Simple SHACL Intro

02 March 2021

MITRE | SOLVING PROBLEMS FOR A SAFER WORLD™

What is SHACL?

- **Shapes Constraint Language (SHACL)** is a **World Wide Web Consortium (W3C)** specification for validating graph-based data against a set of conditions.

- Shape validation is done at the **semantic** RDF level regardless of datafile serialization format (JSON-LD, RDF-XML, Turtle, n-Triples, etc)
  - This means it is more capable and flexible than serialization-based syntactic validation forms such as JSON Schema, XML Schema, etc

- Validation is done by comparing a “shapes” graph defined with SHACL to a datagraph (data file)

What is the “shapes” graph?

- **Shapes graph** is made up of **nodeshapes** and **property shapes**
- **Nodeshapes** apply to defined **RDF/OWL Classes** and define the “shape” that instances of that Class should conform to
- **Property shapes** apply to defined **RDF/OWL Properties** and are themselves asserted as properties of **Nodeshapes**
- **Property shapes** define **constraints** for a given property on a given Nodeshape (type, cardinality, value, value patterns, etc)
- Can also specify logical combinations of constraints
- Extensible for more complex conditions using SPARQL

Standalone shapes content

```
core:ElementShape
   a sh:NodeShape ;
   sh:targetClass core:Element ;
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
            sh:minCount "1" ;
            sh:path core:id],
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
             sh:path core:name],
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
             sh:path core:summary] .
```

Embedded shapes content within RDF/OWL

```
core:Element
   a owl:Class, sh:NodeShape ;
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
            sh:minCount "1" ;
            sh:path core:id],
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
             sh:path core:name],
   sh:property [ sh:datatype xsd:string ;
            sh:maxCount "1" ;
             sh:path core:summary] .
```

Yields cleaner, more expressive and correct RDF/OWL

- **More clear, efficient and flexible manner to associate properties with classes than OWL property restrictions**
- **Also avoids the common misuse of domain assertions on properties to associate properties with classes**
  - Domain assertions are actually intended to infer classification of objects based on the presence of particular properties
- **Validation can be performed using any SHACL validation engine such as pySHACL, rdf4j, the Interoperability Test Bed, GraphDB, etc)**