Relations with other legal subjects, in particular other natural persons (marriage, descent, …)
* Legal statuses of the natural person (citizenship, residency status, civil status, …)
* Association between natural person and spatial entities (residences, links to other jurisdictions, …)
* Evolution of the data points over time. This may include how, why and by whom a given data point has been recorded.
* The records of natural persons should be assignable to the maintenance of a particular public institution (in Belgium most often a municipality). This aspect also includes metadata information such as the status of the record within the registry (e.g.: struck off by official authority, …)

The focus is on terms related to natural person data. We start from terms defined for these entities in the existing person registers[[1]](#footnote-2).

## WHY

A semantic standard makes sharing and exchanging data between different stakeholders easier. Each stakeholder can directly use and interpret the data of the other. This stimulates the exchange and reuse of data and reduces the cost of exchange.

In the semantic web, data is distributed in a different way so that the AI driven machines and the digital gatekeepers of the future such as Siri, Alexa, Cortana, Google Assistant, etc. are able to use and interpret the data. The semantic standard provides machine-readable data.

Opening up semantic data initiates innovation and will enable companies to develop more intelligent products and services. By linking data, we also have richer data. Enriched data from which more knowledge can be obtained.

## USe cases

There are various use cases for which the standardization of **natural persons** provides added value. These use cases will be discussed in the first workshop.

### First use case

The first use case is to exchange data about natural persons in an unambiguous manner. A clear vocabulary allows consumers and re-users of data to have a clear view on what is meant by the incoming data about the person.

### Second use case

Personal data is by nature sensitive[[2]](#footnote-3). In this context, it may be interesting to rely on APIs which expose fragments or parts of a subject’s data, instead of returning a full hierarchical document to consumers. In that regard RESTful APIs tend to lean themselves better to this use case. Before implementing such APIs and setting appropriate routes, it is useful to have a strong understanding of the underlying domain. In the case of personal data, this means primarily a formalization effort of the relationships between facets of personal data.

### Third use case

Personal data is maintained in authentic sources / base registries. The core function of such a registry is to attest of the state of the data points under its scope[[3]](#footnote-4). The data exposed by authentic sources is deemed to be reliable, which allows consumers to build critical processes on top of it. Reliability is dependent on auditability. This means that the current state of the data should be explainable. By designing a base registry as a system of record of the evolution of data over time, it becomes possible to explain why and how the data is in its current state.

### Fourth use case

Personal data should be placed within the surrounding legal context. Public authorities are not interested in mapping all the personal relations between natural persons (friends, …) but want to keep track of the legally relevant ones (marriage, descent,…). The data of natural persons is inherently legal: it is registered through legal processes (administrative, judicial, …) and is about what rights and obligations are derived from specific legal statuses (citizenship , residency status, …). Framing natural persons as legal subjects expose opportunities for future formalization of public services as right allocation mechanisms[[4]](#footnote-5).

# Scope

The following facets are within the scope of this traject:

* Personal designation of the natural person.
* Life events.
* Relations with other legal subjects, in particular other natural persons.
* Association between natural person and spatial entities.
* Evolution of the data points over time.
* Maintenance of records and status within the registry.

# Stakeholders

The stakeholders of this process include:

* [Kruispuntbank Sociale Zekerheid](https://www.ksz-bcss.fgov.be/nl)
* FOD Binnenlandse Zaken – [Rijksregister](https://www.ibz.rrn.fgov.be/nl/rijksregister/)

# 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:

* There is maximum coordination with all stakeholders - mentioned in point 3 - who are represented in at least one of the workshop sessions
* The workshops result in a stable candidate standard that represents a consensus of all participants
* The specification is accepted by the data standards workshop and the Steering Board
* The specification is implemented and at least the framework data is published semantically.

# Deliverables

The workshops will deliver the following deliverables:

* Drawing up an overview of information needs based on analysis of available documentation and existing standards.
* Organizing and facilitating 3 workshops composed of domain experts and processing their feedback.
* Preparation of reusable documentation for the information model and publication on belgif.be:
  + RDF vocabulary
  + HTML documentation for the vocabulary with terms and definitions
  + UML diagram
  + HTML documentation for the UML diagram
  + SHACL validation rules
  + JSON-LD context file
* Integration in the ICEG system of vocabularies
* Integration in the OSLO system of vocabularies

# Milestones and timing

|  |  |
| --- | --- |
| **Date[[5]](#footnote-6)** | **Milestone** |
| **2024-05-23** | ICEG meeting: go/no-go |
| **TBD** | Prepare a letter of intent and invite interested parties to the first workshop. |
| **TBD** | Validation of charter by Workshop Data standards |
| **TBD** | Workshop 1 |
| **TBD** | Workshop 2 |
| **TBD** | Workshop 3 |
| **TBD** | Start of public review period - Recognition "Candidate Standard" – Workshop Data Standards |
| **TBD** | Public review period |
| **TBD** | End of public review period - Recognition of "Standard" - Workshop Data Standards |
| **TBD** | Standard communication to the ICEG Committee |

# dependencies

* ICEG Generic. The person data model is dependent on a robust modeling of domestic and foreign spatial references (addresses,…).

1. https://www.ibz.rrn.fgov.be/nl/rijksregister/reglementering/onderrichtingen/lijst-van-de-informatietypes/ [↑](#footnote-ref-2)
2. Personal data as discussed here squarely falls within the scope of GDPR, among other regulations. [↑](#footnote-ref-3)
3. Authentic sources can apply to various underlying domains: persons, addresses, organizations,… [↑](#footnote-ref-4)
4. Rechtentoekenning, attribution de droits,… [↑](#footnote-ref-5)
5. Specific data te be confirmed after sourcing consultants [↑](#footnote-ref-6)