Applying SHACL to DmfA declarations: a small experiment

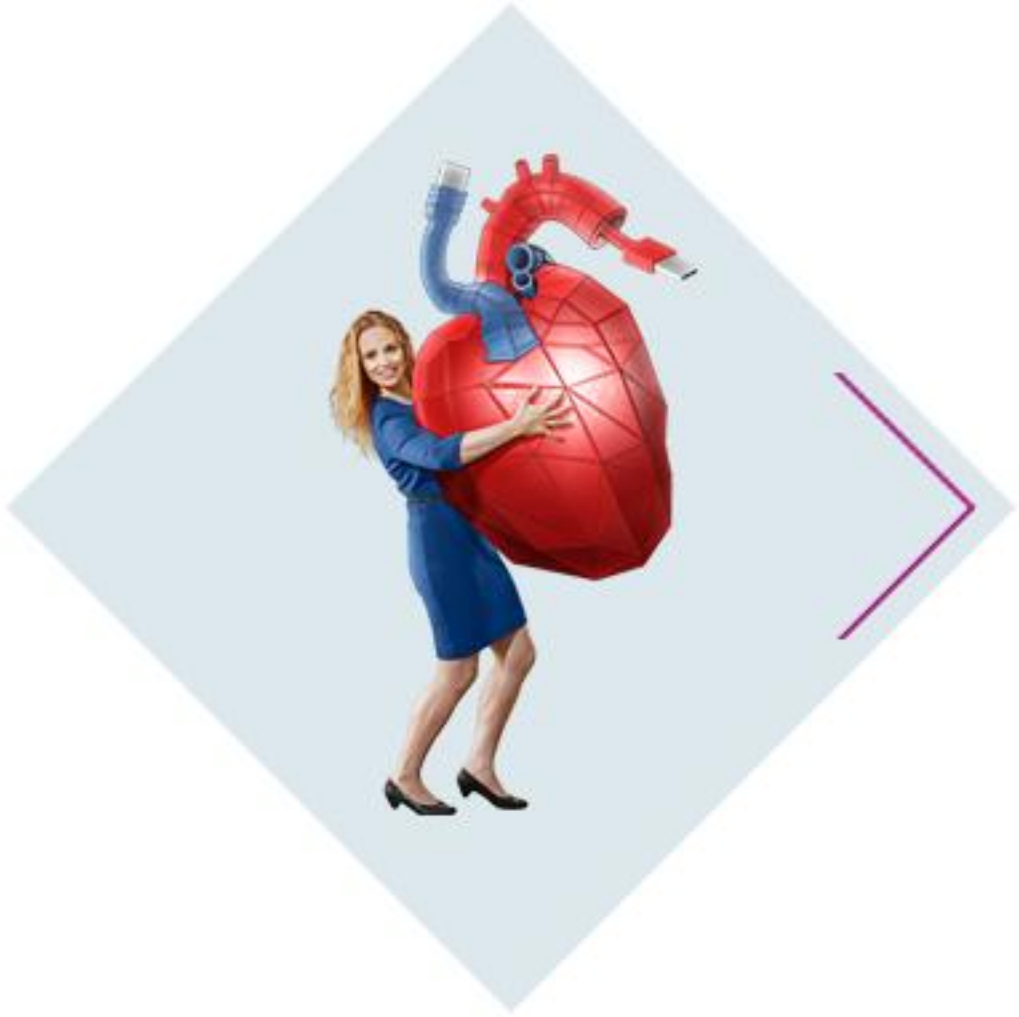
Christophe Debruyne

Smals
ICT for society

Content

- Context: Glossary, XML Schema Definition (XSD) and external sources
- A mini-vocabulary for DmfA declarations
- Translating from DmfA XML to RDF
- SHACL rules
- Demonstration
- Summary

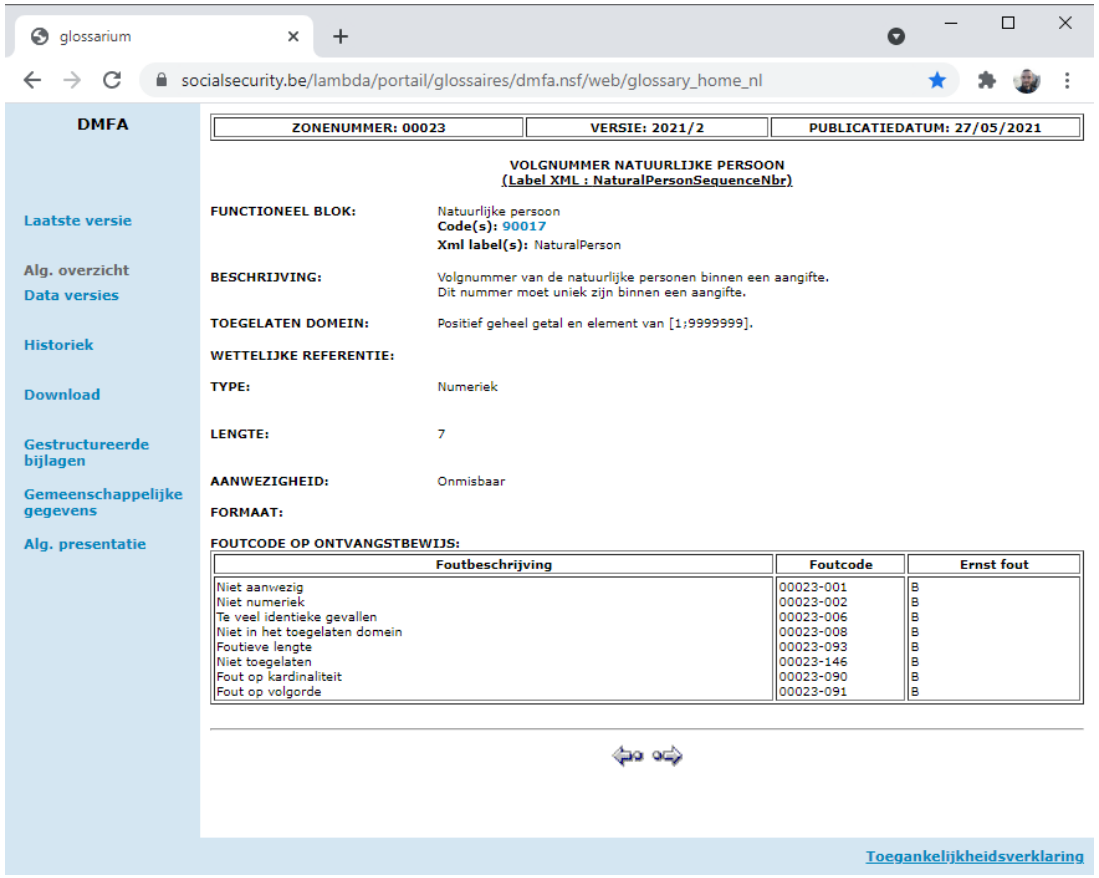
Attention: We are not going to replace DmfA and the glossary. The intention is to demonstrate the possibilities of SHACL for the validation of DmfA declarations. DmfA offers us a very complex matter with which we can demonstrate all the functionalities of SHACL.



Context

Glossarium

- Information about entities (called blocks), relationships and attributes in HTML, PDF, XML, and TXT
- The glossary contains information on validation rules. Simple rules about totality (must exist) and length are structured, but many rules are described in natural language and fall outside the expressiveness of XML. These must be tested by an application.
- In my humble opinion: encoding bias. The model is aimed at producing an XSD.



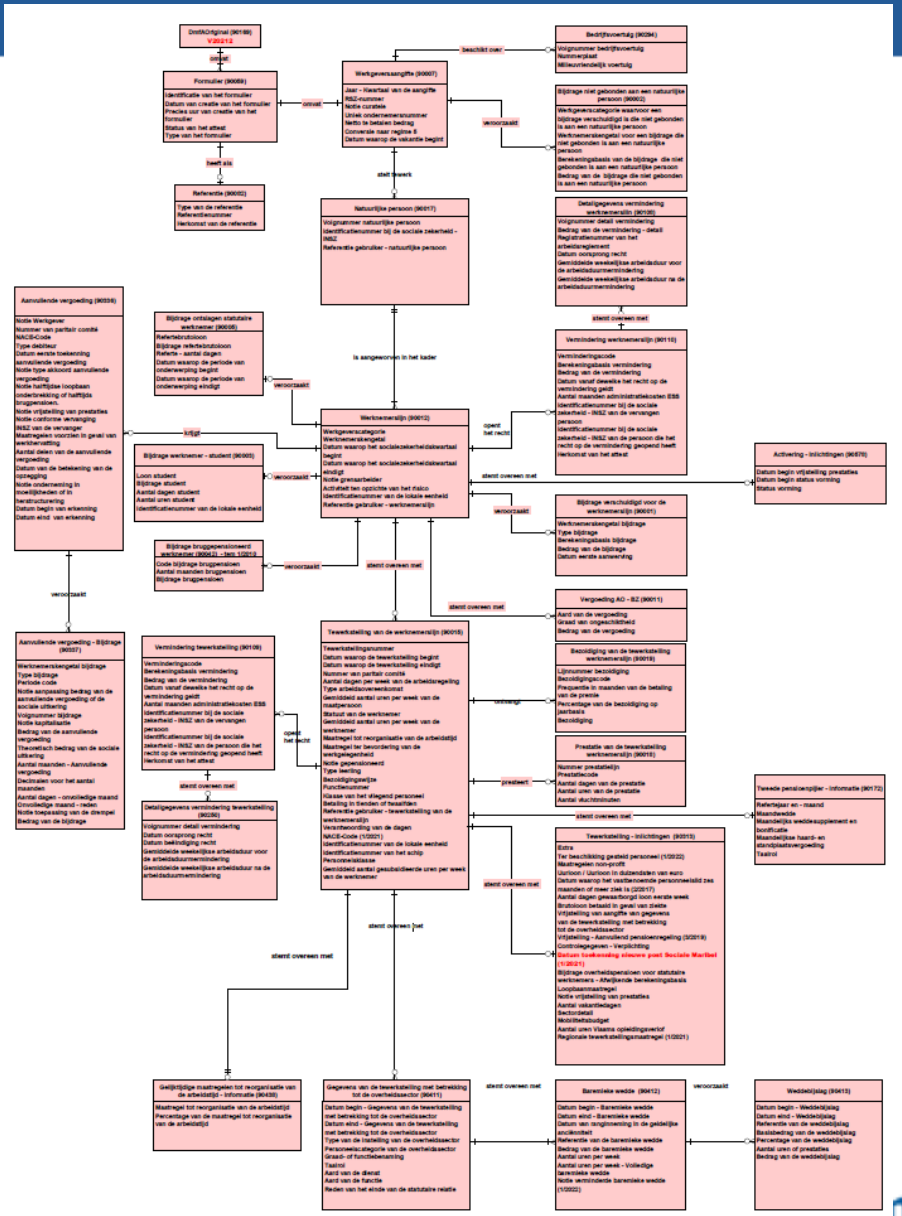
NaturalSequenceNbr is an attribute of the entity person. Sequence numbers are unique within a declaration, but this is not tested by XSD. An application must test this.

Source: https://www.socialsecurity.be/lambda/portail/glossaires/dmfa.nsf/web/glossary_home_nl

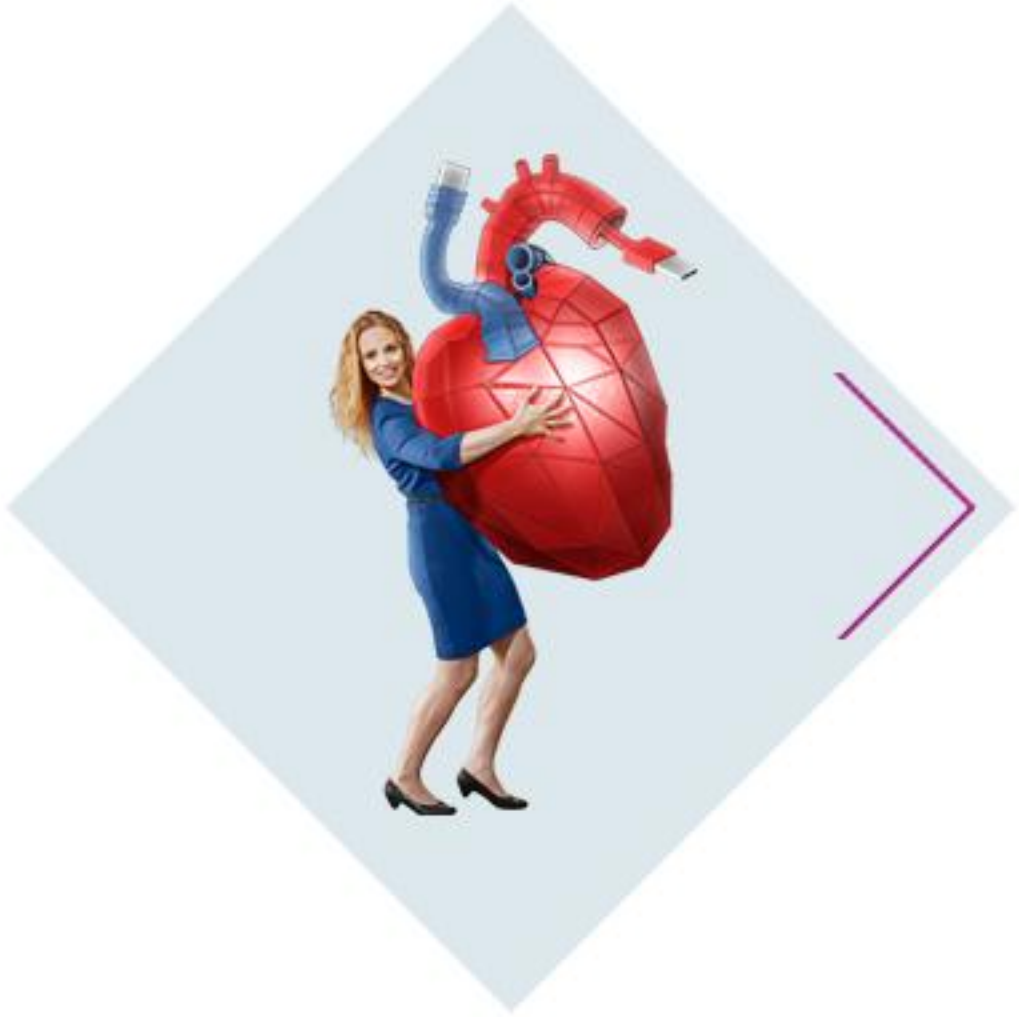
Smals ICT for society

Glossarium

- 29 entities (functional blocks)
- 28 relationships between entities (hierarchy of the blocks)
- 215 attributes (but not all is used)
- Some attributes are historical and documented, but their use is "forbidden"



ERD of a DmfA declaration



A vocabulary for DmfA declarations

Why an ~~ontology~~ vocabulary? (*)

- A direct mapping is a direct translation of data to RDF. The resulting RDF reflects the structure of the source, in this case the XML.
- Relationships in XML are implicit (see e.g. relationship between Form and Reference).
-

```
1  <?xml version="1.0" encoding="UTF-8"?>
2  <!-- edited with XMLSpy v2019 rel. 3 (x64) (http://www.altova.com) by  (Smals) -->
3  <DmfAOriginal xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="DmfAOriginal_20212.xsd">
4      <Form>
5          <Identification>DMFA</Identification>
6          <FormCreationDate>2021-05-28</FormCreationDate>
7          <FormCreationHour>08:14:45.000</FormCreationHour>
8          <AttestationStatus>0</AttestationStatus>
9          <TypeForm>SU</TypeForm>
10         <Reference>
11             <ReferenceType>1</ReferenceType>
12             <ReferenceOrigin>1</ReferenceOrigin>
13             <ReferenceNbr>14506543920212</ReferenceNbr>
14         </Reference>
15         <EmployerDeclaration>
16             <Quarter>20212</Quarter>
17             <NOSSRegistrationNbr>145065439</NOSSRegistrationNbr>
```

Attribute with value of entity Form

Relationship between Form and Reference: here the name of the type is also the relationship.

- Although a direct mapping with SHACL is possible, the SHACL rules would not be intuitive: "Do we think in functions of entities or relationships?" Consequently, a simple vocabulary seemed appropriate to me.

(*) A vocabulary is a "light" ontology. Few axioms and rules, but rich enough to distract interesting things. Examples are class hierarchies, role hierarchies, domains and ranges.

Smals ICT for society

Generating a vocabulary: entities

Glossary contains XML documentation that we can process automatically

Using a Jupyter Notebook (see GitLab)

Creating **types** with:

- BFLabelsXMLdmfa20212FR.xml
- FBLabelsXMLdmfa20212NL.xml

Documentation in NL, FR, and EN.

```
ont:CompanyVehicle a owl:Class ;
    rdfs:label "Company Vehicle"@en,
        "Véhicule de société"@fr,
        "Bedrijfsvoertuig"@nl ;
    dc:identifier "90294" .
```


Smals ICT for society

Generating a vocabulary: attributes

Glossary contains XML documentation that we can process automatically

Creating attributes with:

- ZonesLabelsXMLdmfa20212FR.xml
- ZonesLabelsXMLdmfa20212NL.xml

Documentation in NL, FR, and EN.

Encoding errors in XML require some preprocessing.

```
ont:ReducedScaleSalaryNotion a owl:DataProperty ;
    rdfs:label "Reduced Scale Salary Notion"@en,
        "NOTION TRAITEMENT BARÃMIQUE RÃDUIT"@fr,
        "NOTIE VERMINDERDE BAREMIEKE WEDDE"@nl ;
    dc:identifier "01235" ;
.
```


Smals ICT for society

Generating a vocabulary: entity-attributes

Glossary contains XML documentation that we can process automatically

Relationships between entities and attributes:

VersCplt_BFdmfa20212FR.xml

- ~~VersCplt_FBdmfa20212NL.xml~~

Only the relationships needed.

```
ont:ReducedScaleSalaryNotion a owl:DataProperty ;
    rdfs:label "Reduced Scale Salary Notion"@en,
        "NOTION TRAITEMENT BARÃMIQUE RÃDUIT"@fr,
        "NOTIE VERMINDERDE BAREMIEKE WEDDE"@nl ;
    rdfs:domain ont:ScaleSalary ;
    dc:identifier "01235" ;
.
```


Smals ICT for society

Generating a vocabulary: entity-attributes

Glossary contains XML documentation that we can process automatically

Documentation of the attributes:

- VersCpltdmfa20212FR.xml
- VersCpltdmfa20212NL.xml

Documentation in NL and FR.

```
ont:ReducedScaleSalaryNotion a owl:DataProperty ;
    rdfs:label "Reduced Scale Salary Notion"@en,
        "NOTION TRAITEMENT BARÃ¨MIQUE RÃ©DUIT"@fr,
        "NOTIE VERMINDERDE BAREMIEKE WEDDE"@nl ;
    rdfs:domain ont:ScaleSalary ;
    dc:identifier "01235" ;
    rdfs:comment "Zone qui … réduit."@fr, "Zone … is."@nl ;
    rdfs:isDefinedBy "1 = traitement barémique réduit"@fr,
        "1 = verminderd baremieke wedde"@nl ;
    ont:function "Traitement barémique"@fr, "Baremieke wedde"@nl ;
    ont:length 1 ;
    ont:presence "Obligatoire si …."@fr, "Verplicht indien …."@nl ;
    ont:type "Alphanumérique"@fr, "Alfanumeriek"@nl ;
.
```


Smals ICT for society

Generating a vocabulary: relationships

We processed relationships between entities manually using the ERD.

We could use the XSD, but relationships in XSD had no names. Furthermore, there are relationships with the same names. In the long run, it is better to make a distinction (at least at the level of the vocabulary).

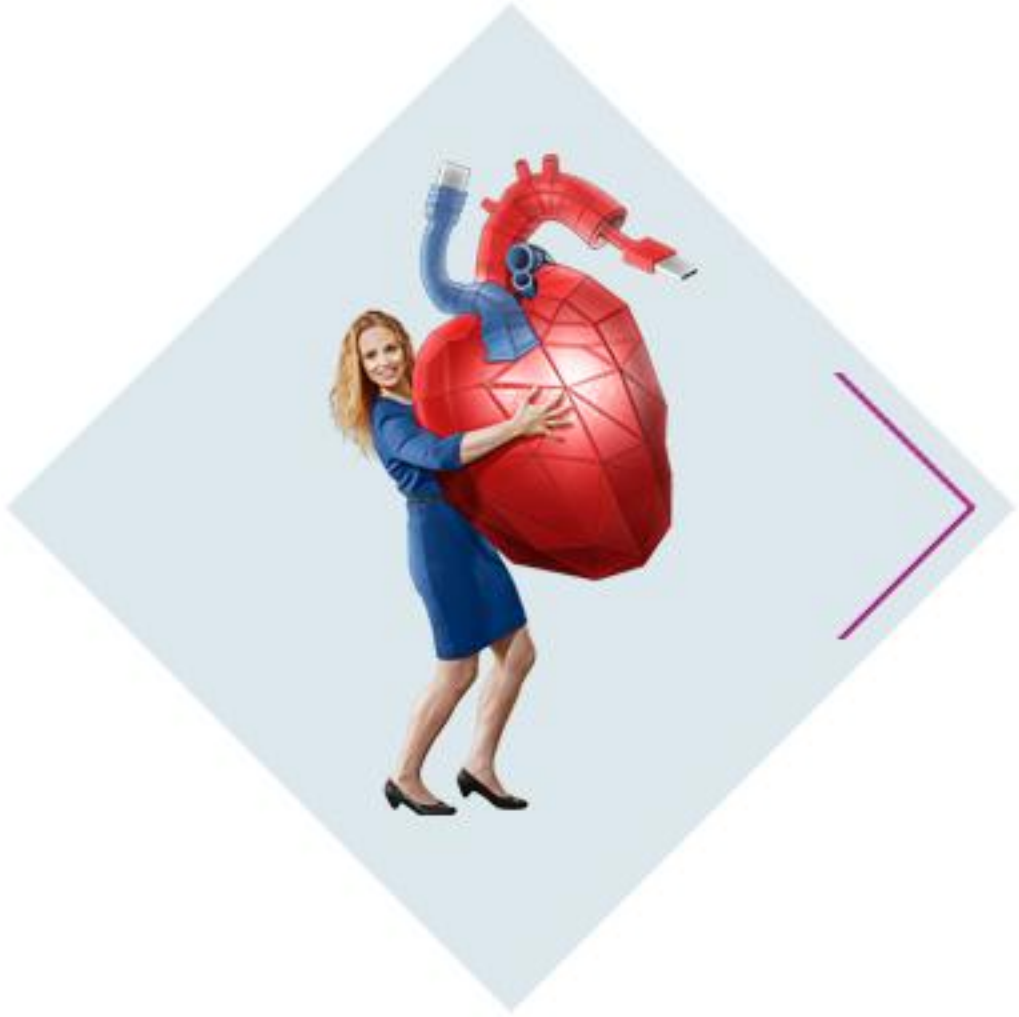
A CSV file was created, and it was transformed:

FROM	TO	LABELNL	LABELFR	LABELEN
90169	90059	omvat	reprent	contains
90059	90082	heeft als	a comme	has
90059	90007	omvat	comprent	contains
90007	90294	beschikt over	dispose de	has
…	…	…	…	…

```
ont:R_9000790294 rdfs:label "has"@en,
        "dispose de"@fr,
        "beschikt over"@nl ;
    rdfs:domain ont:EmployerDeclaration ;
    rdfs:range ont:CompanyVehicle .
```

Note that I combine the codes of two entities for the computer label of the relationship. In this way, different "has" relationships are different.

Smals ICT for society



From a DmfA XML declaration to RDF

From XML to RDF

- Based on the XML file FI.DMFA.101141.20210604.02500.T.1.1 (supplied by Smals)
- Used 11 of the 29 entities. So not everything was mapped.

- Mapping using RML (https://rml.io/) the "successor" of R2RML for relational databases, CSV, JSON, XML,...

- The mapping is simple, it is almost a one-on-one mapping. Data types such as xsd:string, xsd:integer, xsd:date,... must be supplied.
- To be supplemented here and there when SHACL rules are completed.

From XML to RDF

mapping.ttl

dmfa-aangiften > mapping.ttl

```
@prefix rr: <http://www.w3.org/ns/r2rml#> .
@prefix rml: <http://semweb.mmlab.be/ns/rml#> .
@prefix ql: <http://semweb.mmlab.be/ns/ql#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

@prefix ont: <https://kg.socialsecurity.be/ont/dmfa#> .
@base <https://kg.socialsecurity.be/resource/>.

<#DmfAOriginal-Mapping>
    a rr:TriplesMap ;

    rml:logicalSource [
        rml:source "data.xml" ;
        rml:iterator "/DmfAOriginal" ;
        rml:referenceFormulation ql:XPath ;
    ] ;

    rr:subjectMap [
        rr:template "DmfAOriginal-{Form/Identification}" ;
        rr:termType rr:BlankNode ;
        rr:class ont:DmfAOriginal ;
    ] ;

    rr:predicateObjectMap [
        rr:predicate ont:R_9016990059 ;
        rr:objectMap [ rr:parentTriplesMap <#Form-Mapping> ; ] ;
    ] ;
.

<#Form-Mapping>
    a rr:TriplesMap ;

    rml:logicalSource [
        rml:source "data.xml" ;
        rml:iterator "/DmfAOriginal/Form" ;
        rml:referenceFormulation ql:XPath ;
    ] ;

    rr:subjectMap [
        rr:template "Form" ;
        rr:termType rr:BlankNode ;
        rr:class ont:Form ;
    ] ;

    rr:predicateObjectMap [
        rr:predicate ont:AttestationStatus ;
        rr:objectMap [ rml:reference "AttestationStatus" ; rr:datatype xsd:integer ; ] ;
    ] ;
    rr:predicateObjectMap [
        rr:predicate ont:FormCreationDate ;
        rr:objectMap [ rml:reference "FormCreationDate" ; rr:datatype xsd:date ;  ] ;
    ] ;
    rr:predicateObjectMap [
        rr:predicate ont:FormCreationHour ;
```

FI.DMFA.101141.20210604.02500.T.1.1.output.nq

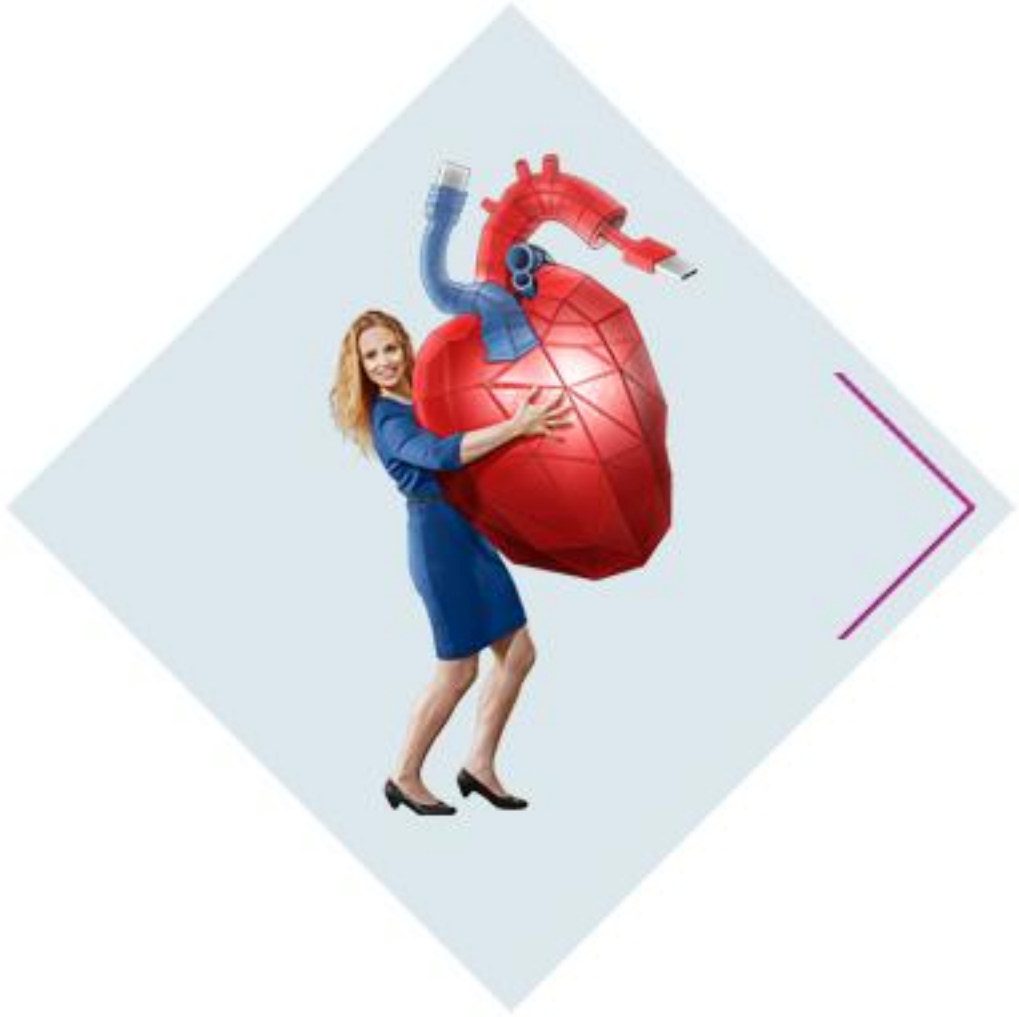
dmfa-aangiften > FI.DMFA.101141.20210604.02500.T.1.1.output.nq

```
_:DmfAOriginal-DMFA <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://kg.socialsecurity.be/ont/dmf
_:DmfAOriginal-DMFA <https://kg.socialsecurity.be/ont/dmfa#R_9016990059> _:Form.
_:WorkerContribution <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://kg.socialsecurity.be/ont/dm
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "471636".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionCalculationBasis> "1080000".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionType> "1".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "015".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "1404".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionType> "0".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "809".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "1512".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "810".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "1080".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "859".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "21060".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "825".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "24300".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "820".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "852".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionAmount> "0".
_:WorkerContribution <https://kg.socialsecurity.be/ont/dmfa#ContributionWorkerCode> "856".
_:OccupationSequenceNbr1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://kg.socialsecurity.be/on
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#ContractType> "0".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#JointCommissionNbr> "112".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#LocalUnitID> "2299890586".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#MeanWorkingHours> "3800".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#OccupationSequenceNbr> "1".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#OccupationStartingDate> "2021-03-22"^^<http
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#RefMeanWorkingHours> "3800".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#Retired> "0".
_:OccupationSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#WorkingDaysSystem> "500".
_:DeductionDetailSequenceNbr1 <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <https://kg.socialsecurity.
_:DeductionDetailSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#AverageWorkingTimeAfterReduction> "285
_:DeductionDetailSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#AverageWorkingTimeBeforeReduction> "38
_:DeductionDetailSequenceNbr1 <https://kg.socialsecurity.be/ont/dmfa#DeductionDetailSequenceNbr> "1".
```

Because we are going to validate declarations locally and do not have enough information, entities of different DmfA documents will receive the same identifier.

If we are going to validate a collection of declarations, then we have to change the approach a bit (e.g., with a UUID). This is for the mapping and the application, and not for the SHACL rules.

nals
or society



SHACL rules

Generating SHACL

- The generation of SHACL from the glossary is limited (type, length, etc.). There was enough for one form of a skeleton but:
  -
  - Contradictions between glossary and rules
  - Sometimes errors in the XSD (?!)
  - Inconsistent approach to some constraints (such as length—exact length vs. maximum length)
- The skeleton therefore contains errors (beyond my abilities) and everything has to be tested and refined manually. I only gained a little time because of this... So I removed the code for the skeleton. This process can be better aligned with the generation of the XSD.

SHACL: examples

kg:DmfAOriginalCountInstancesShape: a document must contain exactly 1 entity of the type ont:DmfAOriginal, otherwise we do not have a declaration.

kg:DmfAOriginalShape: a ont:DmfAOriginal must contain at least 1 ont:Form.

```
kg:DmfAOriginalCountInstancesShape a sh:NodeShape ;
    rdfs:comment "A graph should have exactly one instance of ont:DmfAOriginal" ;
    sh:targetNode ont:DmfAOriginal ;

    sh:property [
        sh:message "Graph does not contain exactly one instance of ont:DmfAOriginal" ;
        sh:path [ sh:inversePath rdf:type ] ;
        sh:maxCount 1 ;
        sh:minCount 1 ;
    ] ;
. ] ;

kg:DmfAOriginalShape a sh:NodeShape ;
    rdfs:comment "Property Shape for DmfAOriginal (code)" ;
    sh:targetClass ont:DmfAOriginal ;

    sh:property [
        rdfs:comment "ont:DmfAOriginal must have at least one ont:Form." ;
        sh:minCount 1 ;
        sh:path ont:R_9016990059

.
```


Smals ICT for society

SHACL: voorbeelden

kg:FormShape

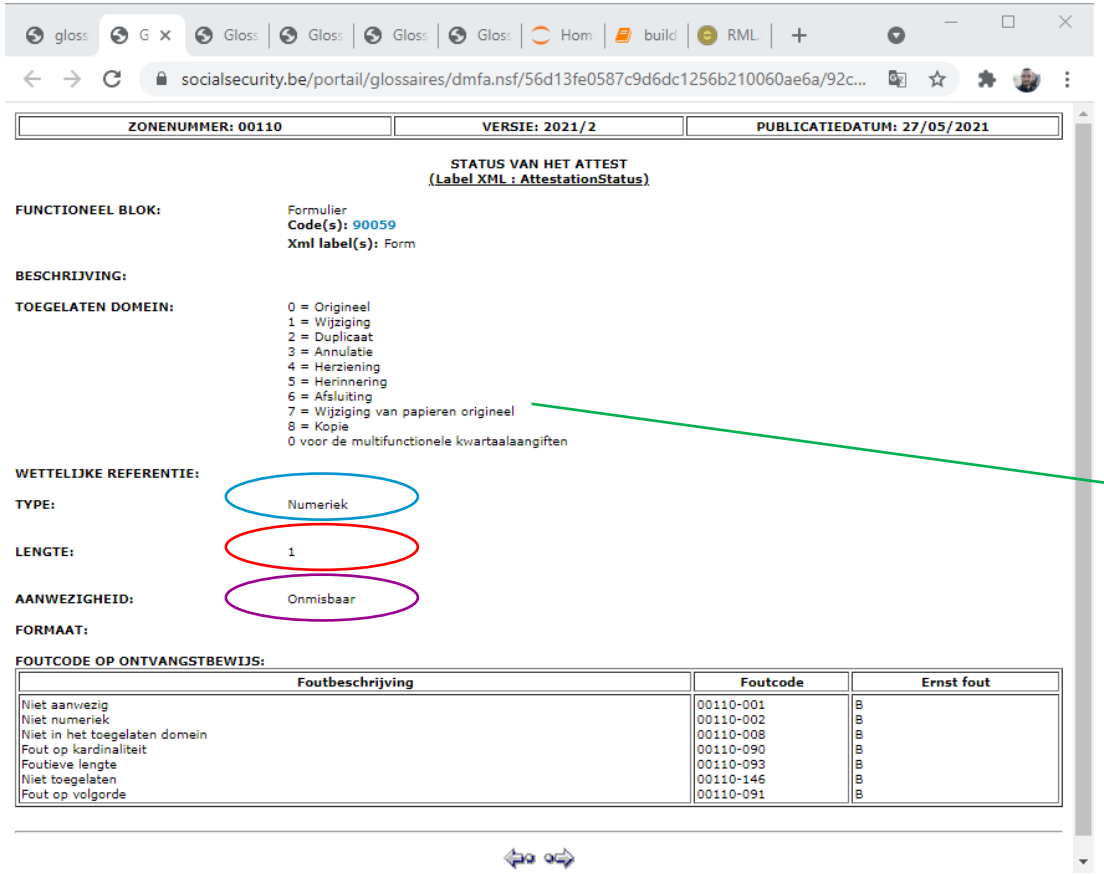
- Constraints on relationships with DmfAOriginal, Reference, and EmployerDeclaration
- Constraints over the attributes

We now choose a number of interesting attributes!

```
kg:FormShape a sh:NodeShape ;
    rdfs:comment "Property Shape for Form (code)" ;
    sh:targetClass ont:Form ;
    sh:property [
        rdfs:comment "Each ont:Form must belong to exactly one ont:DmfAOriginal." ;
        sh:maxCount 1 ; sh:minCount 1 ;
        sh:path [ sh:inversePath ont:R_9016990059 ; ] ;
    ] ;
    sh:property [
        rdfs:comment "ont:Form has at most one ont:Reference." ;
        sh:maxCount 1 ;
        sh:path ont:R_9005990082 ;
    ] ;
    sh:property [
        rdfs:comment "Each ont:Form has exactly one ont:EmployerDeclaration." ;
        sh:maxCount 1 ; sh:minCount 1 ;
        sh:path ont:R_9005990007 ;
    ] ;
    sh:property kg:AttestationStatusShape ;
    sh:property kg:FormCreationDateShape ;
    sh:property kg:FormCreationHourShape ;
    sh:property kg:IdentificationShape ;
    sh:property kg:TypeFormShape ;
.
```


Smals ICT for society

SHACL: examples



socialsecurity.be/portail/glossaires/dmfa.nsf/56d13fe0587c9d6dc1256b210060ae6a/92c...

ZONENUMMER: 00110	VERSIE: 2021/2	PUBLICATIEDATUM: 27/05/2021

STATUS VAN HET ATTEST
(Label XML : AttestationStatus)

FUNCTIONEEL BLOK: Formulier
Code(s): 90059
Xml label(s): Form

BESCHRIJVING:

TOEGELATEN DOMEIN:
0 = Origineel
1 = Wijziging
2 = Duplicaat
3 = Annulatie
4 = Herziening
5 = Herinnering
6 = Afsluiting
7 = Wijziging van papieren origineel
8 = Kopie
0 voor de multifunctionele kwartaalaangiften

WETTELIJKE REFERENTIE:

TYPE: Numeriek

LENGTE: 1

AANWEZIGHEID: Onmisbaar

FORMAAT:

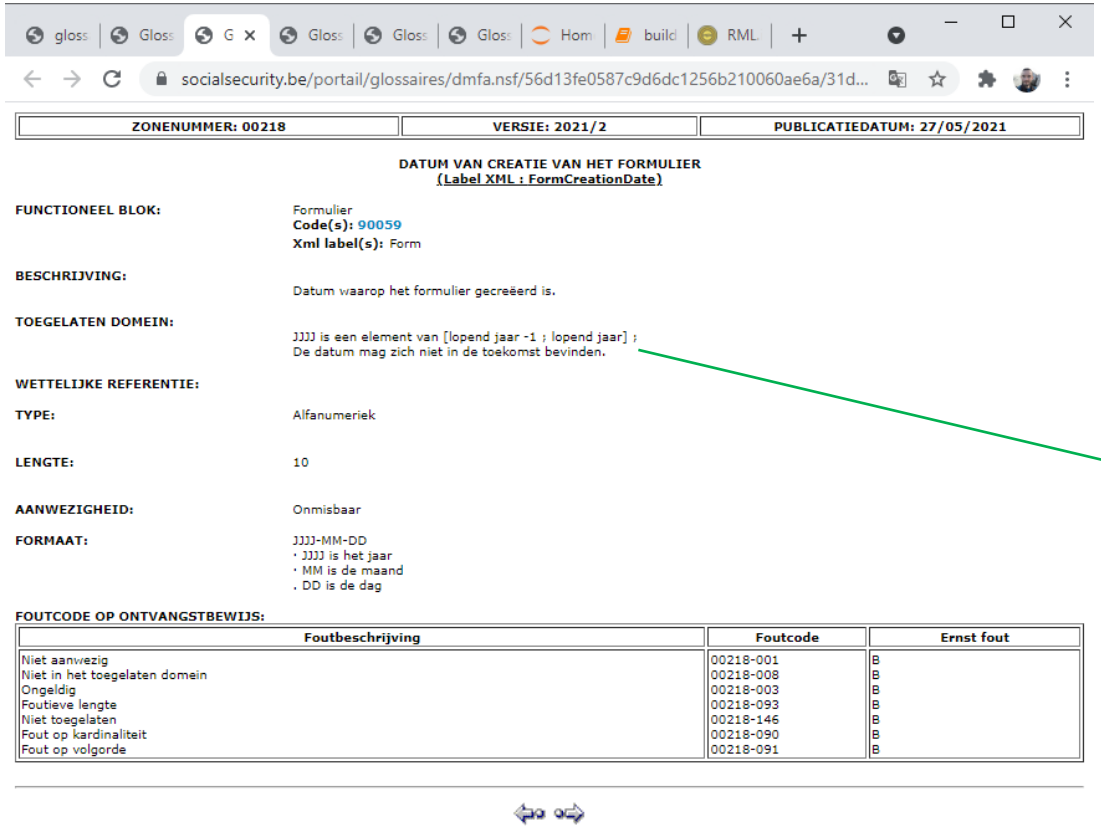
FOUTCODE OP ONTVANGSTBEWIJS:

Foutbeschrijving	Foutcode	Ernst fout
Niet aanwezig	00110-001	B
Niet numeriek	00110-002	B
Niet in het toegelaten domein	00110-008	B
Fout op kardinaliteit	00110-090	B
Foutieve lengte	00110-093	B
Niet toegelaten	00110-146	B
Fout op volgorde	00110-091	B

```
kg:AttestationStatusShape a sh:PropertyShape ;

    rdfs:comment "Property Shape for AttestationStatus
(code)" ;

    sh:path ont:AttestationStatus ;

    sh:datatype xsd:integer ;

    sh:minCount 1 ;

    sh:maxCount 1 ;

    sh:minLength 1 ;

    sh:maxLength 1 ;

    sh:minInclusive 0 ;

    sh:maxInclusive 8 ;

.
```


Smals ICT for society

SHACL: examples



socialsecurity.be/portail/glossaires/dmfa.nsf/56d13fe0587c9d6dc1256b210060ae6a/31d...

ZONENUMMER: 00218	VERSIE: 2021/2	PUBLICATIEDATUM: 27/05/2021

DATUM VAN CREATIE VAN HET FORMULIER
(Label XML : FormCreationDate)

FUNCTIONEEL BLOK: Formulier
Code(s): 90059
Xml label(s): Form

BESCHRIJVING: Datum waarop het formulier gecreëerd is.

TOEGELATEN DOMEIN: JJJJ is een element van [lopend jaar -1 ; lopend jaar] ;
De datum mag zich niet in de toekomst bevinden.

WETTELIJKE REFERENTIE:

TYPE: Alfanumeriek

LENGTE: 10

AANWEZIGHEID: Onmisbaar

FORMAAT: JJJJ-MM-DD
· JJJJ is het jaar
· MM is de maand
. DD is de dag

FOUTCODE OP ONTVANGSTBEWIJS:

Foutbeschrijving	Foutcode	Ernst fout
Niet aanwezig	00218-001	B
Niet in het toegelaten domein	00218-008	B
Ongeldig	00218-003	B
Foutieve lengte	00218-093	B
Niet toegelaten	00218-146	B
Fout op kardinaliteit	00218-090	B
Fout op volgorde	00218-091	B

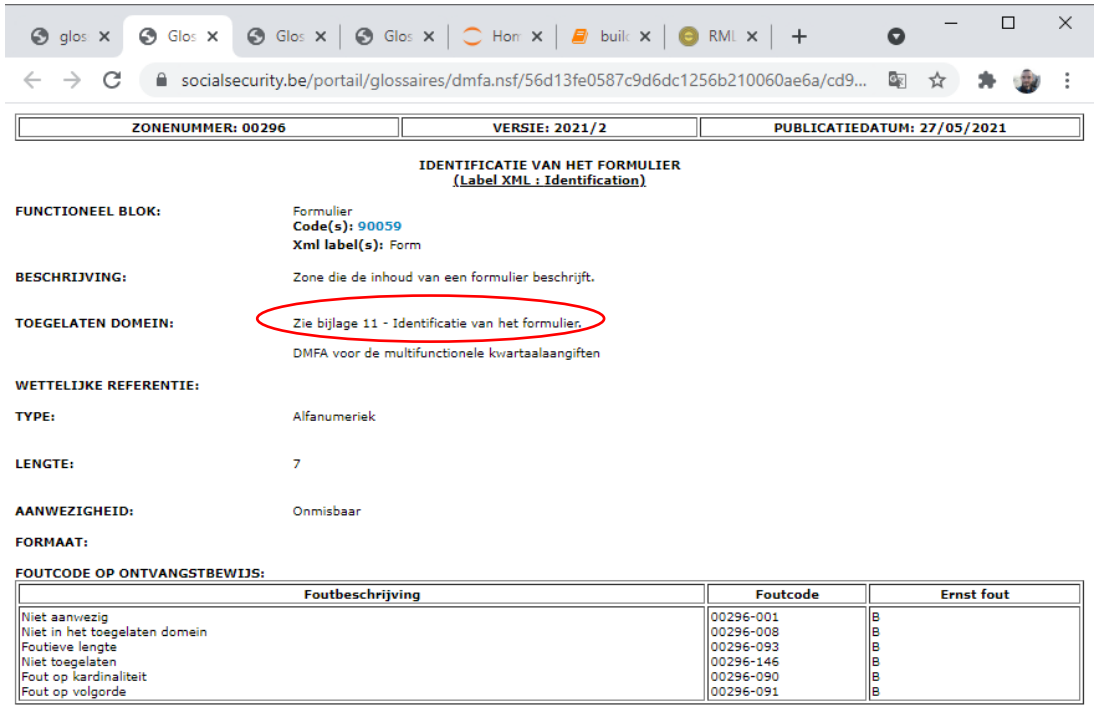
```
kg:FormCreationDateShape a sh:PropertyShape ;
    rdfs:comment "Property Shape for FormCreationDate (code)" ;
    sh:path ont:FormCreationDate ;
    sh:datatype xsd:date ;
    sh:maxLength 10 ;
    sh:minLength 10 ;
    sh:pattern "\\d{4}-\\d{2}-\\d{2}" ;

    sh:sparql [
      sh:message "Year beyond [current year - 1, current year]." ;
      sh:prefixes <> ;
      sh:select """
        SELECT $this (ont:FormCreationDate AS ?path) ?value
        WHERE {
          $this ont:FormCreationDate ?value .
          BIND ( YEAR( ?value ) AS ?y1 )
          BIND ( YEAR( NOW() ) AS ?y2 )
          FILTER ( !( ?y2 - ?y1 = 0 || ?y2 - ?y1 = 1 ) )
      }""" ;
    ] ;
.
```

We use the expressiveness of the graph query language

In this preview, the sh:pattern is actually superfluous, because this is a condition of the xsd:date datatype.

Smals ICT for society

SHACL: examples



socialsecurity.be/portail/glossaires/dmfa.nsf/56d13fe0587c9d6dc1256b210060ae6a/cd9...

ZONENUMMER: 00296	VERSIE: 2021/2	PUBLICATIEDATUM: 27/05/2021

IDENTIFICATIE VAN HET FORMULIER
(Label XML : Identification)

FUNCTIONEEL BLOK: Formulier
Code(s): 90059
Xml label(s): Form

BESCHRIJVING: Zone die de inhoud van een formulier beschrijft.

TOEGELATEN DOMEIN: Zie bijlage 11 - Identificatie van het formulier.
DMFA voor de multifunctionele kwartaalaangiften

WETTELIJKE REFERENTIE:

TYPE: Alfanumeriek

LENGTE: 7

AANWEZIGHEID: Onmisbaar

FORMAAT:

FOUTCODE OP ONTVANGSTBEWIJS:

Foutbeschrijving	Foutcode	Ernst fout
Niet aanwezig	00296-001	B
Niet in het toegelaten domein	00296-008	B
Foutieve lengte	00296-093	B
Niet toegelaten	00296-146	B
Fout op kardinaliteit	00296-090	B
Fout op volgorde	00296-091	B

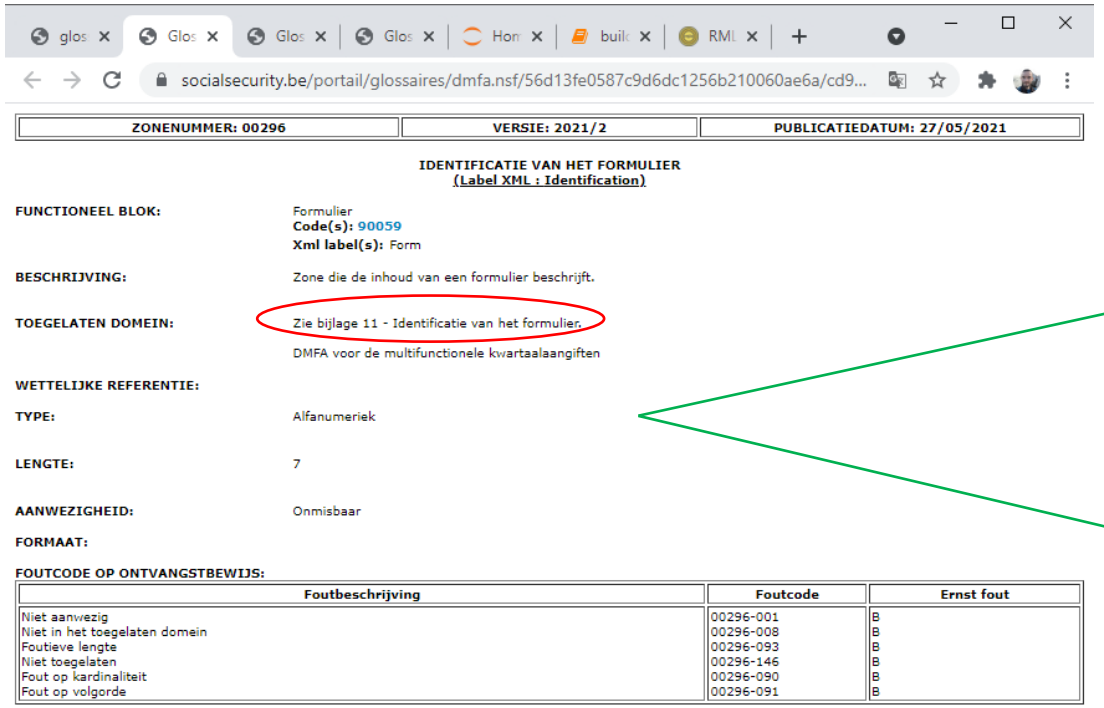
This is an example of how the different identification types in RDF can be represented:

```
kg:IdentificationTypeDMFA a ont:IdentificationType ;
    rdfs:label "DMFA" ;
    rdfs:comment "Multifunctionele ... werkgever" ;
    ont:validFrom "1900-01-01"^^xsd:date ;
    ont:validTo "9999-01-01"^^xsd:date ;
.
```

With such representations in our data graph (or even knowledge graph) we can:

Smals ICT for society

SHACL: voorbeelden



socialsecurity.be/portail/glossaires/dmfa.nsf/56d13fe0587c9d6dc1256b210060ae6a/cd9...

ZONENUMMER: 00296	VERSIE: 2021/2	PUBLICATIEDATUM: 27/05/2021

IDENTIFICATIE VAN HET FORMULIER
(Label XML : Identification)

FUNCTIONEEL BLOK: Formulier
Code(s): 90059
Xml label(s): Form

BESCHRIJVING: Zone die de inhoud van een formulier beschrijft.

TOEGELATEN DOMEIN: Zie bijlage 11 - Identificatie van het formulier.
DMFA voor de multifunctionele kwartaalaangiften

WETTELIJKE REFERENTIE:

TYPE: Alfanumeriek

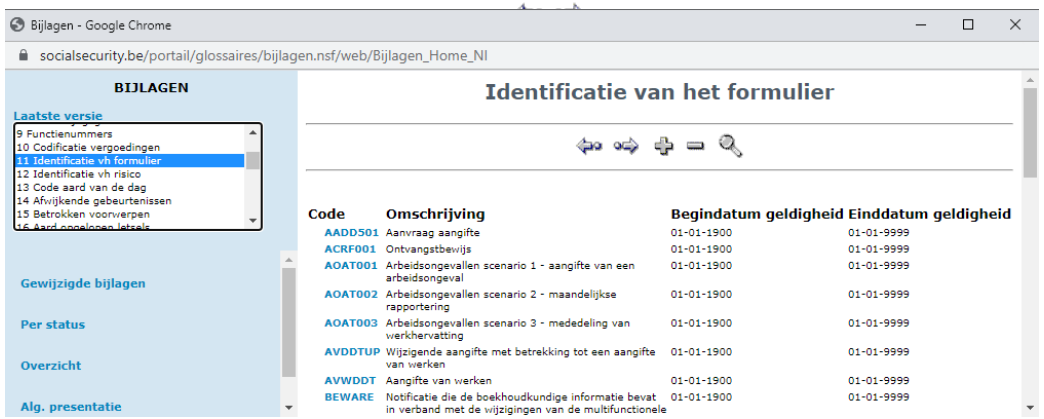
LENGTE: 7

AANWEZIGHEID: Onmisbaar

FORMAAT:

FOUTCODE OP ONTVANGSTBEWIJS:

Foutbeschrijving	Foutcode	Ernst fout
Niet aanwezig	00296-001	B
Niet in het toegelaten domein	00296-008	B
Foutieve lengte	00296-093	B
Niet toegelaten	00296-146	B
Fout op kardinaliteit	00296-090	B
Fout op volgorde	00296-091	B



Bijlagen - Google Chrome

socialsecurity.be/portail/glossaires/bijlagen.nsf/web/Bijlagen_Home_NI

BIJLAGEN

Laatste versie

9 Functienummers
10 Codificatie vergoedingen
11 Identificatie vh formulier
12 Identificatie vh risico
13 Code aard van de dag
14 Afwijkende gebeurtenissen
15 Betrokken voorwerpen
16 Aard opgelopen letsels

Gewijzigde bijlagen

Per status

Overzicht

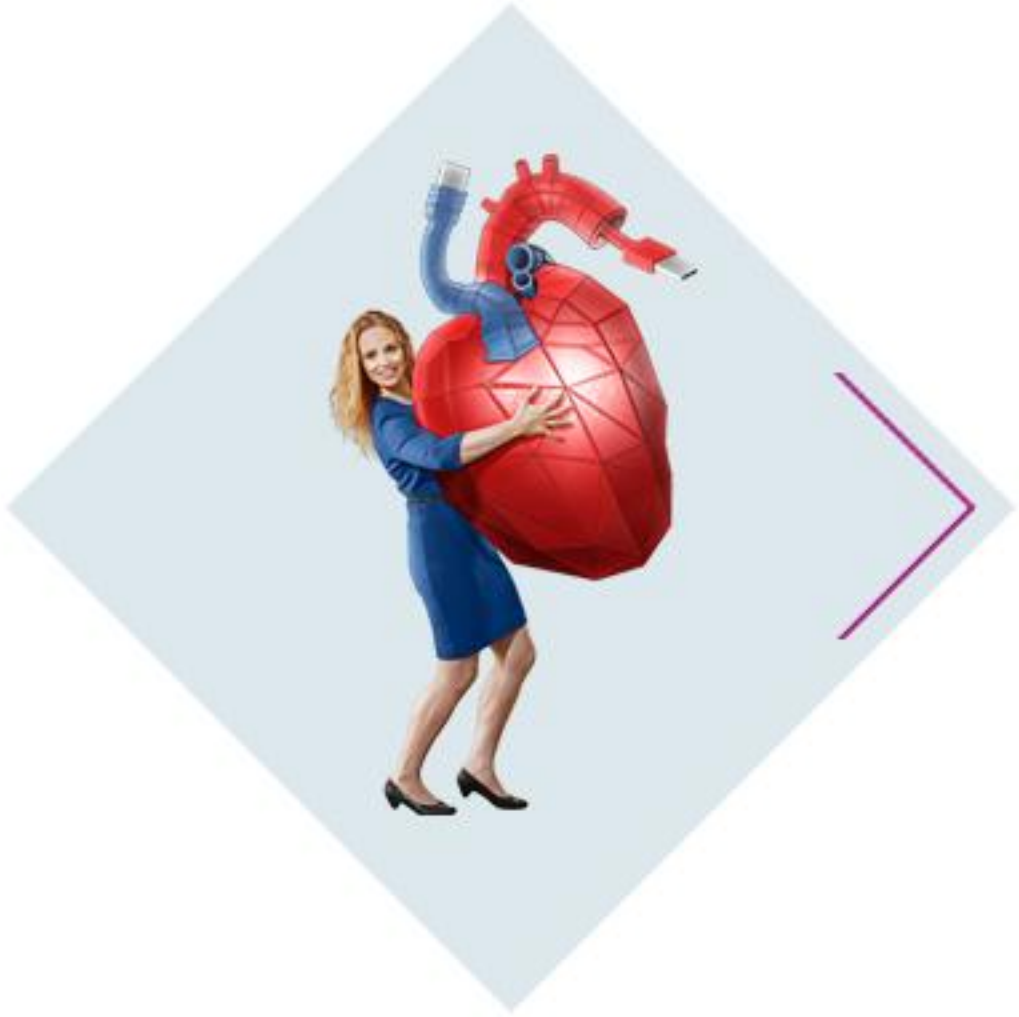
Alg. presentatie

Identificatie van het formulier

Code	Omschrijving	Begindatum geldigheid	Einddatum geldigheid
AADD501	Aanvraag aangifte	01-01-1900	01-01-9999
ACRF001	Ontvangstbewijs	01-01-1900	01-01-9999
AOAT001	Arbeidsongevallen scenario 1 - aangifte van een arbeidsongeval	01-01-1900	01-01-9999
AOAT002	Arbeidsongevallen scenario 2 - maandelijkse rapportering	01-01-1900	01-01-9999
AOAT003	Arbeidsongevallen scenario 3 - mededeling van werkhervatting	01-01-1900	01-01-9999
AVDDTUP	Wijzigende aangifte met betrekking tot een aangifte van werken	01-01-1900	01-01-9999
AVWDDT	Aangifte van werken	01-01-1900	01-01-9999
BEWARE	Notificatie die de boekhoudkundige informatie bevat in verband met de wijzigingen van de multifunctionele	01-01-1900	01-01-9999

```
kg:IdentificationShape a sh:PropertyShape ;
    rdfs:comment "Property Shape for Identification (code)" ;
    sh:path ont:Identification ; sh:datatype xsd:string ;
    sh:maxLength 7 ; sh:minCount 1 ; sh:maxCount 1 ;
    # Geldige waarden komen van bijlage 1
    sh:in ( "AADD501" "ACRF001" ... "ZIMA006" ) ;
    # De geldige waarden hebben ook een geldigheidsdatum,
    # we demonstreren dit voor 1 entiteit
    sh:sparql [
      sh:message "Date not within valid interval of document type" ;
      sh:prefixes <> ;
        sh:select """
          SELECT $this (ont:Identification AS ?path) ?value
          WHERE {
            $this ont:Identification ?value .
            $this ont:FormCreationDate ?date .
            ?idtype a ont:IdentificationType ; rdfs:label ?value ;
                    ont:validFrom ?from ; ont:validTo ?to .
            FILTER ( !( ?date >= ?from && ?date <= ?to ) )
          }""" ;
] .
```

Smals ICT for society



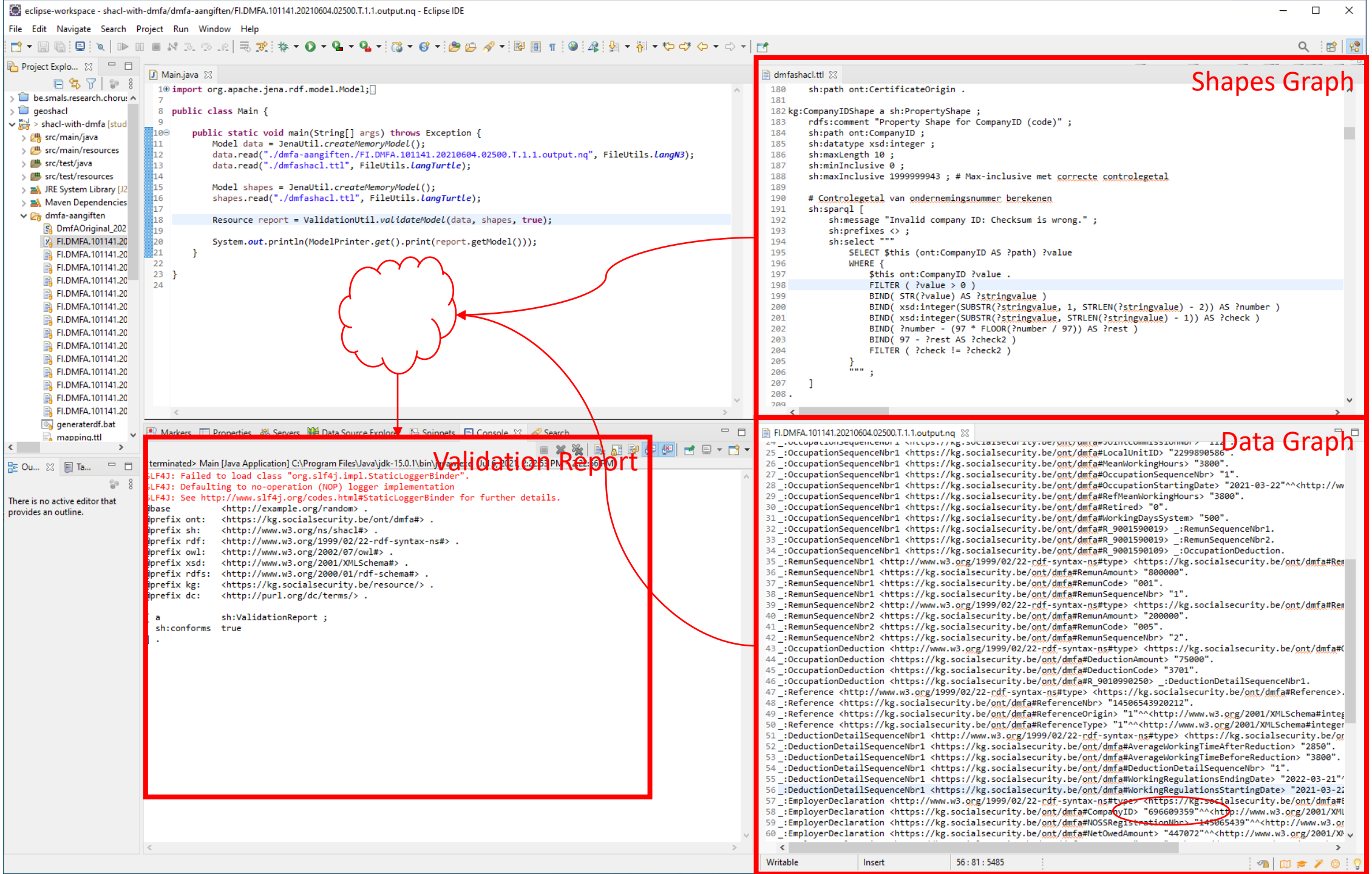
A small demonstration

Demonstratie

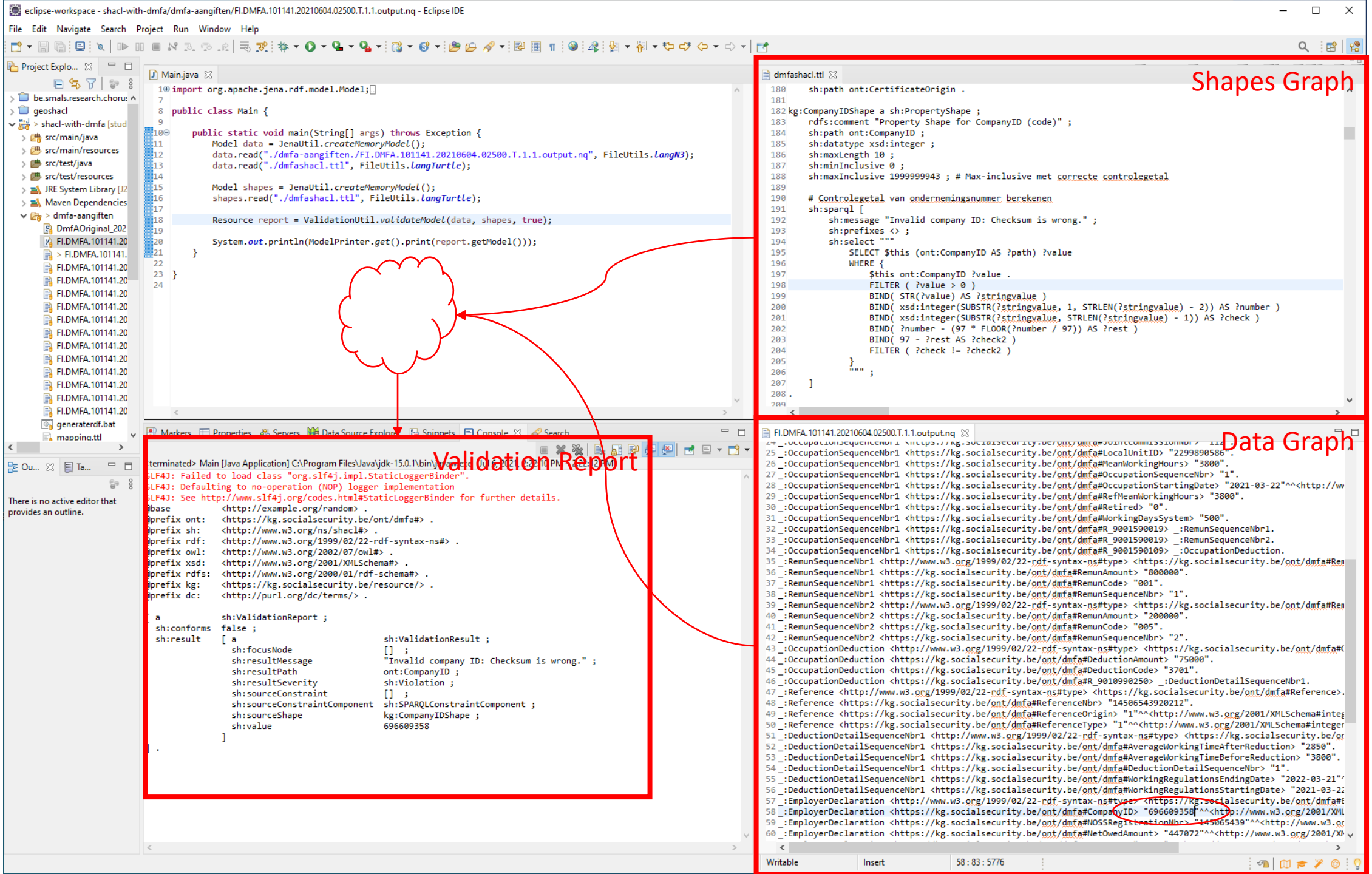
- By means of the TopBraid library
-
- Data graph = DmfA declaration in RDF + extra data
- Shapes graph

- Focus on calculation control number
-

Controle nummer is OK



Controle nummer Is niet OK



SHACL: examples

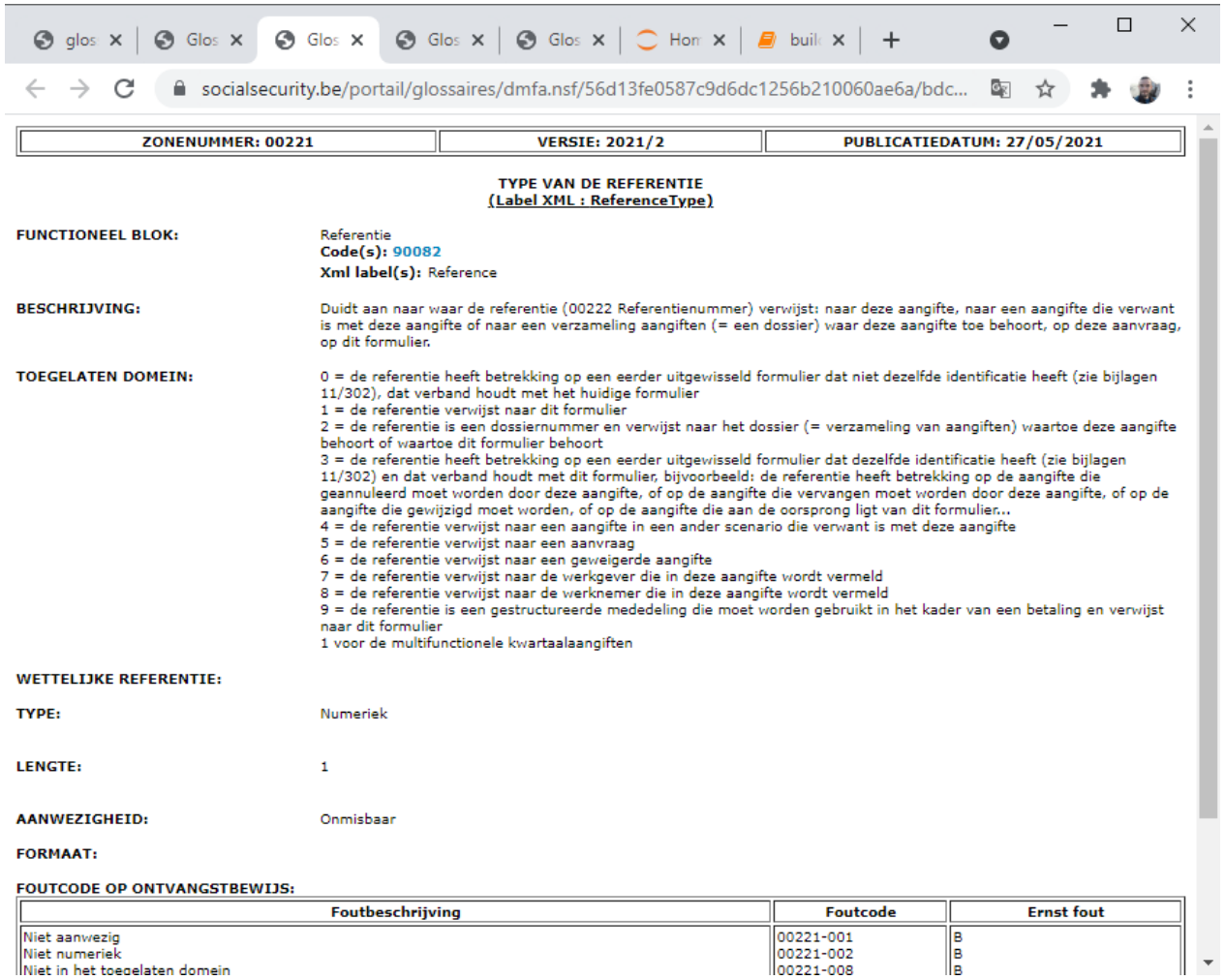
```
kg:CompanyIDShape a sh:PropertyShape ;
    rdfs:comment "Property Shape for CompanyID (code)" ;
    sh:path ont:CompanyID ; sh:datatype xsd:integer ; sh:maxLength 10 ;
    sh:minInclusive 0 ; sh:maxInclusive 1999999943 ; # Max-inclusive met correcte controlegetal

    # Controlegetal van ondernemingsnummer berekenen (ofwel 0, ofwel correct volgens berekening_
    sh:sparql [
      sh:message "Invalid company ID: Checksum is wrong." ;
      sh:prefixes <> ;
      sh:select """
        SELECT $this (ont:CompanyID AS ?path) ?value
        WHERE {
          $this ont:CompanyID ?value .
          FILTER ( ?value > 0 )
          BIND( STR(?value) AS ?stringvalue )
          BIND( xsd:integer(SUBSTR(?stringvalue, 1, STRLEN(?stringvalue) - 2)) AS ?number )
          BIND( xsd:integer(SUBSTR(?stringvalue, STRLEN(?stringvalue) - 1)) AS ?check )
          BIND( ?number - (97 * FLOOR(?number / 97)) AS ?rest )
          BIND( 97 - ?rest AS ?check2 )
          FILTER ( ?check != ?check2 )
        } """ ;
] .
```

Check company numbers! This rule can later be generalized so that it can be reused within an organization.

Smals ICT for society

Into the future?

- Some descriptions refer not only to external data (see example Appendix 11), but also to past declarations.
-
- This experiment involves local validation. Can we validate globally? In theory, yes, but then we have to store all declarations in a graph.
-



socialsecurity.be/portail/glossaires/dmfa.nsf/56d13fe0587c9d6dc1256b210060ae6a/bdc...

ZONENUMMER: 00221	VERSIE: 2021/2	PUBLICATIEDATUM: 27/05/2021

TYPE VAN DE REFERENTIE
(Label XML : ReferenceType)

FUNCTIONEEL BLOK: Referentie
Code(s): 90082
Xml label(s): Reference

BESCHRIJVING: Duidt aan naar waar de referentie (00222 Referentienummer) verwijst: naar deze aangifte, naar een aangifte die verwant is met deze aangifte of naar een verzameling aangiften (= een dossier) waar deze aangifte toe behoort, op deze aanvraag, op dit formulier.

TOEGELATEN DOMEIN:
0 = de referentie heeft betrekking op een eerder uitgewisseld formulier dat niet dezelfde identificatie heeft (zie bijlagen 11/302), dat verband houdt met het huidige formulier
1 = de referentie verwijst naar dit formulier
2 = de referentie is een dossiernummer en verwijst naar het dossier (= verzameling van aangiften) waartoe deze aangifte behoort of waartoe dit formulier behoort
3 = de referentie heeft betrekking op een eerder uitgewisseld formulier dat dezelfde identificatie heeft (zie bijlagen 11/302) en dat verband houdt met dit formulier, bijvoorbeeld: de referentie heeft betrekking op de aangifte die geannuleerd moet worden door deze aangifte, of op de aangifte die vervangen moet worden door deze aangifte, of op de aangifte die gewijzigd moet worden, of op de aangifte die aan de oorsprong ligt van dit formulier...
4 = de referentie verwijst naar een aangifte in een ander scenario die verwant is met deze aangifte
5 = de referentie verwijst naar een aanvraag
6 = de referentie verwijst naar een geweigerde aangifte
7 = de referentie verwijst naar de werkgever die in deze aangifte wordt vermeld
8 = de referentie verwijst naar de werknemer die in deze aangifte wordt vermeld
9 = de referentie is een gestructureerde mededeling die moet worden gebruikt in het kader van een betaling en verwijst naar dit formulier
1 voor de multifunctionele kwartaalaangiften

WETTELIJKE REFERENTIE:

TYPE: Numeriek

LENGTE: 1

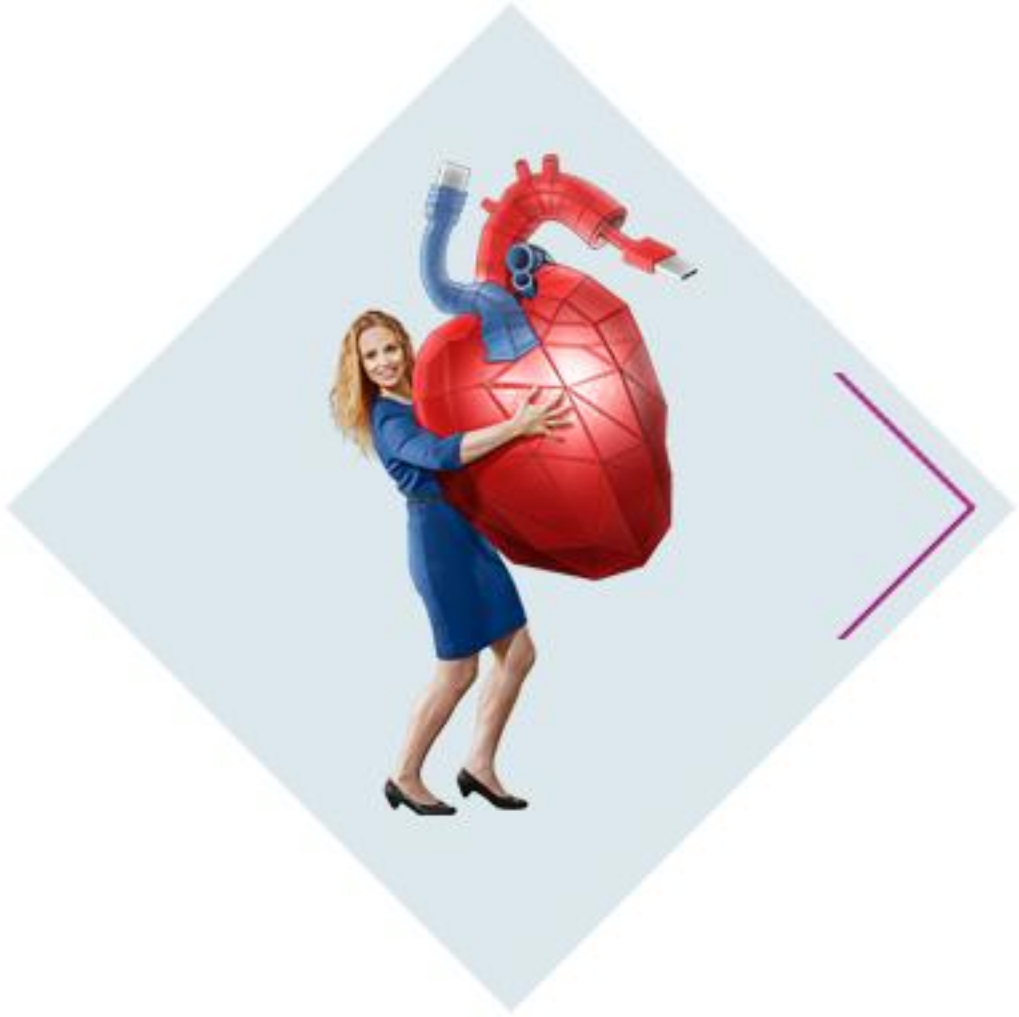
AANWEZIGHEID: Onmisbaar

FORMAAT:

FOUTCODE OP ONTVANGSTBEWIJS:

Foutbeschrijving	Foutcode	Ernst fout
Niet aanwezig	00221-001	B
Niet numeriek	00221-002	B
Niet in het toegelaten domein	00221-008	B

Smals ICT for society



Summary

Summary

- Validating DmfA declarations as RDF is feasible
  - The generation of SHACL must be "aligned" with the generation of XSD
  - Sharing SHACL rules, like sharing the XSD, would allow third parties to validate their data locally.
  -
- SHACL is even more expressive than XSD
-
- SHACL and the vocabulary is compatible with JSON-LD (a JSON serialization of RDF), which is only in more contemporary Web environments (cfr. Flanders)
-
- SHACL can even validate declarations against a KG (future?)
-