The PID Info Type Use Case
RDA P7 - Data Typing Working Group

Ulrich Schwardmann

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)

Am Fassberg, 37077 Göttingen
ulrich.schwardmann [at] gwdg.de

01 March 2016, Tokyo
PID Information Type (PIT)

- is very basic metadata of DO
- used also/primarily for automated processes
- needs as much disambiguity as possible
- use *context free grammar* to define the content of a PIT
- JSON schema is an obvious approach here
DTR

- gives a definition of a type in JSON
- provides schemas defining
  - how these definitions can be made, and
  - which fields need to be filled in.
- But the content of a type instance has to be validated
- and the **schema for a type itself is not defined** and cannot be derived in general
  - usual schema generators, that define schemas for gives json files derive schemas for how to define a type.
  - This is the *upward* schema, which is already known.
Schemas for type instances need to be automatically derived from the type definition.

As an adaption and expansion for DTR one needs definitions for types that are usable to derive schemas for these types and these schemas need to have enough descriptive power: the derivable schemas should cover a considerable part of the JSON-schema possibilities. Test case: is it possible to define JSON as a type?

DTR adaption: give schemas for the definition of types that are suitable to derive a schema from a type definition for instance validation.
Suggestion for DTR schemas

- use **Basic Types**, that have a pattern description by a regular expression
  - it provides
  - **identifier**
  - **regular expression**
  - beside *name, description, standards, provenance*

- and **Derived Types**, that have
  - **identifier**
  - and **properties** defined just by
    - **Basic Types** referenced by their identifier
    - **Derived Types** referenced by their identifier
examples
examples
{"$schema": "http://json-schema.org/draft-04/schema#",
"definitions": {
  "21.T11148_a045f55e2a7fc9d60a5b": {
    "description": "date-time",
    "type": "string",
    "pattern": "^([0-9]{4})-([0-9]{1,2})-([0-9]{1,2})T([0-9]{1,2})\([0-9]{1,2}\)\([0-9]{1,2}\)\([0-9]{6}\)\([0-9]{6}\)$",
  }
},
"type": "object",
"properties": {
  "time-period@21.T11148_a77cd6959b4fff9a9c50": {
    "description": "time-period",
    "properties": {
      "begin-time": {
        "$ref": "#/definitions/21.T11148_a045f55e2a7fc9d60a5b"
      },
      "end-time": {
        "$ref": "#/definitions/21.T11148_a045f55e2a7fc9d60a5b"
      }
    }
  }
},
"required": ["time-period@21.T11148_a77cd6959b4fff9a9c50"]}